Arab World English Journal

INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

AWEJ Volume.16 Number.16 December 2025

www.awej.org

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, Decemper 2025, Pp.-1-2

Contents

Article Titles & Authors	Pages
Contents	1-2
The Role of Integrating AI Tools with Flipped Learning Strategy to Improve Students'English Literature Competence Nasra Ihmaid Jadwe, Hazim Al Dilaimy, Arif Abid Sayil, Mohanad Hazim Hakkoosh & Ali Sabah Jameel	3-18
ESL Undergraduate Students' Perception of the Usefulness of Using Games in Enhancing Grammar Accuracy, Learning Engagement, and Motivation Siti Fatimah Abd. Rahman, Nur Fadhilah Husna Mohd Ramlee, Nurul 'Aqiilah Mohamad Kamal & Qurratull Aini Ahmad Zukhi	19-35
A1 Level Vocabulary Enhancement Schemes for Writing Development Among Foundation Level 1 and Level 2 Students Yasmin Banu Abdul Majeed & Firos Kalarikkal	36-63
The Use of Smartphone Apps for Informal Language Learning Among EFL Students: Focusing on Students' Choices, Purposes, and Perceptions Maha Bader Alghasab	64-89
Applying the File-Card Model in Teaching and Learning English Speaking Skills for English-Majored Students Huynh Thi Bich Phuong, Tran Thanh Du & Nguyen Huu The	90-106
An Investigation of Writing Errors Made by Saudi EFL Learners Rawan Ahmed Alsoweed, Abeer Shujaa Alharbi & Fahad Abdulaziz Alqusumi	107-123
On the Integration of Excellent Traditional Chinese Culture into Public English Teaching in Higher Vocational Colleges in China Xia Peng, Xinjian Hu, Caiyun Cheng & Zhengjun Yang	124-147
Engagement in Thesis Writing: An Appraisal Analysis of EFL Student Writers' Evaluation of Prior Academic Works Lien-Huong Vo, Long Viet Le & Mai-Duyen Pham	148-164
A Ten-Year Systematic Literature Review of English E-Learning Education (2014–2024) Mengyao Wu, Helmi Norman, Nor Hafizah Adnan & Yuanling Hu	165-184
Unlocking the Potential of AI-Enhanced Mixed Reality: Enhancing Technical English Proficiency for Engineering Students in the Gulf Region Shafiq ur Rehman, Rabia Khatoon & Intakhab Alam Khan	185-195
Evaluative Adjectives with Positive Connotations in Modern English Advertising Language Eva Leláková & Nora Kupčíková	196-209
A Retrospective Narrative Inquiry into Vietnamese Learners' Learning Experiences in General English Classrooms Bui Thi Thuc Quyen, Phan Nhat Hao, Nguyen Thi Phuong Hong, Pham Xuan Binh & Huynh Anh Huy	210-223
Conversational Analysis of Learner – AI Chatbot Interactions in Developing Spoken Fluency Anna Matiienko-Silnytska, Nana Mikava, Iryna Savranchuk, Neonila Tkhor & Hanna Poliakova	224-238
From Problems to Progress: Evaluating Dalton Plan-Based Instruction in Chinese College English Reading Tianxiu Qi, Intan Safinas Mohd Ariff Albakri & Mazlin Mohamed Mokhtar	239-256
From Silent to Speaking: Reducing Speaking Anxiety in Low-Proficiency ESL Learners Through Game-Based Learning Mafarhanatul Akmal Ahmad Kamal & Mohd Fadhli Shah Khaidzir	257-272

Contents

Teaching ESP Vocabulary with the Application of Corpus-Based Materials at Universities: A Rapid Review of Literature Marta Lacková	273-288
Exploring Motivational Factors for Learning General and Medical English Among Early-Career Doctors in Ho Chi Minh City: A Mixed-Methods Study Nghi, Tran Tin & Anh, Trang Hoang Nguyen	289-316
The Synergistic Effect of Cognitive and Emotional Intelligence in Foreign Language Learning: Concepts and Pedagogical Perspectives M. Pınar BABANOĞLU, Reyhan AĞÇAM, M. Fatih ADIGÜZEL, Tuçe ÖZTÜRK KARATAŞ & Esin DÜNDAR	317-329
Cross-cultural Pragmatics and Code-switching in Multilingual EFL Classrooms Valeriia Smaglii, Svitlana Yukhymets, Yevheniia Kornielaieva & Inna Kivenko	330-342
Mobile-Assisted Language Learning in Tertiary EFL Education: A Decade of Systematic Evidence (2015–2025) Li Chengli, Harwati Hashim & Supyan Hussin	343-377
Half a Century of Research on Strategies for Vocabulary Learning: A Scientometric Analysis Xuan Wang, Xinyi Shen & Feiyu Liu	378-399
Enhancing EFL Students' Listening Skills with Songs and Films Lisjeta Thaqi-Jashari & Rina Sefiu	400-416
A Comparative Study on the Washback Effects of IELTS and CET-6 Listening Tests: Evidence from Test-Takers Yu Jiashun, Boon Yih Mah & Haida Umiera Hashim	417-437
Negotiating Writer Identity: A Comparative Study of Jordanian Undergraduate Students of English in Jordan and America Amani Salman	438-453
Model-Text Feedback in Large L2 Classes: Student Perceptions from a Bahraini Mixed-Methods Study Hasan Shaikh Husain & Seriaznita Mat Said	454-478
AI Tools in Academic Research: Perceptions, Adoption, and Challenges among Moroccan Student Researchers Lamiae Mamoune, Safae Kharmach, Abderrahim El Karfa & Mohammed Bennis	479-492
Autonomous Digital Language Learning Strategies of EFL Students Beyond the Classroom Hanadi A. Khadawardi	493-514

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp.3-18
DOI: https://dx.doi.org/10.24093/awej/vol16no4.1

The Role of Integrating AI Tools with Flipped Learning Strategy to Improve Students' English Literature Competence

Nasra Ihmaid Jadwe
Dean of the College of Education for Women, University of Anbar, Iraq
Email: nasra.jadwe@uoanbar.edu.iq

Hazim Al Dilaimy
University of Almaarif. Anbar, Iraq.
Email: hazimaldilaimy@uoa.edu.iq

Arif Abid Sayil
Dean of the College of Arts, University of Anbar, Iraq
Email: dr.arif.abd@uoanbar.edu.iq

Mohanad Hazim Hakkoosh
University of Anbar, Ramadi, Iraq
Email: mohanad85@uoanbar.edu.iqa

Ali Sabah Jameel
Department of English, College of Arts, University of Anbar, Iraq
Corresponding Author: alisabah40@uoanbar.edu.iq

Received: 08/21/2025 Accepted: 10/10/2025 Published:12/10/2025

Abstract
The significance of this study is to provide the students with learning strategies that enable them to comprehend a significant amount of literature modules efficiently, saving time and effort. This study attempts to bridge a gap in using two teaching methods by integrating artificial intelligence tools with the Flipped Learning strategy to improve teaching literature modules. The purpose of this study is to provide students with exposure to a broader range of English literature modules by integrating artificial intelligence (AI) tools with a flipped learning strategy, and to increase students' knowledge of the English literature modules. A quasi-experimental design was adopted with 27 English as a foreign language (EFL) students from the fourth grade in the Department of English, College of Arts, University of Anbar. The researchers collected data through a training program that included courses in novels, drama, poetry, short stories, and literary criticism, as well as a pre-test, post-test, and diagnostic test for the experimental group. The researchers verified the validity and the reliability of the instruments. The results revealed that there is a statistically significant difference in the achievement of students in the experimental group and those in the control group, attributed to the AI and the flipped learning strategy for the students in the experimental group. In addition, the results revealed statistically significant differences in students' achievement across the literary modules, with the novel module ranking first. This result is attributed to the students benefiting from using AI tools and the flipped learning strategy to study six novels within nine weeks. The students benefited greatly from applying the flipped learning strategy and artificial intelligence tools, which improved their knowledge in English literature modules and enabled them to study and review a greater number of literary works within the semester.

Keywords: Artificial intelligence educational tools, English literature (drama, novel, poetry, short stories, and literature criticism), fourth-grade students, flipped learning strategy

How to cite: Jadwe, N. I., Al Dilaimy, H., Sayil, A. A., Hakkoosh, M.H., & Jameel, A. S. (2025). The Role of Integrating AI Tools with Flipped Learning Strategy to Improve Students' English Literature Competence. *Arab World English Journal, 16*(4):3-18. https://dx.doi.org/10.24093/awej/vol16no4.1.

Introduction

Teaching English literature modules such as drama, poetry, novels, literary criticism, and short stories requires significant time and effort to cover the modules (Jameel, 2023). The nature of teaching English literature modules at Iraqi universities in the Department of English is based on the fact that the teacher is the center of the teaching and learning process and the only person with knowledge of the modules (Jameel et al., 2024). Therefore, the number of literary modules taught during the academic course ranges from one to two modules, for example, one or two novels or one or two plays, and perhaps the third one is just for general information. This is common among all English language departments. Herein lies the gap, which aims to narrow it and find academic solutions to achieve the objectives of teaching literary modules in English language departments, and to work on exposing students to a reasonable number of literature modules, which in turn contributes to developing students' language and deepening their understanding of the English language culture. The goal of the study is to expose students in one semester to at least six novels instead of studying two novels, five plays instead of two plays, more than 20 poems instead of six to seven poems, and studying more than 15 short stories instead of five to six stories. It is also essential for the students to understand most theories of literary criticism, rather than limiting themselves to only a few, because of their language level as English as a Foreign Language (EFL) learners and the ideological complexity of the theories. Moreover, it is time to link these theories to the literary works that the students study. The researchers seek to achieve these goals by using artificial intelligence (AI) tools as free, open electronic educational resources. To organize student learning, the researchers believe that it must be based on a theoretical framework for the teaching and learning process to be practical and grounded in scientific theory. Therefore, the researchers proposed integrating artificial intelligence tools with the Flipped Learning strategy. Students will receive a plan with the learning objectives, enabling them to prepare the material (drama, poetry, or a novel) according to the specified steps, such as definition, core content, and characteristics.

The current study aims to explore the impact of teaching fourth-year students literature modules using artificial intelligence tools with the Flipped Learning strategy on their achievement. It also seeks to examine the effect of the teaching strategy methods on students' achievement in the novels, drama, poetry, literary criticism, and short stories modules.

To achieve the study objectives, the researchers formulated the following hypotheses:

1. There was a statistically significant difference at the significance level of 0.05 between the achievement of students in the experimental group and those in the control group, attributed to the teaching methods.
2. There were statistically significant differences at the 0.05 level in students' achievement across the literary modules: novel, drama, poetry, short story, and literary criticism, attributed to the integration of artificial intelligence tools with the flipped learning strategy.

The English language in Iraq is considered a foreign language, and its learning takes place as a subject in schools during the primary, intermediate, and secondary stages (Al-Khayyat, 2015). Then, for students who wish to learn English as a specialization, this is done by enrolling in one of the English Departments in the College of Education, the College of Arts, or the College of Languages (Habeeb & Jameel, 2023). Learning English in the College of Arts focuses on literary courses such as drama, poetry, novels, literary criticism, and short stories (Anber & Jameel, 2020). These courses differ somewhat from the English language courses in the College of Education or the College of Languages (Jameel, 2022). Based on these facts and studies, this study adopted the Flipped Learning strategy, which bases on the Constructivist Learning Theory, where students

acquire knowledge through constructs information. The learning process takes place through interaction among them. Additionally, the researchers integrated artificial intelligence tools as a free, open educational resource.

Literature Review

The flipped classroom has a positive effect on learning the English language. The attitudes of students who study the English language using the flipped learning strategy are positive because it is characterized by ease of teaching, control of information, saving lecture time, and being more engaged in knowledge than the passive reception method inside the classroom (Lee & Wallace, 2018). However, the flipped learning strategy provides excellent opportunities to practice English, whether through written communication or direct conversation and group interaction to obtain information (Jadwe, 2024). Not all sources that students rely on are electronic; knowledge can be obtained from nearby individuals, language learning centers, or through communication with others (Shawaqfeh et al., 2024).

Fisher, et al. (2024) investigated the studies from 2015 to 2021that conducted to reveal the effect, role and the significant of using flipped learning strategy (flipped classroom), they found that the participants in these studies ranged between five to 132 and most studies used quasi-experimental design and the most common tools were pre-test and post-test to collect the data, Moreover, they found that flipped learning enhanced idiomatic knowledge, productive performance, increased students' motivation, and raised students' higher-order thinking skills. The study also found that the impact of flipped learning on learning English language skills remains largely unproven.

The use of flipped learning strategies has been widely adopted by numerous researchers and prestigious journals, as noted by Arslan (2020), who examined the benefits and challenges of flipped learning in English language learning through a systematic review of several studies in this field. Seventy-eight applied studies published in reputable academic journals and publishing houses were selected. The results of the study showed that the number of published studies on the use of flipped learning has been steadily increasing, given the importance and positive impact of this strategy on students' English language achievement and their motivation to learn. To implement flipped learning strategy effectively, Lin et al. (2023) emphasized the use of technology and artificial intelligence tools to enhance the learning process, promote autonomous learning and self-regulation, support a student-centred classroom, and foster lifelong learning.

To achieve the objectives of the current study and to support the results with studies that may agree or disagree with the findings, and for the purpose of consolidating and reinforcing the results, the researchers decided to identify studies that investigated the effect of artificial intelligence tools (ChatGPT, DeepSeek, and Gemini, DeepL) on English language learning during the year 2025, provided that the study sample consisted of university students majoring in English language. Other studies did not involve English language major, after surveying studies exclusively through the Scopus database. To clearly demonstrate the differences or similarities between these studies and the current study during the year 2025 with high accuracy, the following variables were extracted: the researcher(s)' name, the artificial intelligence tool, the purpose of using the tool, and the study sample (university level only). Table 1 illustrates this:

Table 1. *Studies conducted to reveal the effect of AI on students' language achievements*

Author	AI Tool	Aims of Using AI	Participants	Database
Al Nafjan, et al. (2024)	AI writing tools	To determine how AI affects the linguistic features of student writing	60 university students	AWEJ
Cotelli Kureth, et al. (2025)	ChatGPT, DeepL, online dictionaries	To develop writing assignments.	EFL students from two Swiss universities	Scopus
Elmotri, et al. (2025)	AI-driven Automated Written Corrective Feedback	To investigate how AI-written corrective feedback facilitates EFL writing.	120 University students	AWEJ
Hwang, et al. (2025)	ChatGPT	Identify user profiles based on ChatGPT Usage Purposes and English Self-efficacy	400 Korean EFL university students	Scopus
Jaafarawi (2025)	ChatGPT, Task-Based Language Teaching Approach	Improving EFL students' communication skills	EFL university students	Scopus
Jaashan, et al. (2025)	LLM-GPT (Language Model GPT)	The effect of LLM-GPT in reducing spelling errors through automated feedback and practice	60 EFL University students	Scopus
Kohnke, et al. (2025)	Grammarly, Quillbot	Improving grammar, writing, vocabulary, and reading skills through personalized feedback and creative support	University First-year Students	Scopus
Neshkovska (2025)	ChatGPT	Developing writing performance through suggestions, corrections, and automated assistance	EFL university students	Scopus
Rovira-Esteva (2025)	ChatGPT	Enhance written production through guided activities	First-year university students	Scopus
Teng (2025)	ChatGPT	EFL students' perceptions of utilizing ChatGPT to seek feedback for writing	40 EFL university students	Scopus
Wardat & Akour (2025)	ChatGPT, ChatPDF	Enhance materials development, personalize learning tasks, and support instructor efficiency.	87 EFL instructors at Jordanian universities	Scopus
Yan (2025)	ChatGPT	Supporting writing feedback.	EFL University students	Scopus
Yu, et al. (2025)	ChatGPT	Investigating the effect of AI in multilingual teaching (educational interaction, student engagement, personalized learning recommendations).	100 international students at a university in Western China	Scopus
Zakaria, et al. (2025)	ChatGPT	To examine how AI affects EEL university students' critical thinking and writing skills	15 articles published between 2023 and 2025	AWEJ
Zakarneh, et al. (2025)	ChatGPT	To support learners' engagement, enhance language competence, and	301 English language University learners	Scopus

		increase students' positive attitudes toward usability.		
The Current study: Jameel, et al.	ChatGPT and Flipped learning Strategy	Improving EFL university students' English literature knowledge and increasing the amount of literature materials.	27 EFL university students at the Department of English, University of Anbar.	Submitted to Arab World English Journal (AWEJ)

To support the previously studies for the year 2025 related to the role of artificial intelligence tools and by briefly returning to the year 2022, the year in which the use of artificial intelligence tools became widespread, Deng et al. (2025) conducted a survey study that included 69 academic articles published between 2022 and 2024 to explore the effect of ChatGPT on students' achievement and the development of the English language performance. The results of the study revealed that most previous research highlighted the positive impact of ChatGPT on university students. Specifically, ChatGPT enhanced their English proficiency, increase motivation to learn, stimulate interest in critical thinking, and reduce both time and effort.

In the same context, Deep et al. (2025) conducted a study that included 29 articles on the role and impact of artificial intelligence tools, published between 2023 and 2025. The study found that most of the studies confirmed the effectiveness of AI tools in improving students' proficiency in English, enhancing the four skills, and linguistic techniques such as grammar, spelling, vocabulary, translation, language functions, student interaction, and providing open learning resources.

Al-Zubaidi et al. (2024) found that university student participants in Morocco demonstrated that ChatGPT is an effective learning tool; they also confirmed that this tool has both positive and negative effects. The study also found that most participants risk using ChatGPT spontaneously without providing feedback or guidance from a supervisor. Therefore, participants expressed confusion about the use of ChatGPT in academic fields and how to use it.

Mali (2025) also conducted a survey study that included 32 articles published during 2024 and the beginning of the first quarter of 2025 in international reputable database journals with a good impact factor, to reveal the mechanism of how university students majoring in English use AI tools in writing skills. The results of the survey indicated that the ethical use of writing is essential when employing AI tools, and that these tools contribute to improving learners' creative writing and enhancing students' ability to construct sentences. The integration of artificial intelligence tools in English language learning was emphasized by Fan et al. (2024), who confirmed the necessity of identifying the challenges of AI tools when integrated into English language learning, particularly in assisting the design of lesson plans, designing supporting curricula, and ensuring equitable access to the classroom for all students through the use of AI tools. Apriani et al. (2025) conducted a study that revealed the impact of ChatGPT on improving writing skills, finding that students' use of the tool helped them formulate ideas and express them in an organized, coherent, and academically clear manner.

The study by Jaafarawi (2025) presented the integration of artificial intelligence tools in English language learning and evaluated their effectiveness. Who found that students' use of AI tools must be under the control and supervision of the instructor to ensure students do not fall into concerns of academic integrity. When AI tools were integrated with the Task-Based Language Teaching strategy, they played a positive and effective role in language learning from the students' perspective. However, the students also required the instructor's support in using these tools.

Through literature reviews, the study of Enamorado (2025) found that it was a unique one that integrated AI with the flipped strategy, as it examined students' learning of English for tourism purposes using ChatGPT as a source of information and through the flipped classroom strategy. This study, in its general tools, is similar to the current research in terms of using artificial intelligence tools with the flipped learning strategy. The study found that ChatGPT has significant importance in generating authentic and realistic materials that helped students learn terminology and dialogues related to tourism in the English language.

Method

To achieve the aims of the study and prove the hypotheses, the participants, the instruments, and the procedures of the study are as follows:

The Participants

The study adopted a quasi-experimental design with 27 male and female students from the fourth grade in the Department of English, College of Arts, University of Anbar. The total number of fourth-grade students was 33, from which six were selected as the pilot sample. The remaining participants were divided into two groups, a control group and an experimental group, using the random sampling method, by using a list of students' names, transferred the first name to the experimental group, the second to the control group, and so on. After every four names, the fifth was assigned to the pilot sample. Table 2 shows the distribution of participants.

Table 2**.** *Participants of the study*

Groups	Male	Female	Total
Experimental	4	9	13
Control	3	11	14
Total	7	20	27

The Instruments

To collect data for the study, the researchers prepared two tools. The first was a training program based on integrating artificial intelligence tools and a flipped learning strategy for teaching modules in novel, drama, poetry, short stories, and literary criticism. Students in the experimental group were taught according to the proposed program, while students in the control group were taught using modern lecture strategies and a communicative approach. In addition, the researchers conducted a pre-test to ensure equivalence between the groups, a post-test to measure the impact of the teaching method on the control and experimental groups, and a diagnostic test to examine the effect of the learning methods on the outcomes of each literature module.

Face and Content Validity of the Posttest

To ensure face validity, the researcher distributed the posttest to jury members who specialise in English language, applied linguistics, methodology, and e-learning. The researchers request from the members to evaluate the test. The jury members had presented some modifications and recommendations, which were taken into consideration.

To verify the content validity, six male and female students were asked to respond to the posttest questions. After conducting the pilot study, the researchers found that the posttest questions were clear and precise, and that the average exam time ranged between 45 and 50 minutes.

The Reliability of the Posttest

The researchers calculated the reliability of the posttest by distributing the tests to a pilot sample consisting of six students. After 12 days, the researchers distributed the same test to the same pilot sample under the same conditions. The researchers used a Pearson Correlation equation, and the internal consistency reliability value was 0.87, indicating a high reliability.

The Procedures

To conduct the study, the researchers identified the problem and selected the participants from the fourth grade from the Department of English, given that students in this stage have a high level of English language background, and that all literature modules are compulsory. After identifying the participants and dividing them into two groups (control and experimental), the literature and the previous related studies of the current year, 2025, were reviewed. The researchers agreed to use artificial intelligence tools such as ChatGPT, Gemini, and DeepSeek since they are readily available, accessible, and free, along with the flipped learning strategy as classification criteria. Then, six Novels, six Dramas, 15 Poems, 20 Short stories, and 10 literary criticism theories and schools were selected to be the materials of the students of the experimental group.

The goal of using artificial intelligence tools is to enable students to obtain information about literature (according to the proposed lesson plans). At the same time, the flipped learning strategy helps students understand a large body of literature through self-study (according to the lesson plan). After identifying the study tools, the researchers compiled a pre-test, a post-test, and a diagnostic test to collect the data. The researchers, each according to their specialization, then compiled lesson plans for the students in the experimental group. These plans consisted of a general structure for the novel, drama, poetry, and short story modules. The plan for the literary criticism module differed slightly from these, given the module's characteristics. The lesson plan included that students extracted the following for the Novel module: plot (events and backgrounds), characters (character details, relationships between characters, purpose of relationships, symbol that the character re-presents, role of the character), place, time, historical and political background of the place and the role of the place in the novel, social and historical context, themes (which are very important as they contain the primary purpose of the novel), narrative style, analysis style, summarizing the chapters of the novel, understanding the literary value of the novel, linking the novel as a plot to the historical contexts and social background of the writer.

The lesson plan for the experimental group students in the Drama module consists of: identify drama terminologies, the most popular plays, the most famous playwrights (different areas), the types of text (lessons, symbols, the progression of the characters, the theme of the play and the sub-themes), the characters and the relationships between them, understanding cultural background, linking the playwright's life with the plot, and focus on the main idea of the scene: "Who said it, the context, what is its lesson?"

The lesson plan for the Poem module is that the researchers asked the participants to: identify the theme of the poem, the background of the poet and thier social life and society, the era, the narrator, the nature of the poem "the tone," irony, symbols, similarity, paradox, personification, anaphora, metaphor, hyperbole, alliteration, stanza form, free verse, poetic devices, connect poetic themes with real life, simplify the poem, and discuss the portrayal of a theme or symbol(s).

Concerning the lesson plan of the "Short Story" module, the researchers asked the participants to: identify the plot, the characters and the relationships among them, the hero, the location "setting," the narrator, symbols, allegory, irony, imagery, writer attitudes and trends, the kind of conflict, the lessons, the message, and the writer's and cultural background.

Finally, the lesson plan of the "Literature Criticism" module, the researchers asked the participants to identify: criticism, textual interpretation, classical criticism, romantic criticism, Marxist criticism, feminist criticism, post-colonial criticism, catharsis, organic unity, ideological imitation, intertextuality, the most prominent critics and their theories starting from Plato to post-Derrida, historical, ethical, Marxist, and post-structuralist theories, a comprehensive understanding of the development of criticism from its inception until now, the common critical schools, linking theory to literature, understanding the most important opinions of critics, and analyzing them and linking them to literary works.

To ensure the equalization between the groups, a pre-test was prepared, and a post-test was prepared to measure differences between the two groups, if any. After implementing the experiment on the experimental group, the researchers faced several challenges, the most significant being students' anxiety about the number of modules they had to study and the diagnostic test, which was consisted of four questions per module and had a total score of 50 marks. Another challenge was how to train students in the experimental group on formulating questions appropriate for artificial intelligence tools to obtain accurate information. The nature of the questions posed by the students led to varying results, some of which were inaccurate and even vague. This depended on the skill of question formulation. Therefore, the students were trained on how to formulate questions for artificial intelligence tools, which resulted in accurate information supported by academic sources published in highly reputable international scholarly journals. The roles of the researchers were to divide the students within the class into groups, each with a leader. The lecturer (researcher) then posed the questions and provided feedback.

After completing the experiment, students in both groups were subjected to an achievement test, which included only the number of literary modules studied by the control group, as they had studied approximately a quarter of what the experimental group had learned. The students then took a test measuring their knowledge of the literary modules they had learned. The arithmetic mean of the scores was calculated, and the effect size was then calculated for which literary modules were most beneficial for understanding using the proposed strategies.

Results

The Results of the First Hypothesis

To verify the first hypothesis, which states that 'there is a statistically significant difference at the 0.05 significance level between the achievement of students in the experimental group and those in the control group, attributable to the teaching methods,' the researchers used the mean score, standard deviation, and t-values. Table 3 shows the results.

Table 3. *Participants' means scores, standard deviation, and t-value of posttest*

Groups	N.	Mean	St.	Df.	T-Values		Sig.
					Calculated	Tabulated	
Experimental	13	84.39	5.98	25	3.39	2.01	0.05
Control	14	71.12	6.28				

Table three shows that the mean score of the participants in the experimental group is 84.39 with standard deviation of 5.98, while the mean score of the participants in the control group is

71.12 with standard deviation of 6.28, and degree of freedom 25. Based on these results, the calculated t-value is 3.39, which is higher than the tabulated t-value, which is 2.01. This means that the first hypothesis is accepted, and there is a statistically significant difference at the significance level of 0.05 between the achievement of students in the experimental group and those in the control group, attributed to the AI and the flipped learning strategy for the students in the experimental group.

Results Related to the Second Hypothesis

To verify the second hypothesis, which states "There were statistically significant differences at the 0.05 level in students' achievement across the literary modules: novel, drama, poetry, and short story, and literature criticism attributed to the integrating artificial intelligence tools with the flipped learning strategy", a diagnostic test with Simulated Sectional Mean Scores and St. was used to collect the data. Table 4 shows the results.

Table 4**.** *Simulated sectional mean scores and St.d*

Modules	**Mean Score**	**Standard Deviation**
Novel	8.89	0.76
Short Stories	7.97	0.79
Literary Criticism	8.51	0.89
Poetry	8.12	1.01
Drama	7.92	1.07
Total	41.41	---

Table four shows that the highest score is in the Novel post-test questions with mean score 8.89 which ranked the first, in the second rank is the short-stories questions with mean score 7.97 and in the last rank is the drama with mean score 7.92, these results are derived from students' marks on the post-test, they get the highest marks on the questions of the novel. The overall achievement is 41.41, which is considered very high out of 50. These results revealed that the second hypothesis was accepted, showing statistically significant differences at the 0.05 level in students' achievement across the literary modules. The novel module ranked first, as students benefited from using AI tools and the flipped learning strategy to study six novels in depth over a nine-weeks.

Discussion

This study aims to benefit from the results and serve as one of the factors that bridge the gap between theory and practice, as well as within Iraqi literature, regarding the role of artificial intelligence tools and the flipped learning strategy. The following is an interpretation of the results of the first and second hypotheses.

The statistical results showed that participants who used artificial intelligence tools and the flipped learning strategy had a higher understanding of English (British and American) literature materials than students in the control group. Although students in the experimental group studied more literary materials than students in the control group, their answers were accurate, organized, and provided precise details. In contrast, the answers of the control group lacked precision, were characterized by weak language that was not convincingly literary, and were also brief. Accurate, diverse, evidence-based, and academically structured answers were provided to the students in the experimental group, who were trained to derive information from AI tools. The answers of the students in the experimental group were professionally formulated, resulting in rich, accurate, and well-organized responses, characterized by clarity, precise information, and a refined literary style.

Students' enhanced recall of this information is attributed to the positive impact of the flipped learning strategy, where each student prepared the literary material according to the study plan. This result is consistent with the results of Fisher, et al. (2024), who found that the flipped learning improves students' language performance and increases students' motivation, also consistent with the study of Arslan (2020), who investigated Seventy-eight applied studies conducted to reveal the role of the flipped learning strategy and found that it has a positive impact on students' English language achievement.

The lesson plan covered all the literary elements for each module, depending on the nature of the scholarly material. This is due to the inherent complexity of the subject matter. The nature, elements, and objectives of the Novel modules differ from those of poetry, drama, short stories, and literary criticism modules. Therefore, the researchers, over the semester of the experiment, provided students with diverse study plans. When students extracted literary elements from AI tools, they reformulated them in their own words. This finding is supported by Apriani et al. (2025), who measured the effect of using ChatGPT on students' academic writing, summarizing skills, and achievement. Therefore, learning was effective, the speed of completion was exact, and the teaching time- and effort-efficient. Self-learning aims to achieve precise and lifelong learning. This is supported by the study of Kohnke et al. (2025), which found that AI improved grammar and writing skills through personalized feedback and intrinsic motivation, leading to independent learning and self-regulation.

The researchers noted that students' preparation for literary subjects, based on lesson plans, was organized, precise, and detailed (metacognitive), consistent with the subject's background, and linked events in a coherent and understandable sequence. This finding is consistent with the results of Wardat and Akour (2025), who found that ChatGPT enhanced materials development, personalized learning tasks, and supported instructor efficiency. Furthermore, Yu et al. (2025) reported that ChatGPT has a positive effect on students' engagement and personalized learning. For example, students meticulously and clearly organized the relationships between the novel's characters, and the network of interconnected relationships was carefully planned from the beginning to the end of the plot. This precise explanation is lacking in lectures given by professors, as the learning process is influenced by numerous external variables that impact education. For example, when a professor explains a play such as Hamlet, s/he defines characters according to their appearance in the play, and their roles are based on the sequence of their appearances. Therefore, students' confusion over the importance of a character's role is dispersed throughout the play's events, requiring significant effort to capture the character's relationship with the other characters and their influence on the events.

However, information extraction must be conducted according to scientifically sound controls and plans to ensure accurate information is obtained. The role of the professor is to guide, develop, motivate, and provide feedback that enriches the learning process. Thus, students acquire raw information from AI tools as a free e-learning resource, as well as from the professor's experience, who disseminates, refines, and re-presents the info. In turn, students retain this knowledge for life, achieving effective learning. This finding supported by Stan, et al. (2025), who found that the use of AI tools supported by lecture's feedback, enhanced real-time practice, continuous learning, task organization, and learning management.

Moreover, the plot in a novel is straightforward and often at the forefront of the story. The characters' connections and relationships are straightforward. Furthermore, the symbols, setting, and time period in the novel often correspond to reality, being drawn from the author's context

and reflecting the social, historical, and political aspects. The narrative style also facilitates understanding the events, morals, and other messages of a novel. All of these factors helped students answer the post-test questions precisely, distinguishing their studies in the novel from other literary modules.

These characteristics also apply to the short story module, which exhibited the same characteristics, but in a reduced form. Therefore, students' scores were high on the short story questions, and the arithmetic mean for their answers came in second rank after the novel module.

The researchers believe this result is because the short story module features numerous stories, multiple characters, and numerous events. However, the students excelled in understanding the stories and the objectives of studying the short story module. This attributed to the role of artificial intelligence tools in presenting the scientific material in an interesting way, the method of studying the material using flipped learning, and the role of the professor in refining and crystallizing the scientific material, providing feedback, and providing students with material that distinguished by academic experience. This approach added spirit to learning English literature. It transformed it from a rigid subject into one that touches the soul of the learner adding the beauty of English literature, motivating students to be creative in it, and making it a spirit that inhabits the bodies and lifestyle of the learners, and a means that refreshes their feelings and motivates them to learn the English language enjoyably way. These results are consistent with those of Alnsour et al. (2025), who found that learning was most effective when learners engaged with an interesting learning method. They also reported that university students successfully used ChatGPT and achieved high scores in English performance. In addition, Jaafarawi (2025), found that AI improved communication skills due to the usefulness of AI in learning English modules.

Pedagogical Implications

Acquiring new information is not an easy task, regardless of the learner's intelligence level (Al-Khayat, 2015). There is no such thing as a lazy or diligent student (Ameen and Sayil, 2023). The rate of learning, the acquisition of new experience, and changes in learner behavior depend on many factors, most notably the level of attention, learning from reliable sources, and an engaging learning style (Yaseen, 2020). When students used the flipped learning strategy, it enabled them to focus on the material and to replay, review, and even summarize it in their own words. This significantly reduced anxiety about acquiring new information. Numerous studies (Habib & Jameel, 2023; Ibraheem et al., 2025; Jameel, 2022; Shukur & Sayil, 2020) have proven that lower anxiety levels are closely linked to higher achievement rates. When the teacher provided feedback, the learning process became more entrenched and sustainable, as confirmed by studies by Bobkina et al., 2025; Jadwe, 2020; and Ravandpour et al., 2022.

Conclusion

The results showed that learning using the flipped learning theory for all literary subjects (novel, drama, poetry, short story, and literary criticism) played a significant role in saving time and effort, while also enabling exposure to a large number of literary texts, which are essential needs that students pursuing a bachelor's degree in English literature must be equipped with. The students' self-learning efforts were supported by following the steps of the lesson plans and extracting information from artificial intelligence tools, after facing several difficulties in obtaining accurate and comprehensive information. This was addressed by training the students on how to formulate questions in an academic manner to obtain highly accurate, organized, and

well-structured information from AI tools. The phrasing of the question significantly affected the quality of the answer that AI tools provided, and students frequently encountered cases where the source provided by the AI tool was hypothetical.

The students benefited greatly from applying the flipped learning strategy and artificial intelligence tools, as the number of novels reached six, and the number of plays reached six as well. Regarding poetry, the students studied more than fifteen poems in depth, as well as more than twenty short stories with all their details, in addition to exploring most of the theories and schools of literary criticism from their inception to recent times.

In light of these results, the researchers recommend that university professors in English departments use artificial intelligence tools integrated with learning strategies so that learning occurs within a scientific, academic, and theoretical framework that adds scientific value and benefits to the learner. The information obtained by the students should be under the management and supervision of the professor, and the use of feedback should be emphasized, as it is the essence of the learning process.

Funding: This study is not funded.

Acknowledgments: We thank the journal's editorial board and the research reviewers for their feedback, their follow-up on the amendments, and their contribution to enriching the research.

Conflicts of Interest: The authors declare that there are no conflicts of interest regarding the publication of this paper.

Authenticity: This manuscript is an original work.

Artificial Intelligence Statement: AI and AI-assisted technologies were used for proofreading only.

About the authors

Prof. Nasra Ihmaid Jadwe (Ph.D). Dean of the College of Education for Women, University of Anbar, Iraq Specialized in Literature and Language. https://orcid.org/0000-0002-8551-1605

Prof. Hazim Al Dilaimy (Ph.D) Head of the Department of English, University of Almaarif. Iraq. Specialized in Linguistics and Phonetics. https://orcid.org/0009-0006-2001-6319

Prof. Arif Abid Sayil (Ph.D) Dean of the College of Arts, University of Anbar, Iraq. Specialized in Literature. https://orcid.org/0000-0003-1854-2644

Dr. Mohanad Hazim Hakkoosh University Presidential Office, University of Anbar, Iraq. Specialized in English literature. https://orcid.org/0000-0002-1628-8438

Prof. Ali Sabah Jameel (Ph.D). Department of English, College of Arts, University of Anbar, Iraq. Additionally, He is the Head of the Department of English at the College of Women, University of Anbar, Iraq. Specialized in English Language Curriculum and Instruction. https://orcid.org/0000-0003-3972-1406

References

Al-Khayyat, A.S.J. (2015). The Effect of CALL proposed program on university students' achievement in English. *Arab World English Journal (AWEJ), Special Issue on CALL No. 2*, p. 183-194. http://dx.doi.org/10.2139/ssrn.2843988

Alnsour, M. M., Qouzah, L., Aljamani, S., Alamoush, R. A., & AL-Omiri, M. K. (2025). AI in education: enhancing learning potential and addressing ethical considerations among academic staff—a cross-sectional study at the University of Jordan. *International Journal for Educational Integrity*, *21*(1), 16. https://doi.org/10.1007/s40979-025-00189-4

Al-Zubaidi, K., Jaafari, M., & Touzani, F. Z. (2024). Impact of ChatGPT on Academic Writing at Moroccan Universities. *Arab World English Journal (AWEJ)* Special Issue on ChatGPT, April 2024:4-25. DOI: https://dx.doi.org/10.24093/awej/ChatGPT.1

Ameen, N. T. M., & Sayil, A. A. (2023). Representations of Utopianism and Dystopia in Ecology: A Study of Subhi Fahmawi's Novels Akhenaten and Nefertiti al Canaanite and Alexandria 2050 through EcoCriticism. *Dirasat: Human and Social Sciences*, *50*(6), 93–106. https://doi.org/10.35516/hum.v50i6.7044

Anber, M., and Jameel, A. (2020). Measuring University of Anbar EFL Students' Awareness of Emoji Faces in WhatsApp and their Implementations. *Dirasat: Human and Social Sciences*, *47*(2). Retrieved from https://archives.ju.edu.jo/index.php/hum/article/view/107471

Apriani, E., Daulay, S. H., Aprilia, F., Marzuki, A. G., Warsah, I., & Supardan, D. (2025). A Mixed-Method Study on the Effectiveness of Using ChatGPT in Academic Writing and Students' Perceived Experiences. *Journal of Language and Education*, *11*(1 (41)), 26-45. https://doi.org/10.17323/jle.2025.17913

Arslan, A. (2020). A systematic review on flipped learning in teaching English as a foreign or second language. *Journal of Language and Linguistic Studies*, *16*(2), 775-797. https://doi.org/10.17263/jlls.759300

Bobkina, J., Baluyan, S., & Dominguez Romero, E. (2025). Tech-Enhanced Vocabulary Acquisition: Exploring the Use of Student-Created Video Learning Materials in the Tertiary-Level EFL (English as a Foreign Language) Flipped Classroom. *Education Sciences*, *15*(4). https:// doi.org/10.3390/educsci15040450

Cotelli Kureth, S., Paliot, E., & Zink, S. (2025). Fostering transparency: a critical introduction of generative AI in students' assignments. *Language Learning in Higher Education*, *15*(1), 63-85. https://doi.org/10.1515/cercles-2024-0105

Deep, P. D., Martirosyan, N., Ghosh, N., & Rahaman, M. S. (2025). ChatGPT in ESL higher education: Enhancing writing, engagement, and learning outcomes. *Information*, *16*(4), 316. https://doi.org/10.3390/info16040316

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, *227*, 105224. https://doi.org/10.1016/j.compedu.2024.105224

Elmotri, B., Harizi, R., Boujlida,A., Elyasa, Y.M., Garrouri, S., Amri, F., Malik, F. H., & Al-humari, M.A. (2025). The Impact of AI-Generated Feedback Explicitness. *Arab World English Journal*, *16*(1):384-402 DOI: https://dx.doi.org/10.24093/awej/vol16no1.24

Enamorado, M. Á. J. (2025). Development of Linguistic Competence in English for Specific Purposes Through ChatGPT: A Case Study. *Journal of Language and Education*, *11*(1 (41)), 85-100. https://doi.org/10.17323/jle.2025.23745

Fan, S., Peng, L., Wu, S., & Yu, X. (2024, November). ChatGPT Empowers Higher Education: —Research Topics Hotspots and Quantitative Visual Analysis. In *Proceedings of the 2024 3rd International Conference on Artificial Intelligence and Education* (pp. 38-45). https://doi.org/10.1145/3722237.3722245

Fisher, R., Tran, Q., & Verezub, E. (2024). Teaching English as a Foreign Language in Higher Education using flipped learning/flipped classrooms: a literature review. *Innovation in Language Learning and Teaching*, *18*(4), 332-351. https://doi.org/10.1080/17501229.2024.2302984

Habeeb, L. S., & Jameel, A. S. (2023). Investigating the translation strategies that English language students' utilized in translation diverse topics. *Dirasat: Human and Social Sciences*, *50*(6), 1-14. https://doi.org/10.35516/hum.v50i6.7029

Hwang, M., Lee, E., & Lee, H. K. (2025). Profiling and Understanding EFL University Students' Purposes for Using ChatGPT: A Latent Profile Analysis. *The Journal of AsiaTEFL*, *22*(1), 1-26. https://www.earticle.net/Article/A466104

Ibraheem, L. H., Mohammed, F. J., & Anber, M. M. (2025). Dialectal Differences in the Perceptive Intelligibility of Iraqi EFL Learners. An-Najah University Journal for Research - B (Humanities), 39(9), 753–764. https://doi.org/10.35552/0247.39.9.2448

Jaafarawi, N. (2025). Revolutionizing Language Education: A Task-Based, Ethical Approach to AI Integration With ChatGPT. In *Prompt Engineering and Generative AI Applications for Teaching and Learning* (pp. 305-328). IGI Global Scientific Publishing. DOI: 10.4018/979-8-3693-7332-3.ch019

Jaashan, H. M. S., & Alashabi, A. A. (2025). Using AI Large Language Model (LLM-ChatGPT) to Mitigate Spelling Errors of EFL Learners. In *Forum for Linguistic Studies* (Vol. 7, No. 3, pp. 328-339).

Jadwe, M. U. (2024). Transfictionality in Barbara Pym's Novels. *An-Najah University Journal for Research-B (Humanities)*, *39*(6), 457-466. DOI:10.35552/0247.39.6.2367

Jadwe, N. (2020). Feminizing Culture: A Reading of Inaam Kachachi's Novel Alnabiza. *Dirasat: Human and Social Sciences*, *47*(2). https://archives.ju.edu.jo/index.php/hum/article/view/107429

Jameel, A. S. (2023). The Role of self-monitoring strategies in developing students' speaking performances via Google Workspace. *Dirasat: Human and Social Sciences*, *50*(6), 25-34. https://doi.org/10.35516/hum.v50i6.7032

Jameel, A. S., Hakkoosh, M. H., Jassim, G. A., & Anber, M. M. (2024). The Innovative Methods of Contemporary Theatre: The Six Characters in Search of an Author as a Model. *Dirasat: Human and Social Sciences*, *51*(5), 119-129. https://doi.org/10.35516/hum.v51i5.10020

Jameel, A.S. (2022). The effects of story mapping and hortatory exposition techniques on students' compositions. *Education Research International*, 2022, 2022, 4300692. https://doi.org/10.1155/2022/4300692

Kohnke, L., Zou, D., & Su, F. (2025). Exploring the potential of GenAI for personalised English teaching: Learners' experiences and perceptions. *Computers and Education: Artificial Intelligence*, *8*, 100371. https://doi.org/10.1016/j.caeai.2025.100371

Lee, G., & Wallace, A. (2018). Flipped learning in the English as a foreign language classroom: Outcomes and perceptions. *TESOL quarterly*, *52*(1), 62-84. https://doi.org/10.1002/tesq.372

Lin, H. C., Hwang, G. J., Chou, K. R., & Tsai, C. K. (2023). Fostering complex professional skills with interactive simulation technology: A virtual reality-based flipped learning approach. *British Journal of Educational Technology*, *54*(2), 622-641. https://doi.org/10.1111/bjet.13268

Mali, Y. C. G. (2025). Exploring the use of ChatGPT in EFL/ESL writing classrooms: A systematic review. *Journal of Language and Education, 11*(2), 137-156. https://doi.org/10.17323/jle.2025.21793

Neshkovska, S. (2025). The Merits and Risks of AI-Assisted Academic Writing: Insights from Current Research. *ELOPE: English Language Overseas Perspectives and Enquiries*, *22*(1), 55-68. https://doi.org/10.4312/elope.22.1.55-68

Ravandpour, A. (2022). The Relationship between EFL Learners' Flipped Learning Readiness and their Learning Engagement, Critical thinking, and Autonomy: A Structural Equation Modelling Approach. *Journal of Language and Education, 8*(3), 97-105. https://doi.org/10.17323/jle.2022.12654

Rovira-Esteva, S. (2025). Teaching or Cheating: The Dark Side of ChatGPT as a Learning Companion for Beginner Chinese Students. *Journal of Technology and Chinese Language Teaching*, *16*(1), 52-84. http://www.tclt.us/journal/2025v16n1/roviraesteva.pdf

Shawaqfeh, A. T., Jameel, A. S., Al-adwan, L. A. Y., & Khasawneh, M. A. S. (2024). Interaction as a mechanism to enhance English language proficiency in the classroom. *Journal of Language Teaching and Research, 15*(1), 229-234. https://doi.org/10.17507/jltr.1501.25

Shukur, H., & Sayil, A. (2021). The Presence of Subject and Self in the Experience of the Iraqi Poet Adnan Al-Sayegh (The Effectiveness of Identification and the Disruption of Separation). *Dirasat: Human and Social Sciences*, *48*(1).

Stan, M. M., Dumitru, C., & Bucuroiu, F. (2025). Investigating teachers' attitude toward integration of ChatGPT in language teaching and learning in higher education. *Education and Information Technologies*, 1-18. https://doi.org/10.1007/s10639-025-13396-w

Teng, M. F. (2025). Metacognitive awareness and EFL learners' perceptions and experiences in utilising ChatGPT for writing feedback. *European Journal of Education*, *60*(1), e12811. https://doi.org/10.1111/ejed.12811

Wardat, S., & Akour, M. (2025). Teachers' use of generative AI in Jordanian universities: Practices and perceptions. *Forum for Linguistic Studies*, *7*(6), 90–102.

Yan, D. (2025). Feedback seeking abilities of L2 writers using ChatGPT: a mixed method multiple case study. *Kybernetes*, *54*(7), 3757-3781. https://doi.org/10.1108/K-09-2023-1933

Yaseen, K. (2020). Flashes of temporal construction in Nagib Mahfouz's (New Cairo). *Dirasat: Human and Social Sciences, 47*(2). https://archives.ju.edu.jo/index.php/hum/article/view/107427

Yu, H., Guo, Y., Yang, H., Zhang, W., & Dong, Y. (2025). Can ChatGPT revolutionize language learning? Unveiling the power of AI in multilingual education through user insights and pedagogical impact. *European Journal of Education*, *60*(1), e12749. https://doi.org/10.1111/ejed.12749

Zakaria, N.Y.K., Hashim, H., & Jamaludin, K.J. (2025). Exploring the Impact of AI on Critical Thinking Development in ESL: A Systematic Literature Review. *Arab World English Journal*, *16*, 330-347DOI: https://dx.doi.org/10.24093/awej/AI.19

Zakarneh, B. I., Aljabr, F., Al Said, N., & Jlassi, M. (2025). Assessing Pedagogical Strategies Integrating ChatGPT in English Language Teaching: A Structural Equation Modelling-Based Study. *World Journal of English Language*, *15*(3). https://doi.org/10.5430/wjel.v15n3p364

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp. 19-35
DOI: https://dx.doi.org/10.24093/awej/vol16no4.2

ESL Undergraduate Students' Perception of the Usefulness of Using Games in Enhancing Grammar Accuracy, Learning Engagement, and Motivation

Siti Fatimah Abd. Rahman
Department of Language and Literacy, Kulliyyah of Education, International Islamic University Malaysia, 53100, Selangor, Malaysia
Corresponding Author: sfarahman@iium.edu.my

Nur Fadhilah Husna Mohd Ramlee
Department of Language and Literacy, Kulliyyah of Education, International Islamic University Malaysia, 53100, Selangor, Malaysia
Email: husnaramlee@gmail.com

Nurul 'Aqiilah Mohamad Kamal
Department of Language and Literacy, Kulliyyah of Education, International Islamic University Malaysia, 53100, Selangor, Malaysia
Email: qilahkamal24@gmail.com

Qurratull Aini Ahmad Zukhi
Department of Language and Literacy, Kulliyyah of Education, International Islamic University Malaysia, 53100, Selangor, Malaysia
Email: qurratullaini.pg@gmail.com

Received:07/15/2025 Accepted:10/22/2025 Published: 12/10/2025

Abstract
This study explores the perceptions of ESL (English as a Second Language) undergraduate students regarding the effectiveness of incorporating games as a pedagogical tool to improve grammar accuracy, enhance learning engagement, and motivation. The research aims to investigate how the integration of games within the ESL classroom setting impacts students' language acquisition, particularly in terms of grammar proficiency. The methodology employed a quantitative approach, which focused on quantitative questionnaires. The quantitative aspect involved administering structured questionnaires to a sample of ESL undergraduate students, assessing their attitudes and opinions regarding the use of games for grammar learning. Overall, this study provides valuable insights into ESL undergraduate students' perceptions regarding the usefulness of integrating games as a means to enhance grammar accuracy, promote learning engagement, and bolster motivation. The findings underscore the potential of utilising gaming strategies in ESL classrooms as an effective and enjoyable approach to facilitate language learning and proficiency development. Further implications for pedagogical practices and recommendations for future research are discussed.

Keywords: English as a second language learners, game-based learning, grammar accuracy, learning engagement, Malay undergraduates, motivation, perception
How to cite: Rahman, S.F.A., Ramlee, N.F.H.M., Kamal, N.A.M., & Zukhi, Q.A.A. (2025). ESL Undergraduate Students' Perception of the Usefulness of Using Games in Enhancing Grammar Accuracy, Learning Engagement, and Motivation. *Arab World English Journal, 16*(4): 19-35. https://dx.doi.org/10.24093/awej/vol16no4.2

Introduction

Teaching and learning pedagogy has shifted from only the talk and chalk method to ten plus two teaching methods, the think-pair-share method and the game-based method in teaching. The diversity of teaching pedagogy started to emerge as time passed. Students in today's world need to be taught in different ways Rahman et al., (2025). The game-based teaching method is considered an effective learning method in class. Hadfield (1999) insists on including games in the language curriculum and stresses the importance of using them effectively rather than just as " an amusing activity for Friday afternoon or the end of the term". The traditional way of learning a language, such as the talk and chalk method, needs to be improved. Hence, the usage of games in learning a language, especially the English language, is highly recommended (Littlejohn, 2012). Furthermore, games can be tailored to different proficiency levels, allowing teachers to cater to the needs of individual students. Despite the potential benefits of using games in ESL classrooms, the effectiveness of this method has not been extensively studied.

Therefore, there is a research gap in this area, and there is a need for more research to understand the usefulness of games in enhancing grammar accuracy, learning engagement, and motivation among ESL students. Psychologists have long acknowledged the importance of play in cognitive development and learning. For instance, Piaget (1962) stated that play is essential to children's cognitive growth and changes as they do. According to Piaget, play becomes more abstract, symbolic, and social as kids progress through various developmental stages. By activating children's schemas in ways that enable them to move beyond their immediate reality, play is believed to aid children's cognitive development. For example, a child can pretend, or "act as if," an eraser is a car while fully knowing that it is not a car. This kind of play enables kids to keep track of several representations of the same thing, a skill necessary for the growth of symbolic thinking, one of early childhood's most important processes (DeLoache, 1987), one of the most significant developments of early childhood.

Hence, this study analyses the ESL undergraduate students' perception of the usefulness of using games in enhancing grammar accuracy, learning engagement, and motivation. By examining the value of games in boosting grammar accuracy, learning engagement, and motivation, this study has the potential to improve the English language learning outcomes for ESL undergraduate students. The results of this study can benefit the field of English language teaching and learning by giving English language teachers and curriculum developers insightful information. To increase the effectiveness of education, the study can specifically assist in the creation of efficient game-based learning activities for English language classes. Additionally, by giving ESL undergraduate students a more interesting and delightful learning experience, this study can directly help them. The motivation, interest, and focus of students can be enhanced through the use of games in the classroom, increasing learning engagement and improving learning outcomes. Furthermore, this study's findings can directly benefit ESL undergraduate students by providing them with more efficient and interesting language learning resources. The study's contribution to the development of effective and engaging language learning materials can have a positive impact on English language learning outcomes.

Literature Review

Game-based learning is a theoretical framework developed from the constructivist theory developed by Jean Piaget. It is a learning theory that suggests learners actively construct knowledge and understanding through their own experiences, interactions, and reflections.

Constructivism considers learning as "a process in which students actively build or construct new ideas or concepts" (Yusoff, 2013, p. 232). Hence, cognitivism in turn is "a response to the behaviourism theory of human thought as organisms that have goals in their behaviour and constantly interact with the environment" (Yusoff, 2013, p. 232). The theory of constructivism encourages students to become designers of their own ideas.

After the development of technology and games, especially e-games, the constructivist theory also developed and led to the emergence of game-based theory. Definitions of game-based learning mostly emphasise that it is a type of game play with specified academic achievements (Shaffer et al., 2005). "The use of play in an educational context and for purposes of learning and development is by no means a new phenomenon" (Plass et al., 2015, p. 258). The e-games that were made absorbed the concept of constructionist learning theory by making the student actively engaged with learning and making their own ideas (Yusoff, 2013). Reports stated that the youth spend much time a day playing digital games, which can be up to seven to 10 hours per week (Lenhart et al., 2008). In an article written by Plass et al. (2015), games are considered either environments that are motivating but likely to require excessive amounts of information to be processed by the learner (cognitive perspective) or, conversely, approaches that provide rich contextual information and interactions needed for learning in the 21st century (sociocultural perspective) (p. 259). This highlights the use of e-games not only for entertainment but also as an educational platform. "Knowledge development can occur whilst playing e-games" (Yusoff, 2013, p. 232). Moreover, Plass et al. (2015) also argued that games for entertainment can motivate learners to stay engaged over long periods through a series of game features that are of a motivational nature" (p. 260). Computer-mediated learning frameworks indeed present highly satisfying results in students' learning (Yusoff, 2013).

Grammar Accuracy and Game-Based Learning

Grammar accuracy is a critical component of language acquisition that assesses a student's ability to apply the proper grammatical rules, whether writing or speaking. It is a skill that calls for both comprehension of grammar ideas and accurate application of them in context. A student who has a high level of grammar correctness can use the proper language structures and construct grammatically accurate sentences. On the other hand, a student who struggles with basic grammatical principles may make mistakes frequently, which might impede the coherence and clarity of their message. To effectively communicate their message and reach their language learning goal, language learners must therefore strive for high levels of grammar accuracy.

There are several grammar games available for English-language learners, as well as for students in kindergarten through 12th grade. Some of these games are designed to appear as workbooks, while others collect a variety of games that can be played individually, collaboratively, or competitively. These games may be exploratory in nature, while others aim to determine correct and incorrect answers. Among the most effective grammar games are those inspired by JeopardyTM. The game board format provides a flexible means of asking different types of questions, promoting collaboration, healthy competition, and encouraging discussion. JeopardyTM style games can be used in both small and large classes, making them an excellent option for educators who are teaching classes of different sizes. This is especially important considering the growing trend of larger classroom sections.

Learning Engagement and Game-Based Learning

Learning engagement plays a vital role in ensuring students' understanding of learning a topic in class. Engagement in learning is important, as according to Mercer (2019), learning requires active involvement on the part of the learner, and action is the defining characteristic of learner engagement. The learning process without engagement is undoubtedly meaningful (Hiver et al., 2021). Philp and Susan (2016) claimed that engagement is the term frequently employed to talk broadly about learners' interest and participation in an activity. Students are considered to be engaged in learning when they are actively immersed in learning and experience the presence of motivation. According to Hiver et al. (2021), "engagement" refers to the degree to which a student is actively immersed in a learning task and the degree to which that student's physical and mental activity is goal-directed and purpose-driven.

Scholars in teaching and learning introduce the four dimensions of learning engagement, which are behavioural, cognitive, emotional, and social engagement. The main strand of work in the field suggests that engagement is manifested not only in its behavioural facet but also in demonstrations of action through the cognitive and social dimensions, as well as in students' emotional responses to learning tasks (Lambert et al., 2017). These four engagements stand together to create a quality learning engagement in class sessions. To effectively promote all four dimensions of learning engagement in class sessions, games have been recommended by Plass et al. (2015) as an effective approach.

Behavioural engagement refers to a student's active participation in learning activities, which includes asking and answering questions, completing assignments and attending classes regularly. This active participation is critical for effective learning, as it allows students to engage with the material, develop a deeper understanding of concepts and improve their skills and knowledge (Philp & Duchesne, 2016). Game-based learning is an effective way to enhance student engagement and increase behavioural engagement in the learning process (Plass et al., 2015). It is said that games are designed to be engaging and challenging; hence, by using games as a tool for learning, students can have an interactive and immersive learning experience, resulting in an active role in the learning activities (Yu et al., 2022).

The second aspect is cognitive engagement. Cognitive engagement refers to the learner's mental effort and mental activity in the learning process (Hiver et al., 2021). Students' mental effort and strategies are used to understand and process information, which includes the process of critical thinking, problem-solving, and elaboration. A study conducted by Brackenbury and Kopf (2022) demonstrated that game-based learning can effectively promote the development of critical thinking and problem-solving skills among students. The authors found that the challenging gameplay and interactive nature of games encouraged students to think critically and develop effective problem-solving strategies. This study managed to highlight the potential of game-based learning as a tool for promoting cognitive engagement and enhancing students' ability to think critically and solve problems in a variety of contexts.

The next aspect of engagement is emotional engagement. Emotional engagement defines students' affective response to learning. This process includes interest, curiosity, enjoyment, and enthusiasm. Affective engagement is related to learners' attitudes towards learning contexts, the members in that context, the learning tasks, and their own participation in learning (Skinner et al., 2009). Emotional engagement can be divided into positive and negative emotions. The affective response is always generated by the positive emotion; meanwhile, anxiety, boredom, frustration and anger demonstrate emotional disengagement or disaffection as a result of negative emotion

(Mercer, 2019). According to Plass et al. (2015), game-based learning can elicit learners' emotions, and empirical research indicates that positive emotion, such as affective response, increases the broad spectrum of cognitive abilities. Mao et al. (2022) researched the effects of game-based learning on students' critical thinking. The study revealed that game-based learning causes positive feelings in learners, which leads to an improvement in critical thinking.

The final dimension of learning engagement is social engagement, which refers to students' interactions and relationships with peers and teachers, including communication, collaboration, and feedback. Mercer (2019) claimed that the social part of engagement is the social kinds of activity and participation that are common in communities of language learning and use, such as interaction with interlocutors and the quality of such social interactions. Maintaining students' attention and engagement in the classroom requires open lines of communication between teachers and students. According to recent research (Philp & Duchesne, 2019), emotional involvement often follows social involvement. Learners' desire to listen to one another or focus on the teacher's discussion is an example of an empathetic discourse move, and it is related to phenomena like reciprocity, mutuality, and other prosocial displays of attachment (Philp & Duchesne, 2016). Plass et al. (2015) stated that there is a positive relationship between social engagement and game-based learning. According to them, game-based learning can facilitate social engagement by providing a fun and interactive platform for students to work together and learn from each other. A recent study on integrating a game–based app to enhance learners' engagement, motivation, and performance has been conducted (Chen, 2023). This study discovered that games play an important role not just in cognitive development, but also in social processes.

Motivation and Game-Based Learning

Motivation is important in students' learning because it drives learners to acquire new knowledge, skills, and attitudes. Motivation can be defined as a psychological construct that refers to the forces that drive an individual's behaviour and influence the direction, intensity, and persistence with which they pursue a specific goal (Berliner & Calfee, 2004). In order to cultivate intrinsic motivation in students, it is important to provide them with instructional materials that are engaging, relevant, and meaningful (Malone & Lepper, 2021). Malone's (1981) work suggested that game-based learning promotes intrinsic motivation in students. According to him, game-based learning provides an engaging and stimulating environment that incorporates all seven elements of intrinsic motivation. The seven elements of intrinsic motivation that are relevant to games are challenge, fantasy, curiosity, control, competition, cooperation, and recognition (Admiraal et al., 2011). Zhang and Yu (2022) conducted a meta-analysis study to investigate and compare the effects of gamification and game-based learning on learning achievement and motivation. The findings revealed that game-based learning has a positive influence on learners' intrinsic motivation.

Studies have also shown that game-based learning increases the self-determination of a student. It is said that game-based learning can help fulfil the innate psychological needs of autonomy, competence, and relatedness (Pratama, 2020). In the self-determination theory (SDT), Vansteenkiste and Ryan (2013) clarified that there are three innate psychological needs that help to facilitate intrinsic motivation, which are autonomy, competence, and relatedness. 'Autonomy' refers to an individual's ability to make choices and decisions about their own behaviour. 'Competence' refers to an individual's feeling of effectiveness and capability in adapting to the world, while 'relatedness' refers to an individual's sense of connection and attachment to others

(Vansteenkiste & Ryan, 2013 proved that these three basic psychological needs are important for human growth, development, and well-being. When these needs are met, individuals are more likely to be motivated and engaged in their activities and behaviours (Alamri, 2020).

Research Problems

Teachers always use the same repeated and boring cycle to teach grammar to students. According to Yaccob (2019), teachers need to be creative and innovative in teaching the English language. In this 21st-century era, the usage of games have been widely used in teaching and learning sessions in schools, as games are known as a source of relaxation and stimulation, and people especially enjoy playing games. While game-based learning has been widely adopted in schools, its effectiveness in promoting adult learners' understanding and retention of information in adult education, particularly in enhancing grammar accuracy, learning engagement, and motivation, remains unclear (Landers et al., 2019). Without a comprehensive understanding of adult learners' perception of game-based learning in enhancing grammar accuracy, learning engagement, and motivation, there is a risk that game-based learning approaches may not effectively meet the needs and expectations of adult learners (Balakrishna, 2023). The issue has arisen as the perception of adult learners in enhancing grammar accuracy, learning engagement, and motivation is not recorded. Hence, this study aims to explore ESL undergraduate students' perception of the usefulness of using games in enhancing grammar accuracy, learning engagement, and motivation.

Research Questions

This study aims to answer the following questions:

1. What are the perceptions of ESL undergraduate students regarding the usefulness of using games in enhancing their grammar accuracy?
2. How do ESL undergraduate students perceive the use of games in improving learning engagement?
3. What are the perceptions of ESL undergraduate students regarding the usefulness of using games in enhancing their motivation?

Methodology

This section explains the methodology used in gathering and analysing data through the method that is relevant to the research. The methodology is selected based on the suitability of the research design, participants, and data collection.

Research Design

A comprehensive survey is conducted to explore the ESL undergraduate students' perception of the usefulness of using games in enhancing their grammar accuracy, learning engagement, and motivation. A quantitative research design utilising a questionnaire survey is used. The questionnaire survey consists of Likert-scale items to assess students' perception of the usefulness of using games for enhancing grammar accuracy, learning engagement and motivation. The researcher is the one in charge of analysing the responses provided by the respondent.

Research Setting

The study is conducted in Malaysia, and all ESL undergraduate students are welcome to participate in the study. The Google Form access was closed as soon as the participation reached 34 students. The study is conducted in semester 1 2023/2024 before the semester ends. This study incorporates game-based learning activities with students' grammar accuracy, learning engagement, and motivation. Participants' confidentiality and privacy are maintained and secured, as the information collected will only be circulated in this study.

Participants and Sampling Procedure

The study includes 34 undergraduate students who are currently pursuing ESL courses. Generally, in Malaysia, everyone is considered an ESL learner because almost all tertiary education in Malaysia uses the English language as the medium of instruction. Hence, everyone is eligible to participate in the study as long as the student is at the undergraduate level. The questionnaire link will be distributed mainly through the WhatsApp application. Anyone who is an ESL undergraduate learner is free to participate in the study. Therefore, each student has an equal chance of being selected to participate in the study.

Research Instrument

This study uses a closed-ended questionnaire that was created using Google Forms. 30 questions are designed to measure the perceptions of ESL undergraduate students about the usefulness of using games in enhancing grammar accuracy, learning engagement, and motivation. The questionnaire uses a 5-point Likert scale to assess the degree of agreement or disagreement with statements related to the usefulness of games, grammar accuracy, learning engagement, and motivation. This 5-point Likert scale gives respondents more options to choose from: strongly agree, agree, neither agree nor disagree, disagree, and strongly disagree, and the participants are asked to select the response that best reflects their opinion.

The questionnaire is divided into several sections, each addressing a specific aspect of the study, such as the use of games in enhancing grammar accuracy, learning engagement, and motivation. Each section contains 10 statements that participants are asked to rate using the Likert scale. The questionnaire is distributed via WhatsApp to 34 ESL undergraduate students, and the data collected will be analysed using statistical methods to conclude about the perceptions of ESL undergraduate students about the usefulness of using games for language learning.

Data Collection

The data collection for this study involves administering a closed-ended questionnaire to 34 ESL undergraduate students. The questionnaire is distributed via WhatsApp and is accessible through a Google Forms link. To ensure the validity of the data, the questionnaire is piloted with a small group of five students who are not included in the final sample. The pilot study helps to identify any issues with the questionnaire's design, wording, or structure and allows for necessary revisions to be made. After piloting, the final questionnaire is sent to the 34 participants via WhatsApp. Participants are provided with a brief explanation of the study's purpose and are informed that their participation is voluntary and confidential. Participants are also assured that their responses will not affect their grades or academic standing in any way.

The data collection period lasts for two weeks, during which participants are given sufficient time to complete the questionnaire. The participants are reminded to complete the

questionnaire via WhatsApp every three days. Upon completion of the questionnaire, the data are exported from Google Forms to Microsoft Excel and SPSS for analysis. The data collection process is conducted with a high degree of transparency, rigour, and ethical considerations to ensure the validity and reliability of the data collected. The data collected is analysed to answer the research questions and objectives of the study.

Data Analysis

The data collected from the closed-ended questionnaire are analysed using descriptive statistics to answer the research questions and objectives of the study. Descriptive statistics such as frequency distributions and means are used to summarise the responses to each item of the questionnaire. The responses to each item of the questionnaire are presented in the form of tables to provide a visual representation of the data.

Results

According to the Cronbach's Alpha value and Range of Reliability in Table 1 and Table 2, the reliability of the data is highly reliable.

Table 1. *Reliability test results*

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardised Items	N of Items
.947	.942	30

Table 2. *Range of reliability and coefficient of Cronbach's Alpha*

Cronbach's Alpha	Internal Consistency
$\alpha \geq 0.9$	Excellent
$0.9 > \alpha \geq 0.8$	Good
$0.8 > \alpha \geq 0.7$	Acceptable
$0.7 > \alpha \geq 0.6$	Questionable
$0.6 > \alpha \geq 0.5$	Poor
$0.5 > \alpha$	Unacceptable

A Cronbach's alpha value of .947 indicates a high level of internal consistency and reliability for the scale used in this study. This suggests that the items in the scale measure the same underlying constructs of attitudes towards language games for improving grammar accuracy, learning engagement, and motivation.

Demographic Characteristics of Participants

The socio-demographic characteristics of the participants are presented in Table 3. Since the study population is university students, the average age of the participants is 23.5 years, with 38.2% male and 61.8% female. The participants were all Malay (100%).

Table 3. *Demographic characteristics of respondents*

Characteristics	Frequency (n)	Per cent (%)
	Gender	
Male	13	38.2%

Female	21	61.8%
	Race	
Malay	34	100%
	Age group (years)	
18 – 24	24	70.6%
25 – 34	10	29.4%
Mean Age (23.5)		

The socio-demographic characteristics of the participants are presented in Table 3. Since the study population are university students, the average age of the participants is 23.5 years; with 38.2% male and 61.8% female. The participants were all Malay (100%).

ESL Undergraduate Students' Perceptions of the Use of Games in Enhancing Grammar Accuracy

Table 4 below demonstrates the variables concerning language games for enhancing grammatical accuracy, which have mean values of more than four, showing a favourable attitude towards practical games. Statements like "I have played any language games to improve your grammar accuracy before?" (s1) and "I have seen improvement in my grammar accuracy after playing language games" (s5) both have mean values of around 4.3 and 4.2, indicating substantial agreement or a good perception.

Table 4. *Results for the usefulness of games in enhancing students' grammar accuracy in the survey*

Grammar accuracy item(s)	SDA (%)	D (%)	N (%)	A (%)	SA (%)	Mean
1. I have played many language games to improve my grammar accuracy before	0.0	2.9	14.7	52.9	29.4	4.38
2. I believe that language games can improve my grammar accuracy	0.0	2.9	5.9	50.0	41.2	4.32
3. I think language games are more fun than traditional grammar lessons	0.0	2.9	5.9	26.5	64.7	4.16
4. I think playing language games can help me identify my grammar mistakes more easily	0.0	2.9	14.7	47.1	35.3	4.26
5. I have seen improvement in my grammar accuracy after playing language games	0.0	5.9	11.8	47.1	35.3	4.24
6. I think playing language games can make learning grammar more enjoyable	0.0	2.9	2.9	29.4	64.7	4.21
7. I think language games should be used in combination with other methods, such as textbooks or online courses, to improve grammar accuracy	0.0	5.9	8.8	35.3	50.0	4.15
8. I think language games should be used excessively in the classroom to improve grammar accuracy	0.0	8.8	11.8	50.0	29.4	4.21
9. My peers have used games before as an alternative instructional material in the class	0.0	14.7	14.7	41.2	29.4	3.97
10. I think games can be beneficial for students to enhance their grammar	0.0	2.9	8.8	41.2	47.1	4.26

Table 4 shows the results of the usefulness of games in enhancing students' grammar accuracy. The table also portrays the percentage of responses by students according to the 5-point Likert scale (i.e., SA = Strongly Agree, A = Agree, N = Neutral, D = Disagree, and SDA = Strongly Disagree) along with the mean.

The results showed that about half of ESL learners agreed, and 41.2% were very agreeable, that language games help increase grammatical accuracy. According to these results, students considered language games to be a successful method of acquiring and maintaining grammar principles. This is consistent with earlier studies that have shown the effectiveness of game-based learning in enhancing linguistic abilities. As Vernon (2006) asserted, when young students play a game, they "focus on the activity and end up absorbing the language subconsciously." Similarly, according to (Naidu et al., 2013) ESL students who practise their grammar by playing games can make gradual improvements in their grammatical fluency over time.

ESL Undergraduate Students' Perceptions of the Use of Games in Enhancing Learning Engagement

The result in Table 5 showed that the usage of game-based learning in classroom settings enhances students' engagement in the teaching and learning process. The result below portrayed that the majority of the respondents strongly agree and agree with the usage of games in enhancing learning engagement.

Table 5. *Results for the usefulness of games in enhancing students' learning engagement in the survey*

Learning engagement item(s)	SDA (%)	D (%)	N (%)	A (%)	SA (%)	Mean
11. Games make it easier for me to understand and apply grammar rules	44.1	8.8	0.0	50.0	47.1	4.35
12. Games help me practise grammar in a fun and interactive way	0.0	2.9	5.9	41.2	50.0	4.29
13. Games enhance my ability to use grammar accurately in real-life situations	2.9	2.9	5.9	47.1	41.2	4.26
14. Games improve my overall understanding of grammar concepts	5.9	0.0	17.6	44.1	32.4	4.15
15. Games motivate me to learn and improve my grammar skills	0.0	2.9	14.7	38.2	44.1	4.18
16. I actively participate in grammar learning games	0.0	11.8	20.6	38.2	44.1	3.85
17. I find grammar-learning games interesting and enjoyable	0.0	5.9	8.8	44.1	41.2	4.24
18. Grammar learning games keep me focused and attentive	2.9	8.8	8.8	52.9	26.5	4.18
19. I feel motivated to perform well in grammar-learning games	0.0	8.8	0.0	44.1	47.1	4.26
20. Grammar learning games challenge me to think critically and apply grammar rules	0.0	5.9	11.8	38.2	44.1	4.15

Table 5. shows the result of questionnaires focused on student learning engagement. The table also portrays the percentage of responses by students according to the 5-point Likert scale (i.e., SA = Strongly Agree, A = Agree, N = Neutral, D = Disagree, and SDA = Strongly Disagree) along with the mean.

Game-based learning proves to be facilitated in these areas; Games make it easier for students to understand and apply the grammar rules (47.1%), games help students practise grammar in a fun and interactive way (50.0%), students feel motivated to perform well in

grammar-learning games (47.1%), grammar motivates students to learn and improve grammar skills (44.1%).

ESL Undergraduate Students' Perceptions of the Use of Games in Enhancing Students' Motivation

The result in Table 6 below showed that the usage of game-based learning in the classroom in enhancing students' motivation was significant. Table 6 shows that the majority of the respondents have positive perceptions of the use of games in enhancing their learning motivation.

Table 6. *Results for the usefulness of games in enhancing students' motivation in the survey*

Learning motivation item(s)	SDA (%)	D (%)	N (%)	A (%)	SA (%)	Mean
21. The game motivated me to progress and get better	2.9	5.9	11.9	47.1	32.4	4.15
22. I enjoyed participating in game-based learning sessions	2.9	5.9	17.6	29.4	44.1	3.91
23. I feel a sense of relatedness and connection with other classmates while engaging in a game-based learning activity	5.9	5.9	8.8	26.5	52.9	3.88
24. I believe that game-based learning positively impacts my real-life motivation and self-determination	2.9	5.9	14.7	38.2	38.2	4.03
25. Game enhances my sense of competence and mastery	2.9	5.9	14.7	32.4	44.1	3.97
26. I believe that using games can lower my anxiety towards learning English	2.9	8.8	14.7	38.2	35.3	3.97
27. I feel a sense of accomplishment when completing tasks in a gamified learning experience	2.9	5.9	11.8	32.4	47.1	4.00
28. Playing games allows me to experience a sense of autonomy and control over my actions and decisions	0.0	5.9	11.8	44.1	38.2	4.21
29. I feel more confident and active in the class with game-based learning	8.8	5.9	11.8	29.4	44.1	3.79
30. Game-based learning has made learning more enjoyable for me	5.9	8.8	8.8	35.3	41.2	3.91

Table 6 shows the results of questionnaires focused on student motivation. The table also portrays the percentage of responses by students according to the 5-point Likert scale (i.e., SA = Strongly Agree, A = Agree, N = Neutral, D = Disagree, and SDA = Strongly Disagree) along with the mean.

Game-based learning facilitated the areas such as: students feel connected to classmates through game-based learning (52.9%), gamified learning provides students with a sense of accomplishment (47.1%), and students become more engaged taking part in a game-based learning session (44.1%).

Discussion

In answering the first research question, what are the perception of ESL undergraduate students regarding the usefulness of using games in enhancing their grammar accuracy, it can be concluded that results showed that 47.1% of respondents agree with this statement, while 35.3% strongly agree. This suggests that language games may help students become better at spotting their own mistakes and fixing them. In statement five, the researchers were interested in investigating whether students improved their grammar accuracy after playing language games. According to the survey results, 35.3% of respondents gave an answer that they strongly agreed with, and 47.1% said that they agreed, indicating that students believe language games are good

for improving their grammar accuracy. Meanwhile, in statement six, this study investigated whether playing a language game increases interest in learning grammar. In this manner, 64.7% of respondents provide a strong agreement, while 29.4% provide an agreement. This indicates that students may become more actively involved and interested in grammar learning. These findings imply that language games have various benefits, including greater error identification and enhanced grammatical accuracy. Students understand the significance of engaging in game-based activities to help them along the path of language acquisition.

Moreover, playing language games for pleasure (s6), a large majority of respondents (64.7%) felt that grammar-learning games increase enjoyment. Students are more likely to be driven, interested, and eager to put in the time and effort to grasp grammar when they enjoy the process of learning. Similarly, Syafiqah Yaccob & Md Yunus, (2019) found that their investigation into the use of games in the instruction of grammar led them to the conclusion that "grammar games encourage, entertain, and promote fluency" (p. 28). Statement seven determines if students think language games should be combined with other approaches, such as textbooks or online courses, to enhance students' ability to correctly apply grammar rules. The majority of respondents favour using language games in the classroom (50% strongly agree, 35.3% agree). This supports the idea that games can be a helpful addition to more conventional educational resources by providing a fun and novel way to practise grammatical skills. In Statement 8, the researchers inquired as to whether or not students think grammatical concepts may be taught through the use of games. According to the data, 79.4% of respondents agreed and strongly agreed with the statement. This data showed that students preferred to be taught grammatical ideas through the use of games.

The next table discussing on wow do ESL undergraduate students perceive the use of games in improving learning engagement, a total of 30 students agree and strongly agree with the statement of (s11) “Games make it easier for me to understand and apply the grammar rules”. It concludes that 88.2% of respondents agree that game-based learning makes it easier for them to understand one topic and apply the grammar rule in learning. The balance of 11.8% of respondents voted for neutral and disagreed. Moreover, 91.2% of the respondents agree and strongly agree that (s12) “Games help me practise grammar in a fun and interactive way”, which left only 18.8% of respondents to disagree with the statement. Game-based learning can result in more enjoyable learning results as it involves interactive technology (Yusoff, 2013).

Several 30 (88.3%) respondents agree and strongly agree that (s13) “Games enhance my ability to use grammar accurately in real-life situations”. The total percentages and responses received proved the usefulness of learning grammar with games. E-games were introduced to encourage students' participation in the learning process (Yusoff, 2013). The involvement of students in the learning process proves that the students’ learning engagement in a real-time context is achieved. Following the number of 23 respondents (67.6%) mostly voted for (s16) “I actively participate in grammar learning games” and left with a total of 38.4% of 7 respondents who voted for neutral and 4 respondents who disagreed. Behavioural engagement is achieved as most of the students are actively participating in the learning activities. The ability for students to actively engage with the content, get a deeper comprehension of topics, and advance their skills and knowledge makes this active involvement essential for effective learning (Philp & Duchesne, 2016).

Furthermore, 85.3% of respondents strongly agree and agree with (s17) “I find grammar learning games interesting and enjoyable”. The result left 14.7% to vote for neutral and disagreed.

In (s18) “Grammar learning games keep me focused and attentive” 79.4% of respondents agree and strongly agree with the statement, and only 20.6% of the respondents voted strongly agree, disagree and neutral separately. Lastly, a number of 82.3% of respondents agree and strongly agree with (s20) “Grammar learning games challenge me to think critically and apply grammar rules” and left only 17.7% agree with 2 disagree and 4 neutral respondents. Cognitive engagement of the learners is achieved when the learners are mentally engaged with the learning process (Hiver et al., 2021). Organising educational games can help students strengthen their critical thinking and problem-solving abilities (Kopf, 2022). The majority of the respondents agree that grammar learning games are able to help them engage well in the teaching and learning process. Based on the data presented above, the result shows that the usage of games in learning grammar is able to support the learning process while indirectly enhancing the students' learning engagement during teaching and learning sessions.

Last but not least, to discuss the research question three; what are the perceptions of ESL undergraduate students regarding the usefulness of using games in enhancing their motivation, 16 students (47.1%) agree and 11 students (32.4%) strongly agree with the statement (s21) “The game motivated me to progress and get better”. It means 79.5% of respondents agreed with the above statement. This high percentage of agreement and strong agreement indicate that game-based learning has effectively stimulated the participants’ motivation to improve and advance in their performance. These results reinforce the notion that the game successfully serves as a source of motivation and encouragement for a substantial portion of the respondents.

This study also found that game-based learning has a positive influence on respondents’ intrinsic motivation. Table 6 shows that 10 students (29.4%) agree and 15 students (44.1%) strongly agree with the statement (s29) “I feel more confident and active in the class with game-based learning”. According to Nayir (2017), students’ intrinsic motivation is closely linked to the value they attach to the learning process. In other words, the findings of this study revealed that the students were only interested and active in the subjects that were interesting to them. Games, in this case, serve as effective learning tools by bridging the gap between the curriculum, the teacher, the classroom environment, and the students themselves. Previous studies have consistently supported this notion by highlighting how games contribute to creating an engaging and stimulating learning environment that encompasses the essential elements of intrinsic motivation (Malone & Lepper, 2021).

It is further revealed that game-based learning fosters students’ self-determination. Looking at the perception of students on self-determination in the table above, it seems that students perceived gaming as an activity that fulfilled their psychological needs for autonomy, competence, and relatedness. The mean value of survey statement (s28) “Playing games allows me to experience a sense of autonomy and control over my actions and decisions” is the highest of all survey statements for Part 3: Motivation, as 82.3% of students agree and strongly agree with this statement. Next, about 76.5% of students claimed that the game enhances their sense of competence and mastery (s25). In addition, 79.4% of the students agreed and strongly agreed with the statement (s23) “I feel a sense of relatedness and connection with other classmates while engaging in a game-based learning activity”. These three statements from the questionnaire revealed that game-based learning does provide opportunities for respondents to exercise autonomy by allowing them to make choices, providing challenges and possibilities for mastery, and enhancing students' sense of competence by making them feel capable and accomplished.

Contributions of the Study

To conclude, this research instils the spirit of ESL towards undergraduate students by explicating its benefits through games. This research delineates the usefulness of games in grammar accuracy and motivation, subsequently fuelling positivity and driving interest in student engagement in language learning. Additionally, this research is able to help educators with teaching planning that suits students' needs. The traditional method is no longer effective to be applied during the teaching and learning session. New methods should be used and practised by educators nowadays to ensure teaching and learning session is a success.

Conclusion

The empirical findings in this study indicate that the incorporation of games in language learning has a positive impact on students' grammar accuracy, learning engagement, and motivation. In essence, this research investigated how undergraduate ESL students view the usefulness of using games to improve grammar accuracy, student engagement, and motivation. This study's results may have significant implications for how the English language is taught and learned in the future. Teachers can improve their students' language learning experiences by integrating game-based learning activities into the curriculum. By offering an enjoyable and low-pressure setting for practice, games can help children improve their grammatical abilities. In addition, games can encourage students to actively participate in the learning process by stimulating their interest and motivating them to do so. As a result of increased motivation and focus, gaming can also lead to better academic achievements. This study tackles a knowledge gap in the literature by investigating the impact of game-based learning on ESL undergraduate students. Educators can improve their students' English language skills by incorporating creative teaching strategies, such as game-based learning, into their lessons.

Funding: This paper is fully funded by the Ministry grant number
FRGS/1/2024/SSI09/UIAM/02/6
Acknowledgements: The authors would like to express their gratitude to the Ministry of Higher Education and the International Islamic University Malaysia (IIUM).
Conflicts of Interest: The authors declare no conflict of interest.
Authenticity: This manuscript is an original work.
Artificial Intelligence Statement: No AI is used in preparing this manuscript.
About the Author:

Dr. Siti Fatimah Abd. Rahman is an Assistant Professor at the Department of Language and Literacy in the Kulliyyah of Education at the International Islamic University Malaysia (IIUM). Her research interests include flipped learning, gamification, technology acceptance, ESL, and technology-enhanced language learning. ORCID: https://orcid.org/0000-0002-1777-3777

Nur Fadhilah Husna is an English teacher under the Ministry of Education, Malaysia, with five years of experience. She is currently pursuing postgraduate studies at IIUM. Her interests include CEFR-aligned teaching, language assessment, and learner engagement. ORCID ID: https://orcid.org/0009-0006-7282-8220

Nurul 'Aqiilah Mohamad Kamal is an English tutor who enjoys helping learners grow through language. She has worked with students from various backgrounds and uses creative methods to make learning enjoyable. Her focus is on building confidence in grammar, vocabulary, and writing.
ORCID ID: https://orcid.org/0009-0005-2928-6385

Qurratull Aini Ahmad Zukhi is an English teacher under the Ministry of Education Malaysia. She is dedicated to helping students improve their proficiency in writing, grammar, and vocabulary. Passionate about creating a supportive and engaging learning environment, she believes in building her students' confidence and inspiring them to love the English language.
ORCID ID: https://orcid.org/0009-0009-9481-5860

References

Admiraal, W., Huizenga, J., Akkerman, S., & Dam, G. ten. (2011). The concept of flow in collaborative game-based learning. *Computers in Human Behavior*, *27*(3), 1185–1194. https://doi.org/10.1016/j.chb.2010.12.013

Alamri, H., Lowell, V., Watson, W., & Watson, S. L. (2020). Using personalised learning as an instructional approach to motivate learners in online higher education: Learner self-determination and intrinsic motivation. *Journal of Research on Technology in Education, 52*(3), 322-352.

Ali, M. W., Iqbal, M., & Ali, A. (2017). Effectiveness of Language Games in Teaching Grammar to Undergraduates: A Case Study at Islamia College Peshawar. *Humanities and Social Sciences*, *24*(1), 21-33.

Balakrishna, C. (2023). The Impact of In-Classroom Non-Digital Game-Based Learning Activities on Students Transitioning to Higher Education. *Education Sciences*, *13*(4), 328. https://doi.org/10.3390/educsci13040328

Berliner, D. C., & Calfee, R. C. (2004). *Handbook of educational psychology*. Lawrence Erlbaum Associates.

Celce-Murcia, M., & Hilles, S. (1988). *Techniques and resources in teaching grammar*. Oxford University Press.

Chen, Y. (2023). Integrating a Game-Based App to Enhance Translation Learners' Engagement, Motivation, and Performance. *International Journal of Instruction*, *16*(2), 759–782. https://doi.org/10.29333/iji.2023.16240a

DeLoache, J. S. (1987). Rapid change in the symbolic functioning of very young children. *Science, 238*(4833), 1556–1557. https://doi.org/10.1126/science.2446392

Fredricks, J. A., Filsecker, M., & Lawson, M. A. (2016). Student engagement, context, and adjustment: Addressing definitional, measurement, and methodological issues. *Learning and Instruction*, *43*(43), 1–4. https://doi.org/10.1016/j.learninstruc.2016.02.002

Froiland, J. M. & Oros, E. (2012). The importance of cultivating intrinsic motivation to learn in students. *Colorado Society of School Psychologists Newsletter*, *26*(1), 12-14. https://doi.org/10.7358/ecps-2019-019-mor1

Hadfield, J. (1999). The role of games in the language classroom. In A. Smith & B. Jones (Eds.), *Innovative language teaching methodologies* (pp. 45–60). Cambridge University Press.

Hiver, P., Al-Hoorie, A. H., Vitta, J. P., & Wu, J. (2021). Engagement in language learning: A systematic review of 20 years of research methods and definitions. *Language Teaching Research*, 1–30.

Kopf, L. (2022). Serious Games and Gamification: Game-Based Learning in Communication Sciences and Disorders. *Perspectives of the ASHA Special Interest Groups, 7*(2), 482-498.

Landers, R. N., Auer, E. M., Helms, A., Marin, S., & Armstrong, M. B. (2019). Gamification of adult learning: Gamifying employee training and development. *The Cambridge Handbook of Technology and Employee Behaviour*, 271-295.

Lambert, C., Philp, J., & Nakamura, S. (2017). Learner-generated content and engagement in second language task performance. *Language Teaching Research, 21*(6), 665–680.

Lenhart, A., Kahne, J., Middaugh, E., Macgill, E. R., Evans, C., & Vitak, J. (2008). T*eens, video games, and civics*. Washington, DC: Pew Internet & American Life Project.

Littlejohn, A. (2012). Language teaching materials and the (very) big picture. *Electric Journal of Foreign Language Teaching*, (*9*), 283-297.

Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive Science, 5(4),* 333–369.

Malone, T. W., & Lepper, M. R. (2021). Making learning fun: A taxonomy of intrinsic motivations for learning. In *Aptitude, learning, and instruction* (pp. 223-254). Routledge.

Mao, W., Cui, Y., Chiu, M. M., & Lei, H. (2022). Effects of game-based learning on students' critical thinking: A meta-analysis. *Journal of Educational Computing Research, 59*(8),1682-1708.

Mercer, S. (2019). Language learner engagement: Setting the scene. In X. Gao, C. Davison, & C. Leung (Eds.), *International handbook of English language teaching*, 643–660.

Naidu, B., Briewin, M., & Amin Embi, M. (2013). Reading Strategy: Tackling Reading through Topic and Main Ideas. *English Language Teaching*, *6*(11). https://doi.org/10.5539/elt.v6n11p60

Nayir, F. (2017). The relationship between student motivation and class engagement levels. *Egitim Arastirmalari - Eurasian Journal of Educational Research, (71),* 59-77.

Piaget, J. (1962). *Play, dreams and imitation in childhood*. Norton.

Philp, J., & Duchesne, S. (2016). Exploring Engagement in Tasks in the Language Classroom. *Annual Review of Applied Linguistics*, *36*(2), 50–72. https://doi.org/10.1017/s0267190515000094

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational psychologist, 50*(4), 258-283.

Pratama, G. A. (2020). Students perception of gamification to promote classroom engagement and motivation in senior high school. *Language Research Society, 1*(1). https://doi.org/10.33021/lrs.v1i1.1040

Rahman, S. F., & Idris, N. F. M. (2025). Investigating ESL Teachers' Experience in Teaching Online Classes. *Arab World English Journal, 16*(2):36-56. https://dx.doi.org/10.24093/awej/vol16no2.3

Shaffer, D. W., Halverson, R., Squire, K. R., & Gee, J. P. (2005). *Video games and the future of learning* (WCER Working Paper No. 2005-4). Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research (NJ1).

Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology*, *100*(4), 765–781. https://doi.org/10.1037/a0012840

Syafiqah Yaccob , N., & Md Yunus, M. (2019, March 1). *Language Games in Teaching and Learning English Grammar: A Literature Review*. Papers.ssrn.com. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3367576

Vansteenkiste, M., & Ryan, R. M. (2013). On psychological growth and vulnerability: Basic psychological need satisfaction and need frustration as a unifying principle. *Journal of Psychotherapy Integration*, *23*(3), 263–280. https://doi.org/10.1037/a0032359

Vernon, S. (2024, January 29). *Home – Teaching English Games*. Teaching English Games. https://teachingenglishgames.com/

Yaccob, N.S., & Yunus, M.M. (2019). Language games in teaching and learning English grammar: A literature review. *Arab World English Journal, 10*(1), 209-217.

Yu, Z., Yu, L., Xu, Q., Xu, W., & Wu, P. (2022). Effects of mobile learning technologies and social media tools on student engagement and learning outcomes of English learning. *Technology, Pedagogy and Education, 31*(3), 381-398.

Whitmore, K. F., & Meyer, R. J. (2020). *Reclaiming Literacies as Meaning Making*. Routledge.

Yusoff, M. H., (2013). Knowledge development through constructionism game-based learning application: an evaluation of students' performance. *International Journal of Future Computer and Communication*, *2 (3)*, 232-236

Zhang, Y., Fan, Y., Mu, Y., Zhang, P., Wang, M., & Li, J. (2022). Reproductive and oncological outcomes of fertility-sparing surgery in patients with stage I epithelial ovarian cancer: A systematic review and meta-analysis. *Medicine*, *101*(31), e29929. https://doi.org/10.1097/md.0000000000029929

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp.36-63
DOI: https://dx.doi.org/10.24093/awej/vol16no4.3

A1 Level Vocabulary Enhancement Schemes for Writing Development Among Foundation Level 1 and Level 2 Students

Yasmin Banu Abdul Majeed
Preparatory Studies Centre
University of Technology and Applied Sciences-Ibra-Oman
Corresponding Author: yasmin.majeed@utas.edu.om

Firos Kalarikkal
Preparatory Studies Centre
University of Technology and Applied Sciences-Ibra-Oman
Emails: firos.kalarikkal@utas.edu.om

Received: 08/05/2025 Accepted: 11/05/2025 Published: 12/10/2025

Abstract
Foundation level students at University of Technology and Applied Sciences encounter great challenges when it comes to vocabulary retention to perform well in writing skills. Therefore, this study aims to examine the effectiveness of current vocabulary teaching practices for A1 and A2 level students in the General Foundation Program at the University of Technology and Applied Sciences in Oman. The purpose of the study is to find the level efficacy of the current systems of vocabulary enrichment schemes and to see how far A1 and A2 level students benefit from these schemes by finding breakthroughs for effective vocabulary enrichment practice to help students' writing skills. Employing a mixed-method approach, the research assessed the impact of existing vocabulary enhancement schemes on writing development among the respondents. Data was collected through surveys, and interviews with instructors. The findings reveal significant gaps in vocabulary enrichment and retention and indicated amendments in current practices. Prioritizing systematic vocabulary acquisition, and the integration of targeted vocabulary in lessons with an emphasis for retention and practical application in writing tasks are recommended. The study advocates for a more structured approach to vocabulary teaching, with the potential to significantly improve writing proficiency among foundation students, thereby contributing to their overall academic success.

***Keywords*:** A1 level vocabulary, curriculum, current practice, general foundation program, vocabulary acquisition, vocabulary enrichment schemes, vocabulary retention

How to cite: Abdul Majeed, Y. B., & Kalarikkal, F. (2025). A1 Level Vocabulary Enhancement Schemes for Writing Development Among Foundation Level 1 and Level 2 Students. *Arab World English Journal*, *16*(4):36-63.
https://dx.doi.org/10.24093/awej/vol16no4.3

Introduction

The foundational stages of learning a new language are critical in developing core skills—speaking, listening, reading, and writing. Among these, writing is often considered one of the most challenging for learners at the A1 level, as it requires the ability to organize thoughts, structure sentences, and employ vocabulary accurately. In the context of English Language Teaching (ELT), vocabulary acquisition plays a pivotal role in the development of these skills, especially writing. Without a sufficient lexicon, students find it difficult to express ideas in written form, leading to fragmented sentences and incoherent narratives.

At the A1 level, according to the Common European Framework of Reference for Languages (CEFR), learners are expected to understand and use everyday expressions and very basic phrases. They can communicate in simple ways on familiar topics and can write short texts on topics of personal relevance. However, even at this early stage, writing tasks often require a foundational vocabulary that many learners struggle to acquire.

In Oman's foundation programs, specifically at the University of Technology and Applied Sciences (Ibra), writing is a key skill that students are expected to develop. Many students at Level 1 and Level 2 face challenges in writing even the most basic texts, which can largely be attributed to their limited vocabulary. Addressing this issue requires not only a focus on teaching grammar and structure but also on enhancing vocabulary in a way that is accessible and relevant to these learners.

Vocabulary is an integral part of language acquisition. Wilkins (1972, p43) wrote ". . . while without grammar very little can be conveyed, without vocabulary nothing can be conveyed". Nation (2016) asserts that knowing a word involves knowing the word form, meaning and use. Concerning word meaning, knowing a word can be further explained by knowing word form and meaning, knowing word concept and referents, and associations. According to Behforouz and Al Gaithi (2022), Omani learners do not employ different vocabulary learning strategies. Many students entering foundation programs have a very limited range of English vocabulary, often resulting from insufficient exposure to the language in their daily lives. Even though they may be familiar with basic vocabulary items, their usage is often passive, meaning they can understand the words but not use them effectively in writing. This problem is more acute in writing tasks, where vocabulary choice plays a critical role in conveying meaning, developing arguments, and ensuring overall coherence.

The current curriculum at the foundation level includes vocabulary learning, but it is often integrated into broader language instruction without a specific focus on how vocabulary impacts writing skills. This study seeks to fill that gap by focusing exclusively on vocabulary enhancement schemes that target writing development. Given that vocabulary is the building block of language, this research aims to provide a targeted approach that can be incorporated into the curriculum to improve student outcomes.

Vocabulary development is a crucial component of language learning, particularly at the A1 and A2 levels of the CEFR. These levels represent the foundational stages where learners acquire essential language skills necessary for effective communication. The General Foundation Program (GFP) at the University of Technology and Applied Sciences (UTAS), Oman, plays a significant role in preparing students for higher education by strengthening their English language proficiency. Vocabulary learning is integral to this process, as it supports other key language skills such as reading, writing, listening, and speaking. Despite the emphasis on vocabulary instruction in GFP, many students continue to face challenges in mastering even basic vocabulary, which

hampers their overall language performance and academic success. These challenges highlight the need for a comprehensive examination of current vocabulary teaching practices to ensure their effectiveness.

In recent years, vocabulary teaching methodologies have undergone significant changes, moving from traditional rote memorization to more communicative and student-centered approaches. Contemporary techniques emphasize contextual learning, task-based activities, and digital resources to engage learners actively in the vocabulary acquisition process. However, the extent to which these methods are applied in Omani GFP classrooms and their effectiveness in enhancing students' vocabulary remains an open question. The learning environment, teacher preparedness, availability of technological resources, and students' exposure to English outside the classroom are critical factors that may influence vocabulary acquisition. Additionally, cultural and educational contexts play a vital role in shaping both teaching practices and student learning outcomes.

Previous studies on vocabulary teaching in EFL contexts have stressed the importance of learner autonomy, motivational strategies, and the integration of technology in language instruction. However, there is limited research specifically focused on Omani GFP students at the A1 and A2 levels. This gap in the literature necessitates a closer investigation into the effectiveness of existing vocabulary teaching strategies at UTAS and the identification of areas for improvement.

Given the diverse academic backgrounds of GFP students, understanding their specific learning needs is essential for developing tailored instructional strategies. The findings of this study will not only provide insights into the current vocabulary teaching practices but also offer recommendations for more effective pedagogical approaches. By addressing the existing challenges and identifying potential remedies, this research aims to contribute to the enhancement of vocabulary instruction and, ultimately, the overall language proficiency of foundation-level students at UTAS.

Many students entering foundation programs have a very limited range of English vocabulary, often resulting from insufficient exposure to the language in their daily lives. Even though they may be familiar with basic vocabulary items, their usage is often passive, meaning they can understand the words but not use them effectively in writing. This problem is more acute in writing tasks, where vocabulary choice plays a critical role in conveying meaning, developing arguments, and ensuring overall coherence.

The current curriculum at the foundation level includes vocabulary learning, but it is often integrated into broader language instruction without a specific focus on how vocabulary impacts writing skills. This study seeks to fill that gap by focusing exclusively on vocabulary enhancement schemes that target writing development. Given that vocabulary is the building block of language, this research aims to provide a targeted approach that can be incorporated into the curriculum to improve student outcomes.

The researchers of the current study have been teaching English at different levels for more than twenty years. During their teaching career, they observed that most of the students at the foundation level lack basic, day-to-day vocabulary knowledge and this affects their writing capabilities.

Hypotheses- The following hypotheses were tested during the current study:

1. The current practice doesn't help learners with A1 and A2 vocabulary acquisition.

2. The exposure of A1 and A2 vocabulary will aid them in improving their writing.

Research Questions

1. Are students from foundation level 1 and 2 aware of A1 and A2 vocabulary?
2. Is the current plan to achieve Vocabulary LOs sufficient?
3. Are students capable of learning and retaining a new set of vocabulary?

It is very important for students to develop strong vocabulary skills in order to express themselves clearly and meaningfully both in written and spoken formats. It is not only imperative for education, language arts, and reading, but it is also to present themselves in public as educated individuals. Developing a wide range of vocabulary and vocabulary skills begins early in a student's life and continues through their education. A good grasp of vocabulary gives the student the ability to say what he or she means. The greater number of words the students know, the more he or she can interpret ideas from others and express their own ideas. In addition to the above, it boosts students' power of persuasion. Thus, in view of the importance of vocabulary development, this study focused on the A1 level vocabulary development of the students.

The present study seeks to evaluate the factors influencing vocabulary enrichment and retention among Foundation Level 1 and Level 2 students at the University of Technology and Applied Sciences (UTAS), Oman. It aims to assess how various instructional, contextual, and learner-specific elements impact vocabulary acquisition and long-term retention. By examining the current vocabulary teaching mechanisms, this research will analyze their effectiveness in meeting the linguistic needs of A1 and A2 learners and fostering sustainable vocabulary growth. The study will also identify gaps and challenges in the existing practices, offering insights into strategies for improving vocabulary learning outcomes. Ultimately, it strives to recommend evidence-based interventions to optimize vocabulary instruction and empower learners with essential language skills for academic success and beyond.

The study shed light on the areas of students' weakness in basic vocabulary power by examining their struggle to use the content vocabulary in context across varied language skills. The study suggests measures for both the tutors and students to follow to enable the learners to use the A1 vocabulary with much comfort. Also, it provides some scholastic activities for students throughout the semester to reinforce A1 vocabulary. The study aims to benefit the Omani student and teaching communities as the main population of the study is the same. Also, the study focusses on the prevailing foundation courses at UTAS. It is hoped that the findings will encourage students to revitalize the current strategy of vocabulary enrichment among Level 1 and Level 2 students.

Literature Review

Vocabulary acquisition is a critical component of language learning, particularly at the A1 level, where learners begin to develop foundational language skills. Research has consistently highlighted the importance of vocabulary knowledge for language proficiency. According to Nation (2013), it plays a vital role in language acquisition particularly in writing. To add to this note, Schmitt (2008) opined that A1 level vocabulary encompasses basic words and phrases essential for everyday communication, and its mastery is crucial for developing writing competence. However, Alqahtani (2015) commented that vocabulary retention and enrichment remain challenging, especially among students in foundation programs, where English serves as a second language.

The Role of Vocabulary in Writing Development

Research underscores that vocabulary is a foundational component for writing development, particularly for learners at the A1 level of the CEFR framework. Nation (2001) emphasized the strong correlation between vocabulary size and writing competence, asserting that a limited vocabulary constrains students' ability to express ideas and construct meaningful texts. Learners at lower proficiency levels often struggle with lexical choices, which can hinder coherence and complexity in their writing. Building a robust vocabulary base is therefore essential to enhancing written communication skills for beginner learners.

Effective Vocabulary Enhancement Strategies

Numerous studies have highlighted strategies for vocabulary enhancement in EFL contexts, including explicit vocabulary instruction, incidental learning, and technology-supported approaches. Thornbury (2002) advocated for the use of explicit teaching techniques, such as direct word explanations, word mapping, and context-based examples. These techniques help reinforce vocabulary retention and comprehension, particularly among learners with limited language exposure. Additionally, Hulstijn and Laufer (2001) argued that task-induced involvement, which requires learners to actively manipulate vocabulary in meaningful contexts, promotes better vocabulary retention.

Contextual Learning and Task-Based Approaches

Task-based language teaching (TBLT) has proven to be an effective approach for vocabulary acquisition, particularly for developing writing skills (Ellis, 2003). When learners engage in writing tasks that require the use of new vocabulary, they not only practice lexical retrieval but also develop contextual understanding. This dual engagement strengthens vocabulary retention and facilitates the transfer of lexical knowledge to authentic writing contexts. Additionally, contextual vocabulary learning fosters deeper semantic processing, making it easier for students to use words accurately in written texts.

Role of Digital Tools in Vocabulary Development

The integration of digital tools and mobile-assisted language learning (MALL) platforms has revolutionized vocabulary teaching and learning. Studies by Stockwell (2010) and Godwin-Jones (2011) revealed that learners exposed to interactive vocabulary apps demonstrated better vocabulary recall and improved writing fluency. These tools provided personalized learning experiences through gamification, spaced repetition, and multimedia content, which make vocabulary acquisition engaging and accessible for A1 learners.

Writing-focused Vocabulary Enhancement Programs

Research by Coxhead (2000) highlighted the effectiveness of vocabulary lists tailored to writing needs, such as the Academic Word List (AWL). For A1 learners, simplified versions or targeted word lists aligned with writing tasks can help scaffold learning and foster confidence in writing development. Additionally, Storch (2013) argues collaborative writing activities and peer feedback have been shown to encourage vocabulary usage and improve writing fluency.

Existing literature suggests that vocabulary teaching methods in foundation programs often rely on traditional rote memorization and isolated word lists. Newton and Nation (2020) warn that these are insufficient for deep learning and retention. In the context of Oman's higher education,

Al-Mahrooqi and Denman (2018) point out that vocabulary instruction often lacks systematic approaches that integrate vocabulary into meaningful contexts, thereby limiting students' ability to apply newly learned words in writing tasks. As suggested by Webb and Nation (2017), vocabulary teaching should involve strategies such as spaced repetition, contextual learning, and active use of vocabulary in writing. The sociocultural context significantly influences language learning, including vocabulary acquisition. In Oman, observe Al-Issa and Al-Bulushi (2012), English is a second language, and students often have limited exposure to English outside the classroom. This context necessitates more robust vocabulary instruction within academic settings to compensate for the lack of natural language exposure. Furthermore, Syed (2003) claims that studies in the Gulf region have highlighted the need for culturally relevant materials that connect with students' everyday experiences, making vocabulary more memorable and meaningful.

Benge & Robbins (2022) rightly observed that positive impact of using the integrated method based on the Mind-Mapping Strategy (MMS) and Key-Word Strategy (KWS) for developing/enriching and retaining vocabulary knowledge. EFL teachers, learners and educational curriculum developers are recommended to integrate KWS and MMS in the Saudi context.

Schmitt (2010) is of the opinion that vocabulary retention is a significant challenge among A1 learners due to the lack of reinforcement and practical application in their curriculum. Studies have shown that vocabulary knowledge directly impacts writing quality, influencing factors like coherence, fluency, and lexical variety. Read (2000) hinted to this grave issue. For students in foundation programs, inadequate vocabulary retention impedes their writing development and as Milton (2009) observes, this leads to simplistic and repetitive language use that fails to meet academic standards. Roediger & Butler (2011) suggested that implementing retrieval practice, where students actively recall words in varied contexts, can significantly enhance retention and improve writing fluency. This approach is especially important at the A1 level, where foundational vocabulary needs to be reinforced consistently.

Collaboration among students can have a lead. According to Zamanian et al., 2022, p. 62, the two groups involved in literature-based interactive (collaborative and collaborative-cooperative) reciprocal input-output activities showed significantly higher levels of vocabulary retention compared to the group exposed to individual/non-reciprocal processing.

As Creswell (2014) stated, mixed methods research, which combines qualitative and quantitative approaches, offers a comprehensive understanding of vocabulary acquisition and its impact on writing. Quantitative data, such as test scores, provide measurable outcomes of vocabulary retention, while qualitative insights from interviews and classroom observations reveal the underlying challenges and needs of learners. Dörnyei (2007) emphasized on qualitative analysis of data to reach flawless conclusions. This approach allows for a holistic analysis of vocabulary teaching effectiveness, highlighting areas for curriculum improvement.

The current literature emphasizes the need for curriculum designs that prioritize vocabulary learning and retention as a core element of language instruction. A set of researchers such as Nation (2013) and Laufer and Hulstijn (2001) said that integrating vocabulary into writing tasks, using authentic materials, and employing technology-based tools like flashcards and language apps can significantly enhance vocabulary acquisition. For foundation programs, such as those in Oman, developing targeted vocabulary courses that focus on practical usage can bridge the gap between learning and application, ultimately improving students' writing skills.

The literature reveals a consistent call for more effective vocabulary teaching strategies in foundation programs, highlighting the direct impact of vocabulary knowledge on writing

proficiency. As the current study investigates the existing practices in the UTAS in Oman, it aligns with broader findings that emphasize the need for vocabulary enrichment and retention within the curriculum. The findings of this research could inform the development of targeted interventions that address these gaps, ultimately enhancing writing development among A1 level students.

Curriculum interventions that emphasize vocabulary acquisition can significantly impact students' writing skills. Integrating vocabulary into all language skills—listening, speaking, reading, and writing—ensures that words are learned in a comprehensive, interconnected manner. Stahl and Fairbanks (1986) are of this point of view. Additionally, Graves (2006) advocates that vocabulary instruction should be systematic, involving deliberate planning, explicit teaching, and continuous assessment of learners' vocabulary progress. Programs that incorporate vocabulary-focused writing tasks, where students apply new words in various writing genres, according to Webb (2005), have shown positive effects on both retention and writing quality.

Methodology

The present study employed a mixed-methods research design, integrating both quantitative and qualitative approaches to gain comprehensive insights into vocabulary enhancement for Omani Foundation Level 1 and Level 2 students. The quantitative component involved the administration of surveys aimed at capturing a descriptive understanding of the students' existing English vocabulary competence. This data provided a snapshot of the current linguistic capabilities and challenges faced by learners. Descriptive analysis of the quantitative findings allowed for a detailed assessment of the prevailing status of vocabulary acquisition among the target population.

Complementing the quantitative inquiry, the qualitative aspect focused on evaluating the effectiveness of intervention strategies for vocabulary improvement. Data was collected through structured interviews and open-ended questionnaires, gathering insights from both tutors and learners. Tutors' perceptions were explored to understand instructional approaches, challenges, and observed outcomes, while learners' responses shed light on their experiences, engagement, and progress with the intervention strategies. The integration of both data types enabled a holistic evaluation of the current vocabulary teaching practices, learner outcomes, and potential areas for pedagogical improvement

Structure of the Study

The current study was undertaken by following the steps mentioned below:

Stage one: In the first stage, the researchers, after the approval from the RCD (Research and Consultancy Department), prepared the research instruments such as surveys, interview questions etc. The stake holders, students, and teachers were approached, and responses were collected online (quantitative data) and face-to-face (qualitative data).

Stage two: In the next stage, the data was processed by using the latest version of SPSS for qualitative data and Nvivo for qualitative data analysis (thematic analysis).

Stage three: The data was carefully evaluated, and the data analysis was done more intensively at this stage, and the researchers have furnished the discussion points to finalize the recommendations.

Sample Size

Fifty level 1 and level 2 students were randomly selected as the subjects for the present study. These students are just beginners at the foundation level at the Preparatory Studies Centre of the University of Technology and Applied Sciences-Ibra, Oman.

In addition, five teachers, each from level 1 and level 2, were approached for interviews to get highlights of their professional experience in connection with the study topic. All the participants in the study showed their consent in a signed consent form which was collected by the researchers.

Table 1. *Age group of respondents.*

Age		Frequency	Percent
18- 21		50	78.1
Under 18		14	21.9

Method of Data Collection and Data Analysis

The core learner respondents administered a structured survey comprising 20 questions designed to capture their perspectives on vocabulary acquisition and the effectiveness of existing enhancement strategies. The quantitative data collected were analyzed using the Statistical Package for the Social Sciences (SPSS) to derive meaningful patterns and trends for formal discussion.

To complement the quantitative findings, qualitative data were gathered through one-on-one interviews with educators. A set of 20 open-ended questions guided the interviews, which were fully voice-recorded and subsequently transcribed using the Otter application. The audio recordings and transcripts were meticulously reviewed multiple times to ensure accuracy and capture nuanced insights. Emerging themes were carefully identified and aligned with the study's objectives.

The qualitative data were then subjected to thematic analysis using NVivo. The similarities and differences in educators' opinions were examined and categorized under five key themes, which provided a structured framework for evaluating the current vocabulary enhancement practices. This rigorous methodological approach ensured that both learner and educator perspectives were integrated, yielding a balanced and in-depth understanding of vocabulary development efforts at the university.

Results

The current study has reached interesting findings related to vocabulary enrichment practices that exist at the University of technology and applied sciences, Ibra.

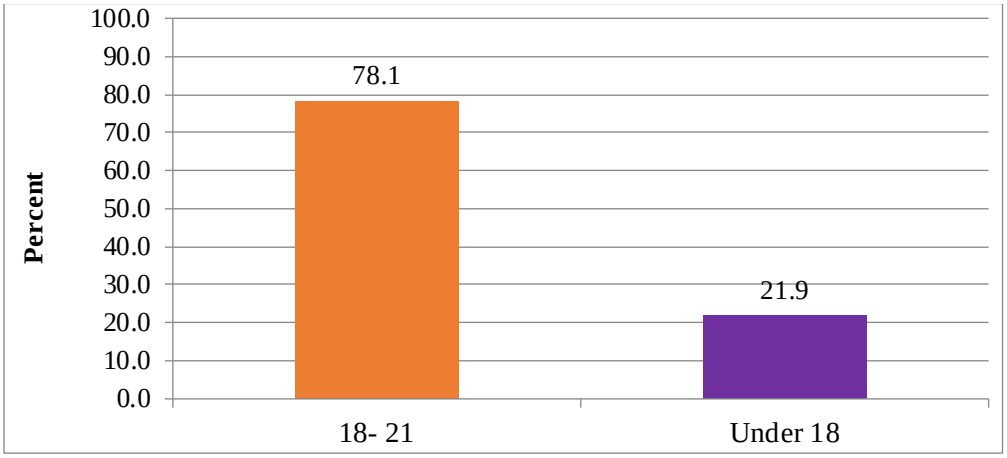


Figure 1. Age group of respondents

Most of the respondents from students' core who participated in this study fall under the age group 18-21. They stand at 78.1 percent compared to 21.9 percent of those who are under the age of 18. This is significant as the learners whose responses were examined were young adults who should have ample range of A1 and A2 level vocabulary.

Table 2. *Duration of Respondents' Language learning?*

	Frequency	Percent
5-7 years	5	7.8
Less than 3 years	29	45.3
More than 7 years	23	35.9
Years 3-5	7	10.9
Total	64	100.0

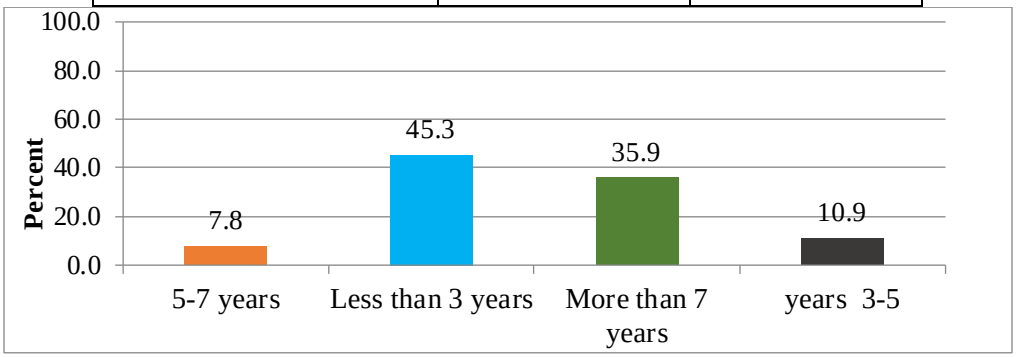


Figure 2. Duration of Respondents' Language Learning

It is noted that 35.9 percent of students have been learning English for more than seven years. More than half of the respondents have been learning the language for more than five years. It means the learners have got enough time to acquire the primary vocabulary. However, this is not the reality. Only 3.1 percent of the respondents stated that they get frequent exposure to English language usage in their everyday life. (See Figure 3 below)

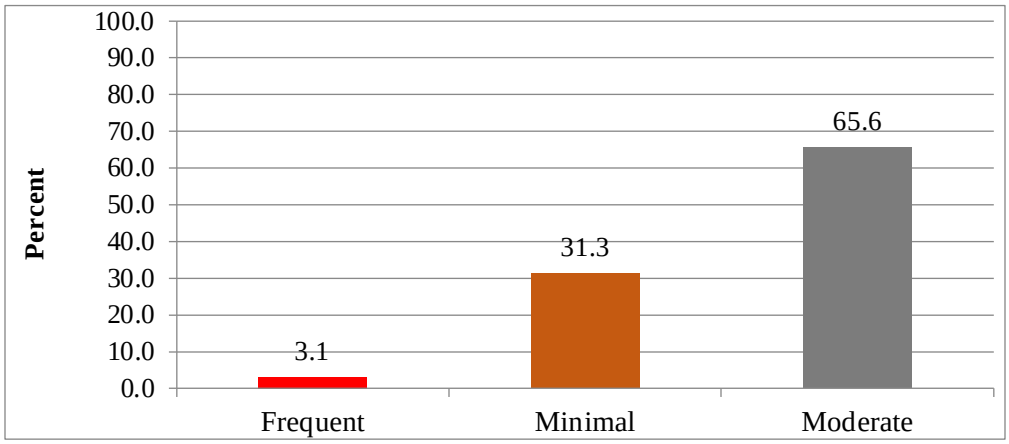


Figure 3. Frequency of exposure to language usage in everyday life

ISSN: 2229-9327

Interestingly, 62.7 percent of the students haven't enrolled in any sort of vocabulary enrichment schemes or training program. This data tells us that tailor-made training schemes should be provided to the learners to make them active in vocabulary enhancement. Moreover, students should be taken to everyday practice of vocabulary as the data reveals the number of everyday users of content vocabulary is only below 15 percent.

Table 3. *Methods used to learn new vocabulary*

Method	Frequency	Percent
Flashcards	2	3.1
Interactive apps	10	15.6
Other	21	32.8
Reading exercises	15	23.4
Vocabulary list	16	25.0
Total	64	100.0

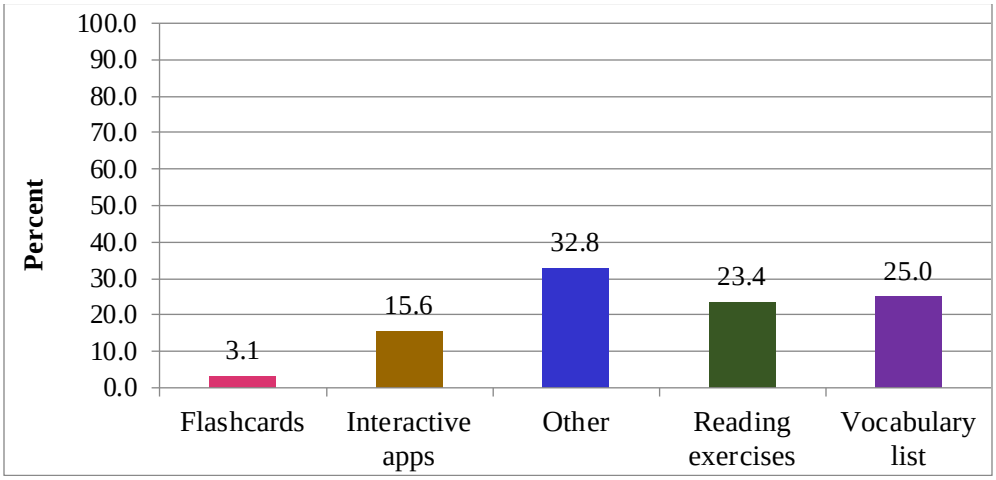


Figure 4. Methods used to learn new vocabulary

Digital learning is a thing of the day, and it must be brought into vocabulary learning. Around half of the respondents revolve around the traditional ways of learning vocabulary with a vocabulary list and reading exercises with numbers of 25 percentage and 23. 4 Percentage respectively whereas users of interactive applications stand at only 15.6 percent. It indicates that the institution should provide more digitized learning experience to bring learners to the environment of e-learning.

Table 4. *Effectiveness of current vocabulary learning methods*

Effectiveness	Frequency	Percent
Neutral	27	42.2
Somewhat effective	16	25.0
Somewhat Ineffective	5	7.8
Very effective	9	14.1
Very ineffective	7	10.9
Total	64	100.0

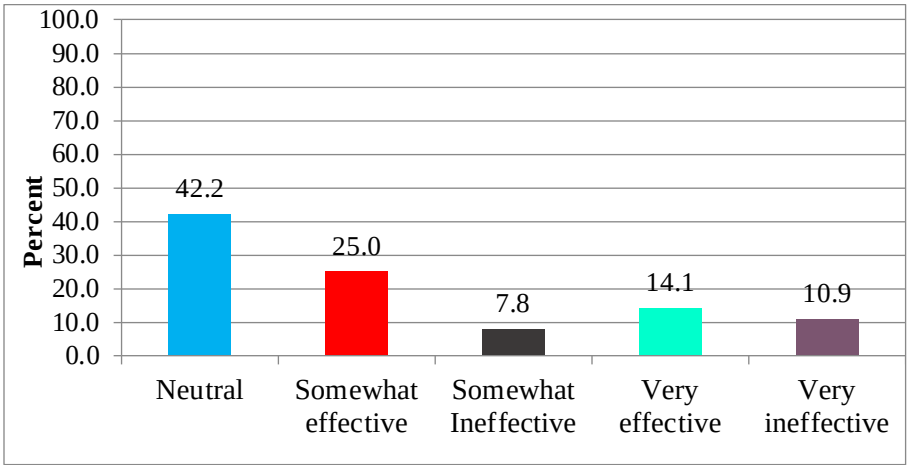


Figure 5. Effectiveness of current vocabulary learning methods

The data from Table No. 4 reveals mixed perceptions among respondents regarding the effectiveness of their current vocabulary learning methods. The largest proportion, 42.2%, stands at a neutral stance, indicating that a significant number of learners are undecided or feel uncertain about their strategies. Meanwhile, 25.0% found their methods somewhat effective, and 14.1% rated them as very effective, suggesting that about 39.1% of the respondents experience some tint of success. However, 18.7% of the participants indicated ineffectiveness, with 7.8% finding their methods somewhat ineffective and 10.9% rating them as very ineffective. Overall, the data suggests a need for improvement in vocabulary learning strategies, as a notable portion of learners either remain neutral or express dissatisfaction with their current approaches. Contextual usage of the content vocabulary should be encouraged as the number of students who use this strategy is less than three percent currently.

Table 5. *Vocabulary enhancement schemes and improvement in writing skills*

Extent	Frequency	Percent
Neutral	18	28.1
Not at all	2	3.1
Slightly	10	15.6
Somewhat	27	42.2
Very much	7	10.9
Total	64	100.0

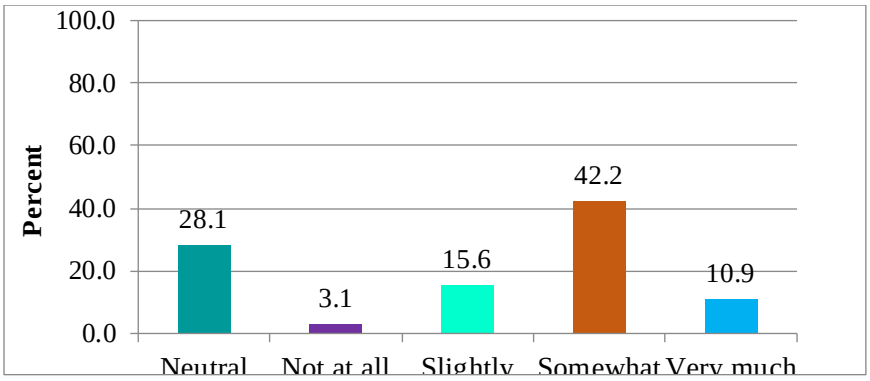


Figure 6. Vocabulary enhancement schemes and improvement in writing skills

Around 42.2% respondents are of the opinion that vocabulary enhancement schemes like vocabulary logs somewhat improve their writing skills while 28.1% remain neutral, indicating uncertainty about their effectiveness. Only 10.9% report that these schemes improve their writing skills very much, while 15.6% observe only slight improvement. A minimal of 3.1% students find no benefit at all. Overall, while many acknowledge some positive impact, the data suggests that the perceived effectiveness of vocabulary logs is limited, highlighting potential areas for refining these strategies. Students need to use the learned vocabulary in their writing assignments. Learners' confidence level must be boosted by giving constructive feedback on the usage of content vocabulary in their writing tasks.

Table *6. Overall writing grades and vocabulary enhancement schemes*

Response	Frequency	Percent
No change	7	10.9
Not applicable	1	1.6
Significant improvement	12	18.8
Some improvement	44	68.7
Total	64	100.0

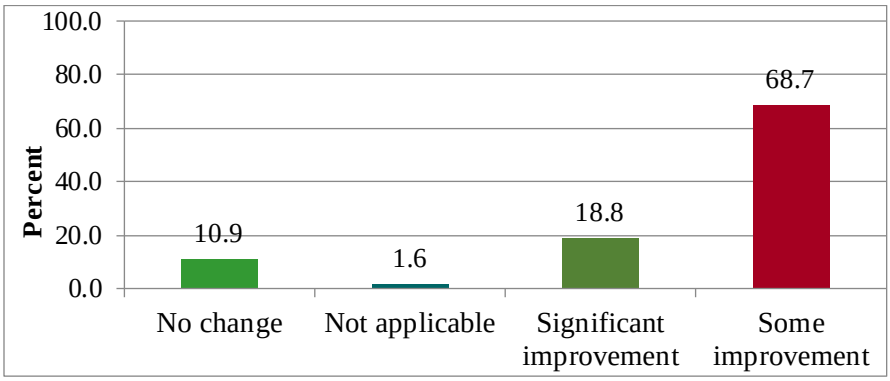


Figure 7. Overall writing grades and vocabulary enhancement schemes

It is interesting to see that 68.7% of respondents have found some improvement in their overall writing grades due to vocabulary enhancement schemes, whereas 18.8% report significant improvement. This suggests that these schemes are generally effective in enhancing writing performance for most learners. However, 10.9% of participants observed no change in their writing grades, and 1.6% considered the question not pertinent. The data points out that while vocabulary enhancement strategies are beneficial for most students, a small portion does not experience a fruitful impact. As the number of students who take additional vocabulary enhancement training is less and it affects the rate of vocabulary retainment.

Table7. *Vocabulary development activities in My ELT/Moodle and writing needs*

Response	Frequency	Percent
Agree	22	34.4
Disagree	3	4.7
Neutral	26	40.6
Strongly agree	12	18.8
Strongly disagree	1	1.6
Total	64	100.0

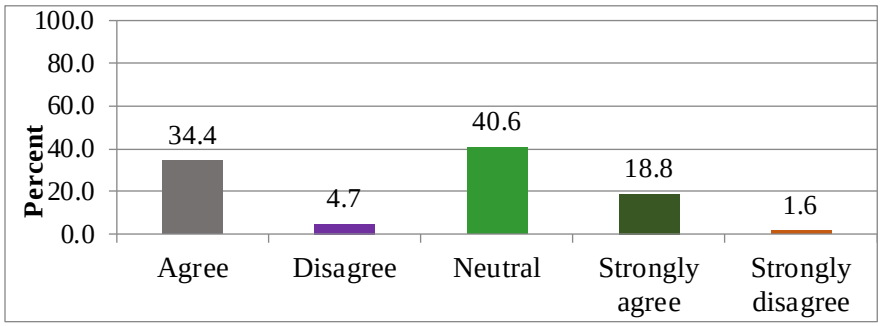


Figure 8. Vocabulary development activities in My ELT/Moodle and writing needs

Mixed responses were given regarding the alignment of vocabulary development activities in My ELT/Moodle with writing needs. 53. 2% is in favor of the statement with 34.4% agree and 18.8% strongly agree. However, a considerable 40.6% stay neutral, suggesting doubt or indifference. Only a small percentage express dissatisfaction, with 4.7% disagreeing and 1.6% strongly disagreeing. This suggests that while more than half find the activities relevant, educators need to cater a considerable group that is either unsure or unconvinced of their alignment with writing requirements.

Suggestions For Improving the Current Vocabulary Enhancement Schemes.

Table 8. *Improving the current vocabulary enhancement schemes*

Response	Frequency	Percent
Greater variety of resources	12	18.8
More focus on writing application	18	28.1
More interactive activities	23	35.9
Other	7	10.9
Regular feedback on progress	4	6.3
Total	64	100.0

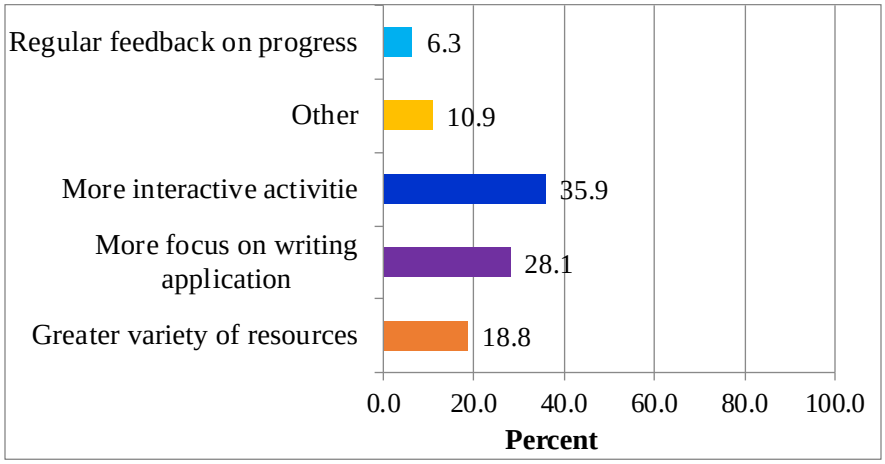


Figure. 9. Improving the current vocabulary enhancement schemes

Figure 9 explains that the most appealing improvement in current vocabulary enhancement schemes is the inclusion of more interactive activities, chosen by 35.9% of respondents. Another 28.1% of participants would like a greater focus on applying vocabulary to writing, emphasizing the need for practical integration. A notable 18.8% ask for a larger range of resources, indicating

a demand for diverse materials while regular feedback on progress is demanded by 6.3% of respondents. Overall, the emphasis on interactivity and writing application points to a preference for more engaging and practical vocabulary learning methods. Currently, most of the vocabulary enhancement materials are largely given on printed materials. Online platforms should be encouraged largely to enable learners to learn and practice new vocabulary on the go.

Additional Support (Tutoring or Workshops) For Vocabulary Enhancement

Table 9. *Additional support (tutoring or workshops) for vocabulary enhancement*

Response	Frequency	Percent
Helpful	27	42.2
Neutral	15	23.4
Not helpful at all	2	3.1
Not very helpful	6	9.4
Very helpful	14	21.9
Total	64	100.0

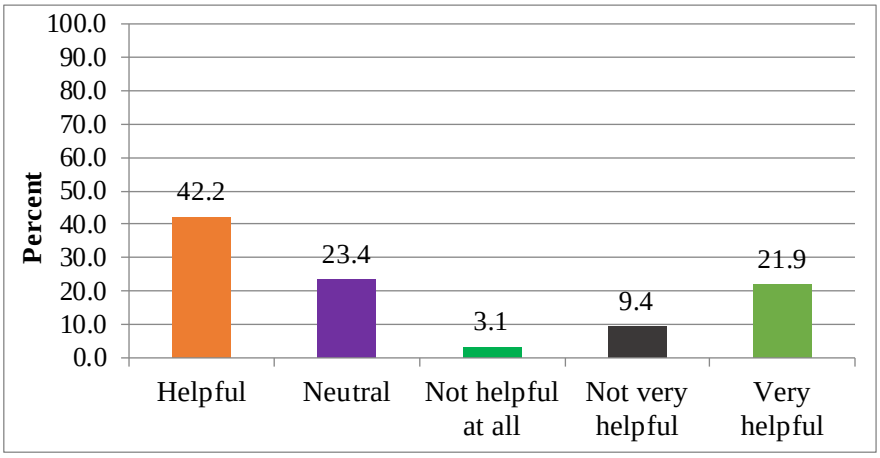


Figure 10. Additional support (tutoring or workshops) for vocabulary enhancement

The data suggests that additional support, such as Academic Support Program (ASP) or workshops, is generally perceived positively for vocabulary enrichment. A combined 64.1% of respondents believe it would be helpful or very helpful, with 42.2% rating it as helpful and 21.9% as very helpful. Meanwhile, 23.4% remain neutral, suggesting that some are unsure about the impact of such support. This poses a serious concern regarding such initiatives. Only a small minority find it unhelpful, with 9.4% considering it not very helpful and 3.1% deeming it not helpful at all. This implies that most learners see value in extra support, though a portion remains skeptical of its necessity.

Interest In Peer-Led Vocabulary Enhancement Group

Table 10: *Interest in peer-led vocabulary enhancement group*

Response	Frequency	Percent
Maybe	24	37.5
No	10	15.6
Yes	30	46.9
Total	64	100.0

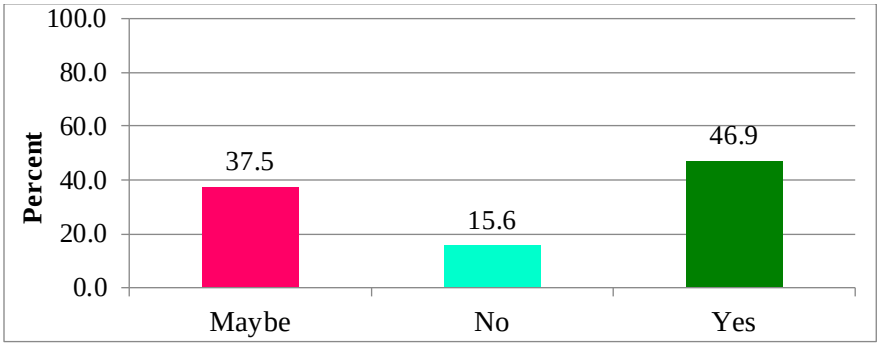


Figure 11. Interest in peer-led vocabulary enhancement group

Interest in participating in a peer-led vocabulary enhancement group is fairly strong, with 46.9% of respondents expressing willingness. Another 37.5% are receptive to the idea but remain unsure, indicating potential engagement if the conditions are attractive. Only 15.6% are not interested, suggesting a relatively small group of definite non-participants. Overall, as there is considerable interest, peer-led vocabulary enhancement groups must be encouraged more.

Overall Effectiveness of the Current Vocabulary Enhancement Schemes in Improving Writing Skills

Table11. *Overall effectiveness of the current vocabulary enhancement schemes*

Rating	Frequency	Percent
Average	18	28.1
Excellent	16	25.0
Good	27	42.2
Poor	2	3.1
Very poor	1	1.6
Total	64	100.0

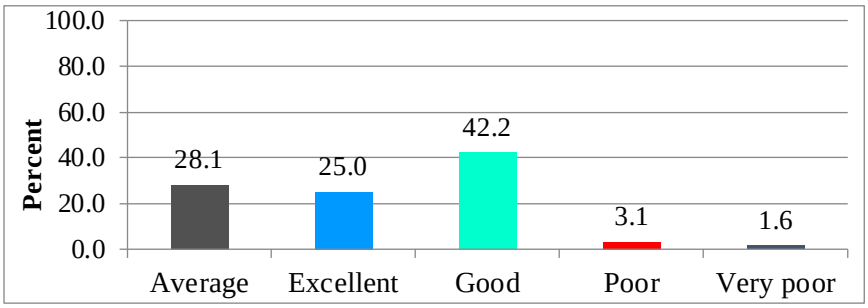


Figure 12. Overall effectiveness of the current vocabulary enhancement schemes

Discussions

The study's findings reveal a notable paradox: while a large majority of students at the UTAS, Ibra have had significant exposure to English, their vocabulary acquisition and application are perceived as insufficient. With 78.1% of respondents aged 18-21 and 35.9% having learned English for over seven years, the cohort consists of young adults with a long history of language education. Despite this extensive background, only a small fraction, 3.1%, reported frequent exposure to English in their daily lives. This low level of practical application outside the classroom is a critical factor undermining vocabulary development and suggests that traditional, school-based learning alone is not enough to build robust language skills.

The data on learning methods highlights a strong reliance on traditional approaches. Approximately half of the respondents, 48.4%, use vocabulary lists and reading exercises as their primary methods, with 25% for vocabulary lists and 23.4% for reading exercises. In contrast, the use of interactive apps, a key component of modern digital learning, is significantly lower at 15.6%. This preference for conventional methods over digital tools may contribute to the perceived ineffectiveness of current learning strategies, as a combined 42.2% of students felt neutral about their methods, and an additional 18.7% found them ineffective. This points to a need for integrating more engaging and technology-driven approaches to cater to the digital-native student population.

Despite the mixed perceptions of current methods, there is a clear consensus on the value of targeted vocabulary enhancement schemes. The study found that 68.7% of students experienced "some improvement" in their writing grades from these schemes, and another 18.8% reported "significant improvement." This strong correlation underscores the positive impact of dedicated vocabulary practice on writing skills. Moreover, a substantial number of students, 64.1%, view additional support such as tutoring or workshops as helpful or very helpful. This indicates a strong demand for structured, extracurricular support to complement their in-class learning and directly address their vocabulary needs.

Student suggestions for improving vocabulary schemes align with the observed gaps in current practice. The top three suggestions are for more interactive activities (35.9%), a greater focus on writing application (28.1%), and a wider variety of resources (18.8%). This feedback reinforces the need to move beyond static, printed materials and toward dynamic, practical, and varied methods. The emphasis on writing application specifically highlights the desire for learners to not just memorize words but to actively use them in context, addressing the low daily exposure rates revealed earlier in the study.

Finally, the survey reveals a strong interest in peer-led learning. 46.9% of respondents are willing to join a peer-led vocabulary group, with an additional 37.5% expressing potential interest. This high level of receptiveness suggests that peer collaboration could be a highly effective, low-

cost strategy for vocabulary enhancement. Implementing such groups would not only provide a platform for interactive practice but also foster a sense of community and shared responsibility for learning, helping to bridge the gap between traditional learning and the practical application of vocabulary in a social, engaging context.

Conclusion and Recommendations

Both the quantitative and qualitative analysis shed important insights related to the current study. Interviews done with the selected educators emphasize the effectiveness of existing materials, the challenges of student motivation, and the importance of supplementary and engaging teaching strategies to improve vocabulary acquisition. The current vocabulary enrichment and retention schemes at UTAS such as vocabulary logs are effective, but they suffer from lack of keen interest in learners, Therefore, it is teacher responsibility to get the learner to mind the vocabulary log in consideration to boost language.

The findings are consistent with and reinforce several key arguments from the existing literature on vocabulary acquisition and its relationship to writing skills. The data reveal a critical disjuncture between learners' prior language education and their practical language usage, confirming observations by Al-Issa and Al-Bulushi (2012) that a lack of exposure outside the classroom is a major challenge in the Omani context. This limited daily use of English (3.1% of respondents) directly supports the claims of Schmitt (2010) and Nation (2013) that without consistent reinforcement and practical application, vocabulary retention and its use in writing will be minimal, leading to the simple and repetitive language use noted by Milton (2009).

The results of the study on learning methods and effectiveness further align with the literature. The finding that half of the students still rely on traditional methods like vocabulary lists and reading exercises echoes the concerns raised by Newton and Nation (2020) and Al-Mahrooqi and Denman (2018) about the prevalence of rote memorization, which is insufficient for deep learning and retention. The low adoption of interactive apps (15.6%) contradicts the benefits highlighted by Stockwell (2010) and Godwin-Jones (2011), who advocate for technology's role in making vocabulary acquisition more engaging and accessible. The mixed or neutral perceptions of effectiveness among a significant portion of the student body (42.2% neutral, 18.7% ineffective) indicate a clear need to move away from these static methods and embrace more dynamic, interactive strategies that involve learners in the active application of words, as advocated by Hulstijn and Laufer (2001).

These findings lead to several key recommendations, grounded in the principles of effective vocabulary instruction outlined in the literature. First, the university should establish targeted vocabulary programs that are integrated into the curriculum, focusing on practical usage to bridge the gap between learning and application, as suggested by Webb and Nation (2017). These programs should move beyond traditional methods and incorporate interactive, technology-based tools to make learning more engaging. Second, the curriculum must emphasize contextual learning and task-based approaches (Ellis, 2003). By integrating vocabulary-focused writing tasks into all assignments, students can actively apply new words, thereby strengthen retention and improve writing fluency, as proposed by Webb (2005). This directly addresses the study's findings that students want to see more focus on writing applications (28.1%).

Finally, the study’s findings on student interest in peer-led groups (46.9% expressing interest) and additional support (64.1% finding it helpful) provide a strong basis for implementing collaborative and supportive learning initiatives. This aligns with Storch's (2013) argument that

collaborative writing activities encourage vocabulary usage and improve fluency. Offering peer-led enhancement groups and workshops would not only cater to students' preferences for more interactive activities (35.9%) but also provide the regular reinforcement and constructive feedback that students need to build confidence and apply their vocabulary effectively. This approach would help address the challenges of limited external exposure and inadequate classroom practice, ultimately enhancing students' overall writing proficiency.

Challenges

Students' perception that vocabulary is unimportant must be actively addressed. Teachers should instill the habit of using newly acquired vocabulary in students' daily communication while emphasizing that vocabulary forms the backbone of language proficiency. The data reveal that students dedicate minimal time to practicing learned vocabulary, and their use of the language beyond classroom settings is infrequent. This lack of consistent engagement poses significant challenges to vocabulary retention and long-term language development.

To address this issue, educators are encouraged to design customized assignments that facilitate the practical application of vocabulary in real-life contexts. By integrating authentic and meaningful tasks, students can develop a sense of relevance and purpose in using the language. MyELT exercises can serve as a foundational resource for crafting tailored assignments, particularly for speaking activities conducted outside the classroom. These tasks should prompt learners to apply their vocabulary in social, academic, or professional situations, thereby strengthening retention and fostering a more natural, confident use of language.

Handling Variations

Monotonous practices can result in lethargic learning of vocabulary and finally end in catastrophic results. In most cases, teachers can overcome the issue of dull attitude among the learners by adopting varied modes of delivery and practice of the vocabulary. One important suggestion is that teachers need to focus on Meaning, form and pronunciation (MFP) when teaching vocabulary and by bringing interesting activities for these three aspects of vocabulary learning.

Motivation and Learners' Background

Motivation plays a pivotal role in vocabulary learning, significantly influencing students' ability to acquire, retain, and effectively use new words. Motivated learners tend to engage more actively in vocabulary activities and show greater long-term retention. To nurture this motivation, teachers should foster a dynamic and interactive learning environment that emphasizes the practical application of vocabulary in meaningful, real-life contexts.

Given the influence of socio-economic status and prior schooling on language acquisition, it is essential to provide tailored support to bridge learning gaps for students from less privileged backgrounds. This can be achieved through targeted interventions such as offering supplementary learning resources, establishing mentorship programs, and implementing differentiated instruction strategies. By ensuring equitable access to vocabulary development opportunities, educators can create a more inclusive environment where all students, regardless of their socio-economic background, can thrive and build a strong foundation for language proficiency.

Socio-Economic Status and Schooling Impact Students' Vocabulary Knowledge

The socio-economic status and schooling background of foundation students at the Preparatory Studies Center, UTAS Ibra, have a significant impact on their vocabulary knowledge. Students from higher socio-economic backgrounds or well-resourced schools often arrive with a stronger vocabulary base, giving them an advantage in language learning. To address these disparities, it is vital to beef up the existing supportive measures, such as Academic Support Program (ASP) and writing center initiatives. In addition, access to diverse vocabulary resources, and mentorship programs should also be introduced. Vocabulary log should be more intensified, and students should prepare and maintain a bank of verbs with three inflexions (V1, V2 and V3). These initiatives can help level the playing field and ensure that all students, regardless of their background, have an equal chance to succeed in building their vocabulary.

Use of Technology

The integration of technology, such as the MyELT app, has proven highly effective in enhancing vocabulary learning for foundation students by providing interactive and visually appealing methods that sustain their interest and engagement. These digital tools allow students to learn at their own pace, reinforce vocabulary through repetition, and access multimedia content that aids memory retention. Additionally, gamified elements such as quizzes and challenges foster motivation and encourage students to practice more frequently.

However, to fully leverage the potential of these technological resources, greater institutional support is essential. This includes investing in premium educational apps that offer comprehensive content tailored to different proficiency levels, including A1 and A2. Equally important is the provision of continuous training for teachers on how to effectively integrate digital tools into their teaching strategies, ensuring that technology complements traditional methods rather than merely replacing them.

Reliable access to technological infrastructure, such as high-speed internet, sufficient devices, and tech support, is crucial for seamless implementation. Educational institutions must prioritize the development of a well-equipped digital learning environment to eliminate technical barriers that may hinder student engagement and learning continuity.

In addition, tutors need to play an active role in monitoring and guiding students' progress with vocabulary acquisition and retention. This includes tracking student engagement on apps, analyzing data on vocabulary usage, and providing timely feedback to reinforce learning. Personalized interventions based on app-generated insights can help address specific learner challenges and ensure more effective outcomes. Through such coordinated efforts, technology can significantly enhance vocabulary teaching and learning outcomes, ultimately leading to better language proficiency among foundation students.

Peer Interaction

Peer interaction is highly effective for vocabulary development among foundation students. However, Preparatory Studies Center, UTAS Ibra needs to conduct vocabulary games other than Kahoot or WordWall exercises which have been played during lessons. Department or level wise vocabulary contests should be propagated and administered to encourage teams or peers of students to participate. In this way, more interaction can be promulgated. By pairing stronger students with those who need more support, both groups benefit stronger students by reinforcing their knowledge through teaching, while weaker students gain from peer assistance and increased

exposure to vocabulary use in context. This collaborative learning approach fosters a supportive environment, encourages active communication in English, and promotes mutual improvement. To maximize its effectiveness, structured peer activities and teacher monitoring can ensure meaningful and productive interactions.

Curriculum

The current curriculum at the Preparatory Studies Center, UTAS Ibra, has the issue of higher graded vocabulary set for the lessons. However, the focus should be on shifting student attitudes rather than making extensive curriculum changes. The current curriculum already includes comprehensive vocabulary enhancement components though some of the vocabulary items presented are far above the level of the learners. Many students lack the motivation and commitment to engage fully with these resources. Efforts should be directed towards encouraging students to value vocabulary learning, perhaps through awareness sessions, motivational workshops, and practical applications of language skills. By fostering a positive and proactive approach to vocabulary acquisition, students can better utilize the existing curriculum to improve their language proficiency.

Enrichment Schemes

The existing vocabulary enrichment schemes at the center, such as vocabulary logs, are valuable tools for reinforcing word retention and usage. However, their effectiveness is often hindered by students' lack of self-motivation. While motivated students actively engage with these logs, writing their own sentences and reviewing vocabulary regularly, others merely complete the task superficially. To address this, strategies such as gamifying vocabulary logs, incorporating peer reviews, and providing incentives for consistent usage could help boost student engagement and make vocabulary learning a more dynamic and rewarding experience.

Limitations of the Study

The present study is delimited to the Omani students of Foundation level – 1 and level 2, University of Technology and Applied Sciences, Ibra only. In addition to this, the study considered student's vocabulary knowledge vis-à-vis English language. The study was carried out at a single college. The responses of the students during the survey were based on their experience, and exposure to learning environments and activities related to their knowledge acquisition throughout their education. Since these aspects were likely to be different for different respondents, it was considered as a limitation. Lack of complete control over the factors affecting students' response was also a limitation. The personal habits of the students which affect their response were not under control of the researchers.

Funding: The research doesn't incur any funds to conduct the study and publication.
Acknowledgments: The researchers acknowledge the support of Dr. Hilal Al Hajri, the HoC of PSC, UTAS-Ibra and the faculty members and students who participated in the study.
Conflicts of Interest: The authors declare no competing financial or personal interests that could have appeared to influence the work reported in this paper.
Authenticity: This research paper is written with full authenticity of its content and is done basically with the data collected from the participants.

Artificial Intelligence Statement: In conduction of the study and preparation of this research paper, the authors had minimal assistance from AI to check information. However, with the nature of this study, this dependence needs to be limited.

About the Authors

Dr. Yasmin Banu Abdul Majeed holds a doctorate degree in Applied Linguistics from RTM Nagpur University, Nagpur, India. She has over fifteen years of English teaching experience at different levels. Her research interest lies in the field of Applied Linguistics, ESL/EFL and Language Assessment. Currently, she is working at the University of Technology and Applied Sciences, Ibra, Oman, as a senior lecturer.https://orcid.org/0009-0000-1678-9016

Dr. Firos Kalarikkal is a writer and well experienced ELT professional who has been in the field for the last 20 years. He holds a Ph.D in ELT, CELTA, a Master's in English Language and Literature, and a B.Ed. in Teaching English. He currently serves as a senior lecturer at the University of Technology and Applied Sciences (UTAS) at Ibra in the Sultanate of Oman. He has authored a book, book chapters, and published research papers. His area of interest in research is assessment schemes, approaches, and trends in English Language Teaching. https://orcid.org/0000-0001-6533-7972

References

Al-Issa, A., & Al-Bulushi, A. (2012). English language teaching reform in Sultanate of Oman: The case of theory and practice disparity. *Educational Research for Policy and Practice*, *11*(2), 141–176.

Alqahtani, M. (2015). The importance of vocabulary in language learning and how to be taught. *International Journal of Teaching and Education*, *3*(3), 21–34.

Al-Mahrooqi, R., & Denman, C. (2018). Issues in teaching English in the Arab world. In R. Al-Mahrooqi & C. Denman (Eds.), *English education in Oman: Current scenarios and future directions* (pp. 15–32). Springer.

Behforouz, B., & Al Gaithi, A. (2022). Omani EFL learners' vocabulary learning strategies: An investigation into usage and effectiveness. *International Journal of Language Education*, *6*(2), 45–60. https://doi.org/10.18205/ijle.v6i2.1559

Benge, C., & Robbins, M. E. (2022). Integrating learning strategies for EFL learners' vocabulary enrichment and retention. *LingCure Journal of Language, Literature and Culture*, *6*(1), 384–398.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*(2), 213-238. https://doi.org/10.2307/3587951

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE.

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

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Godwin-Jones, R. (2011). Emerging technologies: Mobile apps for language learning. *Language Learning & Technology*, *15*(2), 2-11.

Graves, M. F. (2006). *The vocabulary book: Learning and instruction*. Teachers College Press.

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, *51*(3), 539-558. https://doi.org/10.1111/0023-8333.00164

IBM SPSS Statistics (Version 18.0).

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, *22*(1), 1–26. https://doi.org/10.1093/applin/22.1.1

Milton, J. (2009). *Measuring second language vocabulary acquisition*. Multilingual Matters.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, *63*(1), 59–82. https://doi.org/10.3138/cmlr.63.1.59

Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.

Nation, I. S. P. (2016). The four strands: A practical framework for teaching listening, speaking, reading, and writing. *TESL-EJ*, *19*(4). https://doi.org/10.55593/ej.19.4.921

Newton, J., & Nation, I. S. P. (2020). *Teaching ESL/EFL listening and speaking* (2nd ed.). Routledge.

Read, J. (2000). *Assessing vocabulary*. Cambridge University Press.

QSR International Pty Ltd. (2023). *NVivo* (Version 14). QSR International Pty Ltd. https://www.qsrinternational.com/nvivo/

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*(1), 20–27. https://doi.org/10.1016/j.tics.2010.09.003

Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, *12*(3), 329–363. https://doi.org/10.1177/1362168807086821

Schmitt, N. (2010). *Researching vocabulary*. Palgrave Macmillan

Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, *56*(1), 72–110. https://doi.org/10.3102/00346543056001072

Stockwell, G. (2010). Using mobile phones for vocabulary activities: Examining the effect of platform. *Language Learning & Technology*, *14*(2), 95-110.

Storch, N. (2013). *Collaborative writing in L2 classrooms*. Multilingual Matters.

Syed, Z. (2003). The sociocultural context of English language teaching in the Gulf. *TESOL Quarterly*, *37*(2), 337–341. https://doi.org/10.2307/3588226

Thornbury, S. (2002). *How to teach vocabulary*. Longman.

Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition*, *27*(1), 33–52. https://doi.org/10.1017/S027226310505002X

Webb, S., & Nation, I. S. P. (2017). *How vocabulary is learned*. Oxford University Press.

Wilkins, D. A. (1972). *Linguistics in language teaching*. Edward Arnold.

Zamanian, M., Hojatpour, H., & Heydar-Tabaee, S. (2022). Enhancing English vocabulary retention among Iranian trainee teachers through literature-based interactive input-output tasks. *Arab World English Journal (AWEJ)*, *13*(4), 62–74

Appendix A

Students' Questionnaire

Section 1. Demographic Information

1. **What is your age?**
ما هو عمرك؟
o Under 18
أقل من 18
o 18-20
18-20
o 21-23
21-23
o 24 and above
24 فأكثر

2. **What is your current level of study?**
ما هو مستواك الدراسي الحالي؟
o Foundation Level 1
المستوى التأسيسي 1
o Foundation Level 2
المستوى التأسيسي 2

3. **How long have you been studying English?**
منذ متى وأنت تدرس اللغة الإنجليزية؟
o Less than 3 years
أقل من 3 سنوات
o 3-5 years
3-5 سنوات
o 5-7 years
5-7 سنوات
o More than 7 years
أكثر من 7 سنوات

4. **How would you describe your exposure to English outside the classroom?**
كيف تصف تعرضك للغة الإنجليزية خارج الفصل الدراسي؟
o Minimal
ضئيل
o Moderate
متوسط
o Frequent
متكرر
o Extensive
كبير

5. **Have you participated in any formal vocabulary enhancement programs before?**
هل شاركت في أي برامج رسمية لتعزيز المفردات من قبل؟
o Yes
نعم
o No
لا

Section 2. **Current Vocabulary Learning Methods**

6. How often do you engage in vocabulary-building activities?
كم مرة تنخرط في أنشطة بناء المفردات؟

o Daily
يوميًا
o Weekly
أسبوعيًا
o Monthly
شهريًا
o Rarely
نادرًا

7. **Which method do you primarily use to learn new vocabulary?**
ما هي الطريقة الأساسية التي تستخدمها لتعلم مفردات جديدة؟
o Flashcards
بطاقات الذاكرة
o Vocabulary lists
قوائم المفردات
o Reading exercises
تمارين القراءة
o Interactive apps
التطبيقات التفاعلية
o Other
أخرى

8. **How effective do you find your current vocabulary learning methods?**
ما مدى فاعلية طرق تعلم المفردات الحالية التي تستخدمها؟
o Very effective
فعالة جدًا
o Somewhat effective
فعالة إلى حد ما
o Neutral
محايدة
o Somewhat ineffective
غير فعالة إلى حد ما
o Very ineffective
غير فعالة جدًا

9. **Do you use any specific strategies to remember new vocabulary?**
هل تستخدم أي استراتيجيات محددة لتذكر المفردات الجديدة؟
o Yes
نعم
o No
لا

10. **If yes, which strategies do you use?**
إذا كانت الإجابة بنعم، فما هي الاستراتيجيات التي تستخدمها؟
o Memory
الذاكرة
o Association with images
الربط بالصور
o Repetition
التكرار
o Contextual usage
الاستخدام في السياق
o Other
أخرى

ISSN: 2229-9327

Section 3. **Impact of Vocabulary Enhancement Schemes**

11. **To what extent do vocabulary enhancement schemes like vocabulary logs improve your writing skills?**
إلى أي مدى تحسن برامج تعزيز المفردات، مثل سجلات المفردات، مهاراتك الكتابية؟
o Very much
كثيرًا جدًا
o Somewhat
إلى حد ما
o Neutral
محايد
o Slightly
قليلاً
o Not at all
ليس على الإطلاق

12. **How often do you use newly learned vocabulary in your writing assignments?**
كم مرة تستخدم المفردات التي تعلمتها حديثًا في مهامك الكتابية؟
o Always
دائمًا
o Often
غالبًا
o Sometimes
أحيانًا
o Rarely
نادرًا
o Never
أبدًا

13. **How confident are you in using new vocabulary in your writing?**
ما مدى ثقتك في استخدام المفردات الجديدة في كتاباتك؟
o Very confident
واثق جدًا
o Confident
واثق
o Neutral
محايد
o Unconfident
غير واثق
o Very unconfident
غير واثق تمامًا

14. **Do you believe that vocabulary enhancement schemes help you express ideas more clearly in your writing?**
هل تعتقد أن برامج تعزيز المفردات تساعدك على التعبير عن الأفكار بوضوح أكبر في كتاباتك؟
o Strongly agree
أوافق بشدة
o Agree
أوافق
o Neutral
محايد
o Disagree
لا أوافق
o Strongly disagree
لا أوافق بشدة

ISSN: 2229-9327

15. **Have you noticed an improvement in your overall writing grades due to vocabulary enhancement schemes?**
هل لاحظت تحسنًا في درجات كتابتك العامة بسبب برامج تعزيز المفردات؟
o Significant improvement
تحسن كبير
o Some improvement
بعض التحسن
o No change
لا يوجد تغيير
o Decline
تراجع
o Not applicable
غير قابل للتطبيق

Section 4. **Engagement with Vocabulary Enhancement Schemes**

16. **How engaged are you with the vocabulary enhancement activities provided in your course?**
ما مدى تفاعلك مع أنشطة تعزيز المفردات المقدمة في دورتك؟
o Very engaged
متفاعل جدًا
o Engaged
متفاعل
o Neutral
محايد
o Disengaged
غير متفاعل
o Very disengaged
غير متفاعل تمامًا

17. **How often do you participate in additional vocabulary enrichment activities outside of class?**
كم مرة تشارك في أنشطة إضافية لإثراء المفردات خارج الصف؟
o Frequently
بشكل متكرر
o Occasionally
أحيانًا
o Rarely
نادرًا
o Never
أبدًا

18. **How satisfied are you with the variety of vocabulary enhancement activities available to you during your course?**
ما مدى رضاك عن تنوع أنشطة تعزيز المفردات المتاحة لك خلال دورتك؟
o Very satisfied
راضٍ جدًا
o Satisfied
راضٍ
o Neutral
محايد
o Dissatisfied
غير راضٍ
o Very dissatisfied
غير راضٍ تمامًا

19. **Do you think the vocabulary development activities in MyELT/Moodle align well with your writing needs?**
هل تعتقد أن أنشطة تطوير المفردات في MyELT/Moodle تتماشى مع احتياجاتك في الكتابة؟
o Strongly agree
أوافق بشدة
o Agree
أوافق
o Neutral
محايد
o Disagree
لا أوافق
o Strongly disagree
لا أوافق بشدة

20. **How likely are you to recommend the vocabulary improvement schemes to your peers?**
ما مدى احتمالية أن توصي ببرامج تحسين المفردات لزملائك؟
o Very likely
من المحتمل جدًا
o Likely
محتمل
o Neutral
محايد
o Unlikely
غير محتمل
o Very unlikely
غير محتمل على الإطلاق

5. **Suggestions for Improvement**

21. **What improvements would you most like to see in the current vocabulary enhancement schemes?**
ما هي التحسينات التي تود رؤيتها في برامج تعزيز المفردات الحالية؟
o More interactive activities
المزيد من الأنشطة التفاعلية
o Greater variety of resources
تنوع أكبر في الموارد
o More focus on writing application
تركيز أكبر على تطبيق الكتابة
o Regular feedback on progress
ملاحظات منتظمة على التقدم
o Other
أخرى

22. How do you prefer to receive vocabulary enhancement content?
كيف تفضل تلقي محتوى تعزيز المفردات؟
o Online platforms
المنصات الإلكترونية
o Printed materials
المواد المطبوعة
o In-class activities
الأنشطة الصفية
o One-on-one sessions
جلسات فردية
o Other
أخرى

23. **How helpful would additional support (e.g., tutoring or workshops) be for your vocabulary enhancement?**
ما مدى فائدة الدعم الإضافي (مثل التدريس الخصوصي أو الورش) في تعزيز المفردات؟
o Very helpful
مفيد جدًا
o Helpful
مفيد
o Neutral
محايد
o Not very helpful
غير مفيد جدًا
o Not helpful at all
غير مفيد على الإطلاق

24. **Would you be interested in participating in a peer-led vocabulary enhancement group?** هل تهتم بالمشاركة في مجموعة لتعزيز المفردات يقودها الأقران؟
o Yes
نعم
o No
لا
o Maybe
ربما

25. How would you rate the overall effectiveness of the current vocabulary enhancement schemes in improving your writing skills?**
كيف تقيم فعالية برامج تعزيز المفردات الحالية في تحسين مهاراتك الكتابية؟
o Excellent
ممتاز
o Good
جيد
o Average
متوسط
o Poor
ضعيف
o Very Poor
ضعيف جدًا

ISSN: 2229-9327

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp.64-89
DOI: https://dx.doi.org/10.24093/awej/vol16no4.4

The Use of Smartphone Apps for Informal Language Learning Among EFL Students: Focusing on Students' Choices, Purposes, and Perceptions

Maha Bader Alghasab

English Department, College of Basic Education, The Public Authority for Applied Education and Training (PAAET), Kuwait

Email: mb.alghasab@paaet.edu.kw

Received:09/11/2025 Accepted:11/04/2025 Published: 12/10/2025

Abstract

The advancement in mobile technologies has paved the way for personalized and self-directed language learning beyond conventional classroom settings. Despite the availability of numerous studies on mobile-assisted language learning (MALL) in the higher education context, there has been a lack of research exploring its potential within the informal learning context in Kuwait. This paper, therefore, aims to (1) investigate how English as a Foreign Language (EFL) students use MALL to develop their language skills in informal learning settings and (2) investigate their perceptions and attitudes towards using MALL. Adopting a mixed methods design, the research involved 54 undergraduate students, majoring in English language, at the College of Basic Education in Kuwait. Questionnaires were distributed to all participants, and semi-structured interviews were conducted with 20 students. The findings revealed that students generally held positive perceptions towards using mobile apps in learning English. YouTube, TikTok, and Instagram emerged as the most frequently used apps for informal English language learning. Additional findings indicated that among the language skills developed through smartphone apps, writing is the least targeted skill. In contrast, other skills, such as listening and enhancement of vocabulary knowledge, receive comparatively greater emphasis. These findings underscore the significant role of smartphone apps in enhancing students' engagement and motivation in language learning, extending beyond the traditional classroom boundaries.

Keywords*:* English as a foreign language, informal learning, higher education, mobile assisted language learning, smartphone apps

How to cite: Alghasab, M. B. (2025). The Use of Smartphone Apps for Informal Language Learning Among EFL Students: Focusing on Students' Choices, Purposes, and Perceptions. *Arab World English Journal*, *16*(4): 64-89. https://dx.doi.org/10.24093/awej/vol16no4.4

Introduction

MALL has had a tremendous impact on language learning pedagogy, with students no longer depending on their teachers as the primary source of exposure to the target language (Burston & Giannakou, 2022; Klimova, 2018). MALL refers to the use of handheld mobile devices such as iPads, smartphones, tablets, and similar technologies to support language learning (Stockwell, 2022). It focuses on leveraging the mobility, connectivity, and multimedia capabilities of mobile devices to promote language learning. In its broader meaning, MALL is not a stable concept referring to movement across locations, but rather it enables time-shifting and boundary crossing (Kukulska-Hulme, 2009).

Smartphone applications, as examples of MALL, have had a significant impact on the development of language skills in informal online language learning contexts beyond the classroom boundaries (Kusyk et al., 2025). Empirical studies have reported the effectiveness of using smartphone applications on the development of language skills, including writing (AlSaied & Akhtar, 2021; Awada, 2016), speaking (Machmud & Abdulah, 2017), and reading (Naderi & Akrami, 2018). Furthermore, research findings indicated that smartphone applications promote learners' autonomy and self-directed study (Roy & Gandhimathi, 2025), vocabulary knowledge (Godwin-Jones, 2011; Kacetl & Klimova, 2019; Nami, 2020; Xodabande et al., 2022), and willingness to communicate (Zhao et al., 2024). Previous studies have also shown that students held positive attitudes towards using smartphone applications and tended to use them to practice listening and speaking more often than reading and writing skills (Bin-Hady & Al-Tamimi, 2021) and that dictionary and lexical apps were the most widely used apps (Nami, 2020).

Despite these promising findings, other studies have reported challenges to using smartphones, indicating that EFL learners tend to use them primarily for information gathering or communication purposes rather than developing their language skills (Metruk, 2022). Furthermore, there has been an observed pattern of students focusing more on receptive activities, such as reading emails, listening to music, and watching videos, rather than focusing on productive/interactive activities, such as writing emails or communicating with others using the target language (Alzougool & AlMansour, 2017). These practices have been attributed to students' lack of readiness and unfamiliarity with using smartphones in the learning of English.

Similarly, empirical studies conducted in Kuwait have documented the widespread use of smartphones among students and instructors in higher education contexts (Al-Hunaiyyan et al., 2016; Alzougool & AlMansour, 2017; Dashti & Aldashti, 2015) and secondary school students (Dashti & Abdulsalam, 2025), highlighting patterns of usage that traverse academic and social domains. The findings of these studies have shown that, although students and instructors generally have positive perceptions of mobile learning, particularly when using smartphones, their actual use in teaching and learning tends to be limited. Findings from more recent studies conducted by Dashti and Abdulsalam (2025), with Kuwaiti secondary school students and teachers, have shown that social media platforms developed reading, listening, and speaking skills, respectively, with no significant effect on writing.

While these regional studies provided promising results, their focus tended to be limited to mobile learning broadly (e.g., Al-Hunaiyyan et al., 2016) or within formal courses (e.g., Dashti & Abdulsalam, 2025). Given the swift pace of technological advancements and the rise of informal language learning contexts (Inayati et al., 2024), there is a need to re-examine students' language learning practices and perceptions towards smartphone apps in Kuwait, where the use of smartphones is widespread and continually evolving. This study aims to fill this gap in the

literature by exploring how undergraduate students use smartphones in informal, everyday contexts for language learning and their perceptions about using MALL in learning the English language. To achieve these aims, the following research questions are proposed :

1. What types of smartphone language learning apps do students use?
2. For what language learning purposes are these apps used?
3. What are students' perceptions of using smartphone apps to develop their English language skills?

Literature Review

Informal Language Learning

In adult education, the learning process can occur in three different forms: formal education, non-formal education, and informal education (Livingstone, 2001). Formal education is typically teacher-centered, occurring in a structured and planned manner, and involves a pre-established body of knowledge (e.g., learning in schools). Non-formal education refers to studying voluntarily to acquire further knowledge and skills with the assistance of an instructor (e.g., workshops). Informal education involves informal learning, i.e., "any activity involving the pursuit of understanding knowledge or skill which occurs without the presence of externally imposed curricular criteria" (Livingstone, 2001, p.4). It also includes self-directed learning, which refers to "learning that is undertaken on the learner or learners' own terms without either prescribed curricular requirements or a designated instructor" (Livingstone, 2001, p.3). Informal learning recognizes the social significance of learning from other people in a variety of settings, such as the workplace, rather than formal education or training (Eraut, 2004; Lave & Wenger,1991). Knowles (1970), in his pioneering work, emphasized the importance of self-directed and experiential learning outside formal educational settings. He argued that people learn best when they have control over their learning process and can apply their knowledge in real-life situations.

In a language learning context, informal learning posits that language learning occurs naturally and incidentally outside the structured classroom settings through everyday interaction and experiences with native speakers or through exposure to media and cultural content (Benson, 2001). It challenges the traditional, formal, and instructed classroom and proposes that learning a language can also occur in naturalistic and informal settings outside the classroom context. In some cases, informal language learning occurs unintentionally or even unconsciously as a byproduct of activities where learning is not the primary goal. Lightbown and Spada (2001) argued that while formal learning focuses on teaching the structural aspect of the language itself, informal language learning emphasizes meaning focus since it enables learners to develop their communicative skills by engaging in meaningful interactions using the target language (Larsen-Freeman & Long, 1991). It also promotes cultural awareness and intercultural competence since it exposes learners to different cultural norms, values, and perspectives (Liddicoat, 2007). Benson (2001) discussed four main key dimensions when discussing informal learning: location (where learning occurs), formality (how structured/unstructured the learning is), pedagogy (types of teaching and learning practices), and locus of control (who controls the learning: the learner, teacher, or others). Kusyk et al. (2025) explained that the informal language learning term is a broad term that includes language acquisition and usage and takes place not just beyond the classroom in both online and offline contexts, but also without the use of materials designed explicitly for language learning. The following section narrows the focus to the online context as one context for informal language learning.

Informal Online Language Learning

As technologies continue to advance, investigating online informal language learning opportunities has become a significant topic in the field of MALL (Sockett, 2014; Soyoof et al., 2023). Stemming from research on how adults learn naturally in workplace settings, online informal learning can include "situations of unintentional, incidental or natural language exposure, in which learners perform activities in the target language for purposes other than learning" (Inayati et al., 2024, p.2). In this paper, the term online informal language learning refers to all intentional, unintentional, and incidental language learning activities that learners engage in using smartphone apps to improve their English language beyond their formally organized language classrooms. The ubiquity of mobile devices has indeed expanded opportunities for informal online language learning by providing constant, flexible access to a variety of digital language resources and social platforms. The following section discusses MALL and presents related empirical research findings.

Mobile Assisted Language Learning (MALL)

MALL is a subdivision of Computer- Assisted Language Learning (CALL), which refers to the learners' capacity to leverage different devices (e.g., mobile phones, smartphones, tablets, iPad, or laptops) to study language at any time and from anywhere (Duman et al., 2015). It refers to learning that is happening across multiple contexts, through social and content interactions, using personal and electronic mobile devices (Crompton, 2013), which enables spontaneous access to information and interactivity across different contexts of use (Yang, 2013). MALL has transformed the learning process since it has removed traditional barriers of location and time, offering more opportunities for informal online language learning. The effectiveness of MALL in improving language skills has been well documented in previous studies (see review by Burston & Giannakou, 2022), with smartphones identified as the most widely utilized mobile devices for this purpose (Garzon et al., 2023; Lai & Zheng, 2017; Klimova, 2018; Peng et al., 2021). Furthermore, smartphones support self-regulated language learning beyond traditional classroom settings, enabling learners to engage autonomously with authentic language content (Chung & Lo, 2025). Kessler et al. (2025) classified MALL into two main categories: formal classroom-based learning, in which mobile devices are used in educational settings, or informal opportunities using mobile apps and commercial platforms. The focus of this paper is on the second category, i.e., using smartphone apps by students to informally learn the English language. The following section presents research findings about using smartphone apps in language learning contexts.

Smartphone Apps and Language Learning

Reviewing the literature on the use of smartphone apps revealed that studies either focus on apps that offer exercises based on grammar translation and behaviorism principles, such as Duolingo, or apps that focus primarily on constructivist theory and communicative language learning, such as Hello Talk (Zhao et al., 2024). In their systematic review of commercial MALL apps, Kessler et al. (2025) reported that previous research included empirical investigations for several apps, including Duolingo, Babbel, Busuu, Speaky, Hello Talk, Rossetta Stone, with Duolingo being the most frequently studied app. This finding has also been reported in Shortt et al. (2021), as it has been found that Duolingo was the most dominant and influential mobile learning application on the market for language learning.

Studies that focused on apps that present behaviouristic principles have reported positive effects of using apps such as Quizlet, Duolingo, and Babbel (e.g., Chung & Lo, 2025; Kessler, 2023; Zhao et al., 2024) on language development. These apps typically employ core behaviorist learning techniques such as retrieval practice, feedback, and repetition, which could possibly contribute to effective vocabulary retention, grammar acquisition, and skills development (see review Kesseler et al., 2025). For example, Chung and Lo (2025) reported the effectiveness of using Duolingo in promoting learners' autonomy and motivation. Furthermore, they found that using Duolingo improved discrete language skills such as vocabulary and grammar through repetition and feedback. Kessler et al. (2025) reported similar results, indicating the effectiveness of Duolingo in improving various language skills (e.g., reading, writing, speaking, listening, vocabulary, and grammatical competence). Their comparative design study demonstrated the effectiveness of the Babbel app in promoting grammatical knowledge, cultural awareness, and speaking /pronunciation compared to the Duolingo app. Kessler (2023) investigated the implementation of a weekly reflective e-journal writing activity on students' metacognitive awareness. Results have shown a positive effect of using Duolingo and reflective writing on students' more profound reflections in different domains, including target language linguistic features. Chung and Lo (2025) found that Quizlet, which is also behaviouristic in nature, similarly supports learners' vocabulary retention and confidence through structured, autonomous, and frequent practice. Zhao et al. (2024) demonstrated that using Duolingo was helpful in developing learners' willingness to communicate in the target language.

Other studies have highlighted the distinct advantages of smartphone apps designed around constructivist and socially interactive frameworks. Unlike behaviouristic apps that emphasize repetitive drills, these constructivist, socially oriented apps foster student collaboration, communication, and active meaning-making through digital interaction. Examples of these apps include WhatsApp (Albogami & Algethami, 2022; Andujar, 2016; Awada, 2016), Telegram (Aghajani & Adloo, 2018; Nadei & Akrami, 2018), Hello Talk (Chung & Lo, 2025; Zhao et al., 2024), Cambly (Alshammary, 2020), and WeChat (Dai & Wu, 2022). Findings of previous research have shown the effectiveness of using these apps to promote language skills. For example, studies of Awada (2016), Andujar (2016), and Alikbari and Mardani (2022) found that compared with regular writing instruction, using WhatsApp in writing instruction has increased learners' motivation, reduced their anxiety levels, and improved their writing skills. In developing speaking skills and pronunciation, studies of Dai and Wu (2022), Alshammary (2020), Albogami and Algethami (2022) reported the positive effect of using WeChat, Cambly and WhatsApp respectively. Dai and Wu (2022) found that the exchange of feedback using WeChat improved learners' pronunciation, while Alshammary (2020) found that Cambly had a positive influence on students' speaking proficiency, motivation, anxiety levels, speaking opportunities, autonomy, social relationships, and cultural awareness. Similarly, Albogami and Algethami (2022) reported that using WhatsApp not only improved students' speaking performance but also increased their motivation and attitudes towards speaking and decreased their anxiety and fear of making mistakes. Naderi and Akrami (2018) reported that reading comprehension instruction via Telegram is efficacious in improving reading skills, while Aghajani and Adloo (2018) concluded that Telegram encourages cooperative learning and enhances students' writing skills.

Other studies have highlighted the role of various smartphone social media platforms in supporting language learning. YouTube, Instagram, and TikTok are among the platforms found to offer informal and immersive learning environments. Shanthi et al. (2025) and Tran (2023) found

that learners who use TikTok, particularly the short videos, as supplementary language learning tools reported higher motivation levels and more positive attitudes toward vocabulary learning. Likewise, Aldossari (2024) found that TikTok is effective in enhancing students' speaking skills.

To sum, empirical studies have found that smartphone apps help to develop all language skills, primarily retaining new vocabulary (Burston & Arispe, 2022; Kessler et al., 2025; Klimova, 2018; Kucuk & Daskan, 2024), writing and listening (Peng et al., 2021), as well as speaking skills (Li, 2024; Sari & Ersanli, 2024). Kessler et al. (2025) found that the focus of 32 % of previous empirical research was primarily directed toward vocabulary development, with few investigating listening skills (17.9%) and grammar/speaking skills (14.3%). Studies that focus on reading and writing were less common.

The following section narrows the focus on MALL studies in the Kuwaiti sociocultural context.

The Case of Kuwait

In Kuwait, there has recently been considerable interest in mobile technologies in education generally (Al-Hunaiyyan et al., 2016; Alzougool & AlMansour, 2017) and in language instruction particularly (Dashti & Aldashti, 2015; Dashti & Abdulsalam, 2025). Findings from this empirical research show that smartphones have become indispensable tools for accessing multimedia content, social networking, and language learning apps that can enhance linguistic skills. Al-Hunaiyyan et al. (2016) and Alzougool and AlMansour (2017) found that while mobile devices, especially smartphones, are widely used by Kuwaiti students, mobile learning adoption is still developing and not yet fully utilized. Students appreciated the convenience and accessibility that mobile devices offer for learning anytime and anywhere. However, the study revealed challenges such as technical issues, lack of infrastructure, limited institutional support, and some resistance from educators who were unfamiliar with mobile learning pedagogy. These challenges impeded the full adoption of mobile learning in Kuwait.

In the language instruction context, Dashti and Aldashti (2015) investigated the attitudes of EFL college students toward mobile learning, focusing on the use of mobile devices for language education. The findings indicated that the majority of students have had positive perceptions of mobile devices as beneficial tools for enhancing their knowledge in areas such as vocabulary and grammar. A more recent study by Dashti and Abdulsalam (2025) explored the influence of social media applications on EFL high school students in Kuwait. The study found that daily use of social media platforms notably improved students' listening, speaking, and reading skills. However, writing skills did not show any significant improvement. In terms of students' usage, the study showed that WhatsApp, Instagram, Snapchat, and YouTube were the most used apps, respectively.

Given the continuous development of smartphone apps, there is an obvious need for more up-to-date research on the use of smartphone applications for language learning in Kuwait. Using mixed methods, as will be discussed in the following section, this study aims to explore usage patterns and students' perceptions of using smartphone apps for developing their language learning in informal settings.

Methodology

The Design of the Study

The study employed a mixed methods design by first gathering quantitative data to explore patterns of smartphone usage and the nature of students' perceptions towards using smartphone

apps. This is followed by collecting qualitative data to provide further detailed insights. This research design integrates the strengths of both quantitative and qualitative approaches, enabling the collection of broad, generalizable data through a questionnaire, while gaining deep, contextualized insights through interviews. Furthermore, a quantitative questionnaire allows the identification of usage patterns and perceptions broadly across participants while the interview explores nuanced explanations and experiences behind the quantitative trends, which strengthens the understanding and interpretation of the results.

Context and Participants

The study was carried out at the College of Basic Education in Kuwait. The participants included 54 undergraduate students majoring in the English language who volunteered to take part in the study. The sample comprised 22 male and 32 female participants, aged between 19 and 23. All participants were enrolled in a four-year teacher education program at the College of Basic Education. Among the 54 pre-service teachers participating in the study, 6% were in their second year, 14% were in their third year, and 34% were in their fourth year of the four-year teacher education program indicating that the majority of participants were in the upper years of their academic training likely reflecting more advanced stages of both theoretical and coursework and practical teaching experience. Throughout their studies, these students undertake a variety of courses related to the English language, linguistics, teaching methodology, curriculum design, and educational psychology. In addition to this, the program includes a supervised practicum component, allowing students to gain hands-on teaching experience.

All participants in this study reported owning a smartphone. Regarding the frequency of their use of language learning applications, the demographic questionnaire questions revealed that 39% of them use smartphone apps daily, 7% use these apps weekly, 2% reported monthly use, and 6% of them reported occasional usage. None stated that they did not use any language learning apps. This result suggests a high level of engagement with smartphone apps among participants.

To select participants, two sampling techniques were employed: convenience sampling for the questionnaire and purposive sampling for the semi-structured interviews. Both are non-probability sampling methods; while the former refers to choosing participants who are most available and willing to participate, the latter refers to intentionally selecting participants because they have specific characteristics relevant to the study (Creswell, 2013). Both sampling methods have advantages since employing convenient sampling is quick, easy, and cost-effective, while purposive sampling helps to gather an in-depth understanding from rich telling cases that can provide further insights to answer the research questions. However, both can introduce significant bias because, in the case of convenience sampling, the sample may not represent the larger population, while purposive sampling may introduce bias because it is based on the researcher's judgment.

Participants were fully informed about the study and provided their consent through a formal consent form before participation. To ensure privacy and confidentiality, all participants' names were changed and anonymized in the data records. All ethical practices, including voluntary participation, the right to withdraw, and secure data storage, were carefully followed to protect participants' rights and privacy.

Data Collection Instruments

Driven by the research questions, two instruments were used to collect the data from participants: a questionnaire and semi-structured interviews. The questionnaire was distributed first via an online link using Google Form, allowing students to provide quantitative data efficiently and conveniently. Subsequently, for deeper qualitative insights, semi-structured interviews were conducted with selected students who presented the extremes of smartphone app usage, i.e., those who used the apps most and least. Together, these instruments facilitated a comprehensive data collection approach combining broad quantitative data with focused qualitative insights.

Prior to designing the questionnaire, previous literature was carefully considered, including Lai and Zheng (2018), who investigated self-directed use of mobile devices for language learning beyond the classroom, and Nami (2020), who explored students' choices and perceptions of educational smartphone apps in higher education. Based on insights from these studies, the current study's questionnaire was crafted to include four main sections (Appendix A). First, demographic information is gathered to provide background information about the participants. The second section included questions on frequency and types of smartphone apps students use, as well as the specific language skills or areas of linguistics knowledge for which they utilize these apps. To explore students' general perceptions of the effectiveness of smartphones for language development, a set of statements was presented in section three using a 5-point Likert scale. These statements were organized around thematic areas identified in the literature: personalization, development of language skills, authenticity of the learning experience, and connectivity. The last section included two open-ended questions that asked students to describe how they use smartphones to enhance their language skills and the challenges they encounter when using smartphone apps. To avoid the possibility of receiving biased responses, some statements were negatively worded and recorded for data analysis. This structured approach aimed to capture both usage patterns and learners' attitudes toward the effectiveness of smartphone apps for language learning, aligning with the research questions and the focus of the current study. Taking the questionnaire was voluntary, and no time limitation was set to let students respond at their convenience.

Prior to data collection, the instrument was piloted to check the validity and reliability. To ensure the validity of the questionnaire, expert validation was employed. The researcher gave the instrument to two experts in the field of applied linguistics to check every item against the objectives of the study. With regard to the reliability of the questionnaire, Cronbach's alpha was calculated. The result indicated a high reliability of 0.75.

The semi-structured interview questions also revolve around similar themes, including questions about describing personal experiences (personalization), students' perceptions of the usefulness of using smartphones for language skills development, authenticity, and connectivity (Appendix B).

Data Analysis

The data analysis involved both quantitative and qualitative approaches. For the quantitative, the questionnaire items were coded using numerical values and analyzed using IBM SPSS 25. Descriptive statistics were employed to illustrate usage patterns and frequency of smartphone app use, as well as to provide an overview of students' perceptions regarding the effectiveness of these apps in enhancing their language skills.

For qualitative data, a thematic analysis approach was employed to identify common themes and patterns in open-ended questions and the interview transcripts (Braun & Clarke, 2006). The author has performed the primary data coding, and to enhance the reliability and reduce bias, an inter-rater reliability check was performed by a colleague. The inter-rater reliability check resulted in a consistency rate of 85%, indicating a high level of agreement in coding among the researchers. The thematic analysis was used to interpret the quantitative findings further and provide insights into participants' perceptions of the apps' usefulness.

Results

Students' Choices and Usage of Smartphone Apps outside the Classroom Context

To answer research question 1, regarding the type of smartphone apps students commonly use and their frequency patterns, it was observed that YouTube (47%), TikTok (27%), and Instagram (23%) were the most widely used apps (see Figure 1). Additionally, 17% of participants reported the use of Google Translate. A small number of students reported using other apps, including Twitter (10%), Snapchat (10%), Duolingo (12%), Learning English vocabulary (4%), BBC Learning English (1%), Quizlet (1%), and Cambly (5%). Despite the popularity of apps such as Orai, Rosetta Stone, The English Club, FluentU, and Babbel, none of the students reported using these apps for language learning.

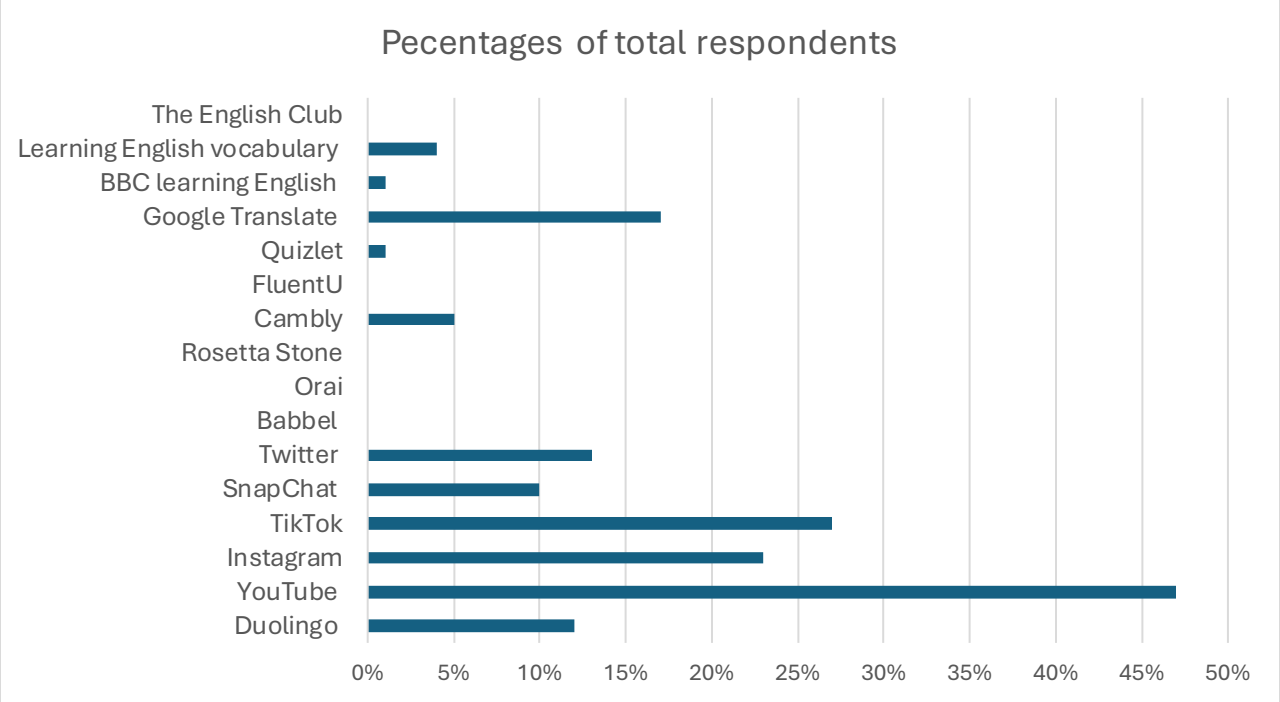


Figure 1. Students' usage frequency of smartphone apps

Figure 2 below indicates that most students used smartphone apps to develop various language skills. Smartphone apps are primarily used to develop listening skills (32%) and vocabulary knowledge (30%), while writing (16%) remains the least targeted skill for improvement among other skills.

Speaking skills are also significant, with 27% of students reporting app use for this purpose. Pronunciation development is reported by 25% of respondents, showing moderate emphasis. Skills such as reading, cultural awareness, and grammatical knowledge attract roughly similar attention, each chosen by about 22-23% of students.

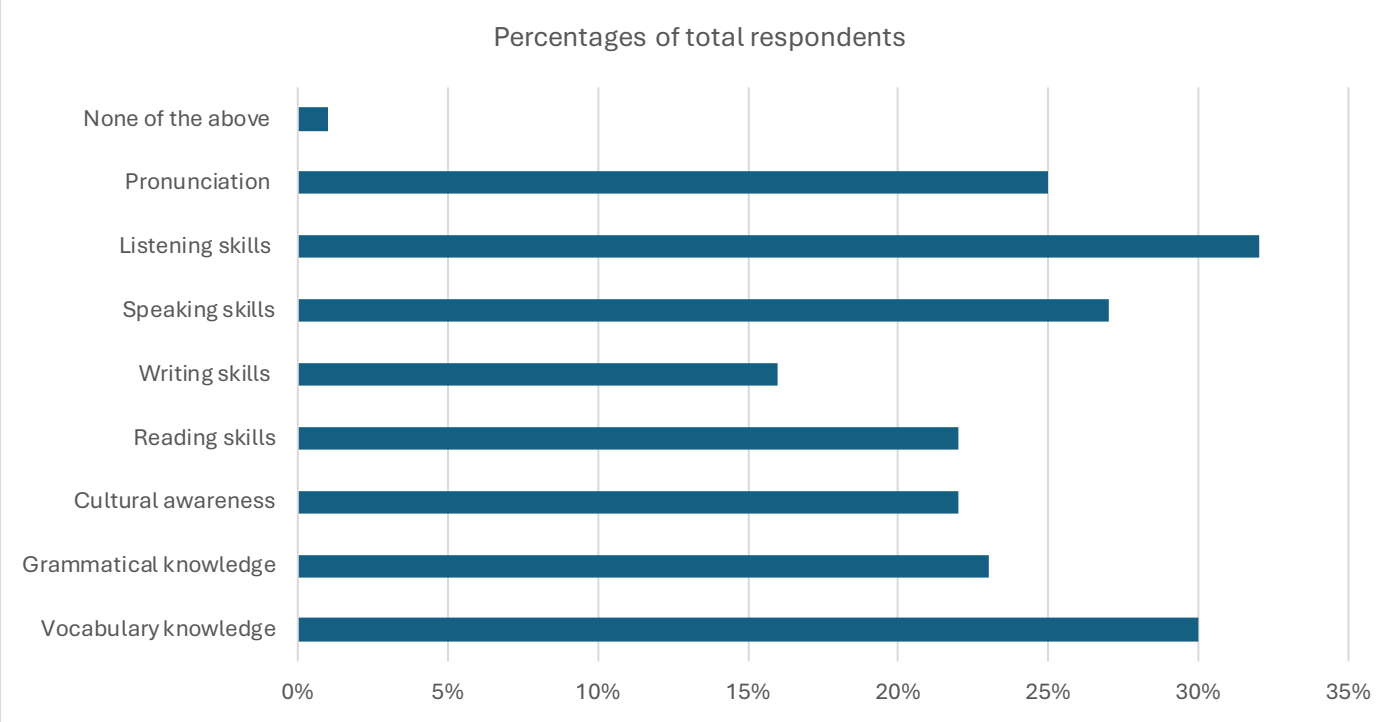


Figure 2. Purposes of using smartphone Apps

Students' Perceptions of Using Smartphone Apps

Overall, questionnaire results demonstrate that learners not only appreciate the accessibility and conveniences of mobile-assisted language learning but also strongly recognize its role in fostering autonomy, extending learning opportunities beyond the class, and supporting tailored self-directed study habits. This implies that smartphone apps are viewed as highly effective tools for enhancing personalized learning experiences in language acquisition. In addition to that, students also believed that smartphone apps are helpful to develop their language skills, expose them to authentic language content, and connect them with other speakers of English. The following sections discuss these findings with reference to the descriptive statistical analysis of the questionnaire.

Smartphone Apps and Personalized Learning

Table 1 indicates that the students strongly perceive smartphone language learning apps as highly effective tools for personalized learning. The majority of students expressed positive attitudes, indicating both agreement and strong agreement across all questionnaire statements. Across all five statements, the mean scores ranged from 3.95 to 4.36 on a five-point scale, signaling high levels of agreement. The highest rated statements (a3 and a5) suggest that learners value the accessibility offered by smartphone apps, fitting learning seamlessly into their daily lives. The high agreement levels observed in statement a5 also indicated students' strong positive beliefs that

smartphone apps support independent learning (mean=4.18), and promote self-paced, personalized learning outside the classroom (mean=4.11). The least strongly rated statement (i.e., a1) still reflects a majority belief (80%) that smartphone apps provide immediate support for learning. The relatively low standard deviation (0.64-077) across statements suggests consistent responses among participants.

Personalization													
		Strongly disagree		Disagree		Neutral		Agree		Strongly agree			
	Questionnaire statements	F	%	F	%	F	%	F	%	F	%	Mean	Std. Deviation
a1	Smartphone language learning apps give immediate support to my language learning	0	0	1	1.8	10	18.2	35	63.6	9	16.4	3.95	0.650
a2	Smartphone language learning apps extend my language learning experience beyond the language class.	0	0	4	7.3	4	7.3	35	63.6	12	21.8	4.00	0.770
a3	Smartphone language learning apps enable me to learn the language at any time and from any place	0	0	2	3.6	1	1.8	27	49.1	25	45.5	4.36	0.704
a4	Smartphone language learning apps enable me to have self-paced, personalized language learning outside the classroom.	1	1.8	1	1.8	3	5.5	36	65.5	14	25.5	4.11	0.737
a5	Smartphone language learning apps provide opportunities to act independently in learning the target language	0	0	0	0	7	12.7	31	56.4	17	30.9	4.18	0.641

Table 1. *Using smartphone apps for personalized learning*

Interviewing students indicated that the majority expressed a strong appreciation for the availability and accessibility of smartphone apps. They emphasized that mobile apps allow them to address their academic, personal, and social needs seamlessly, without being limited by physical location or specific times of day. For example, some students mentioned that smartphone apps allowed them to explore their needs and independently find online resources that enhance their language skills. Often, students find themselves learning in random moments, scrolling through articles and apps, without consciously labeling these activities as learning.

"Being able to access learning materials and explore unlimited videos on YouTube, no matter where I am, helps me to develop my language skills, especially listening and speaking, and reduce stress" (S19)

"For me, the best thing about learning using a smartphone is the flexibility of the learning process, for example, I like to download YouTube movies and while watching these videos, I found myself learning new vocabulary and developing my listening skills" (S21).

"I enjoy developing my English language through the use of smartphone mobile apps, in the car, on a break, or in between classes. I use some apps to listen to English podcasts about a topic that I like, which eventually helped me to have better English" (S8)

Smartphone Apps and Language Skills Development

The data in Table 2 illustrates that students have notably positive perceptions towards using smartphone language learning apps for developing a broad spectrum of linguistic skills, especially vocabulary, reading comprehension, and grammar. Students responded positively to all statements, indicating consensus and confidence in the effectiveness of smartphone apps for developing language skills. As shown, strong agreement is evident for vocabulary development (mean=4.35), reading comprehension (mean=4.18), as well as grammar and writing (mean=4.00). Students' disagreement with b8 and b10, the negatively worded statements, means that most participants believe that smartphone apps can be helpful in improving listening skills and learning accurate pronunciation. Interestingly, the negatively worded statement b8 and b14 received the lowest mean score (mean = 2.24, SD= 1.186 and mean=2.36, SD=0.988 respectively), indicating an overall disagreement with the idea that smartphone apps are ineffective for listening and writing development. Overall, the findings suggest that smartphone apps are widely valued as effective tools for core language skills development.

Table 2. *Using smartphone apps for language learning*

Development of language skills and areas of linguistic knowledge													
		Strongly disagree		Disagree		Neutral		Agree		Strongly agree			
		F	%	F	%	F	%	F	%	F	%	Mean	Std. Deviation
b6	Smartphone Language learning apps can be effective for learning and improving my knowledge of English vocabulary.	0	0	0	0	3	5.5	30	54.5	22	40.0	4.35	0.584
b7	Smartphone Language learning apps can be effective for learning and improving my knowledge of English grammar and writing.	0	0	4	7.3	7	12.7	29	52.7	15	27.3	4.00	0.839
b8	I don't think that Smartphone language learning apps can be effective for improving my listening skills.	17	30.9	21	38.2	7	12.7	7	12.7	3	5.5	2.24	1.186

Development of language skills and areas of linguistic knowledge													
		Strongly disagree		Disagree		Neutral		Agree		Strongly agree			
b9	Smartphone Language learning apps can be effective for improving my reading comprehension skills.	0	0	1	1.8	4	7.3	34	61.8	16	29.1	4.18	0.641
b10	Smartphone Language learning apps cannot be effective for learning the correct pronunciation of English words.	12	21.8	24	43.6	5	9.1	9	16.4	5	9.1	2.47	1.260
b11	Smartphone language learning apps can be effective for improving my speaking skills.	0	0	4	7.3	11	20.0	30	54.5	10	18.2	3.84	0.811
b12	Smartphone language learning apps enhance my interaction with others in the English language.	1	1.8	3	5.5	15	27.3	26	47.3	10	18.2	3.75	0.886
b13	Smartphone language learning apps can be effective in developing my cultural awareness of the foreign language	0	0	3	5.5	16	29.1	22	40.0	14	25.5	3.85	0.870
b14	I don't think that Smartphone language learning apps can be effective for improving my writing skills	10	18.2	25	45.5	10	18.2	10	18.2	55	100.0	2.36	0.988

Interviewing students indicated that (n=15) of students reported that they use smartphone apps to develop their listening skills and vocabulary knowledge by watching YouTube videos with subtitles and listening to podcasts. The following interview excerpts exemplify some students' quotes.

"I think listening to podcasts or watching videos with subtitles is very useful for developing my listening and vocabulary knowledge" (S14)

"I personally use YouTube videos to pick up on brand new vocabulary and listen to the pronunciation of certain words and phrases. I believe that this is the best way to develop my language outside the classroom" (S17).

Some of the interviewees (n=13) mentioned that they use YouTube videos to understand grammar, and they use Google Translate daily to understand the meaning of some challenging vocabulary. The following interview excerpts provide examples for students' responses.

"Occasionally I watch YouTube videos explaining grammar I do not know or using Google Translate for a word I do not know the meaning of, it's around half an hour a day" (S19)

"I use YouTube to watch English videos like vlogs and lessons. It helps me improve my listening and learn how to speak naturally. I also use Google Translate to translate some texts from English to Arabic" (S7)

The majority of students (18) declared that they use reels on TikTok and Instagram to develop their speaking and listening skills. For example, the following students reported their frequent use of TikTok live and Instagram posts to develop their speaking and listening skills as well as vocabulary knowledge. They stated that:

"I found TikTok live interesting and effective. I follow native speakers of English, and when they go live, I immediately join them to listen to different accents like American or British. In this way, I think my language developed effectively" (S8).

" For me, I like to use Instagram frequently, I enjoy following foreigners who speak English as a first language, I enjoy watching makeup tutorials and fashion designers, I noticed that my language is developing rapidly (…) I usually use Instagram to communicate with my foreign friends in English to improve my writing/speaking skills" (S11)

Some students (n=8) declared that they use apps such as Duolingo and audiobooks to develop their grammatical knowledge and reading skills. For example, one of the students enjoys the structured nature of Duolingo's activities, and the other one is interested in navigating audio. The following interview excerpts present some of the students' opinions.

"I use apps like Duolingo to learn new vocabulary, and I like the structured nature of its activity. It is very motivating and shows me my current level, as well as the progress I have made. For me, Duolingo is very effective for language learning" (S5)

" I believe that my reading skills have developed as a result of using audiobooks on the Apple Store. In my free time, before I go to sleep, I like to listen to some stories. I enjoy the interactivity of these apps, and this helped me a lot to develop reading and listening skills" (S13)

Smartphone Apps, Authenticity, and Connectivity

Students' responses to the three statements about smartphones' effectiveness in offering an authentic language learning experience show a clear positive trend with mean scores above 3.50 (see table 3). These generally positive and consistent responses reflect a favorable view of the impact these apps have on offering real-world language engagement. Statement c16 has the highest mean score (3.87), indicating that students feel most empowered to write in English on social media through smartphone apps. Also, respondents generally agree (60%) or strongly agree (10.9%) that the use of these apps adds authenticity to their learning process. Overall, students have positive attitudes about the effectiveness of smartphones in enhancing their English language communication skills by engaging in authentic language experience.

Table 3. *Smartphone apps and authenticity*

Authenticity													
		Strongly disagree		Disagree		Neutral		Agree		Strongly agree			
		F	%	F	%	F	%	F	%	F	%	Mean	Std. Deviation
c15	Smartphone language learning apps enhance my willingness to participate in English language discussions	0	0	3	5.5	16	29.1	32	58.2	4	7.3	3.67	0.695
c16	Smartphone language learning apps enhance my willingness to write in English when participating on social media platforms.	0	0	2	3.6	11	20.0	34	61.8	8	14.5	3.87	0.695
c17	Smartphone language learning apps enhance the authenticity of my language learning resources and activities.	0	0	2	3.6	14	25.5	33	60.0	6	10.9	3.78	0.686

In addition to exposing students to authentic language content and supporting authentic language experience, students valued the affordances of smartphone apps to connect them with other native speakers of English. As shown in Table 4 below, students generally had favourable attitudes towards using smartphone apps to communicate with other people using the target language.

Table 4. *Smartphone apps and connectivity*

Connectivity													
		Strongly disagree		Disagree		Neutral		Agree		Strongly agree			
		F	%	F	%	F	%	F	%	F	%	Mean	Std. Deviation
d18	Smartphone language learning apps enhance my interaction with others in the target language	0	0	2	3.6	7	12.7	39	70.9	7	12.7	3.93	0.634
d19	Smartphone language learning apps enhance my connection with peer learners of the target language	0	0	2	3.6	12	21.8	37	67.3	4	7.3	3.78	0.629
d20	Smartphone language learning apps connect me with speakers of the language	0	0	2	3.6	10	18.2	35	63.6	8	14.5	3.89	0.685
d21	Smartphone language learning apps enhance my connection with the target language community	1	1.8	3	5.5	6	10.9	37	67.3	8	14.5	3.87	0.795
d22	Smartphone language learning apps make my language learning experience in and outside the class more connected	0	0	2	3.6	3	5.5	42	76.4	8	14.5	4.02	0.593

The findings in Table 4 indicate that students perceive smartphone language learning apps as highly effective tools for fostering interaction and connection within various aspects of the language learning process. The highest agreement was recorded for item d22 (mean=4.02, SD=0.593), indicating that learners strongly believe that these apps help them to integrate their language learning experiences both inside and outside the classroom. Likewise, students' positive responses to items d18 (mean=3.93) and d20 (mean=3.89), and d21 (mean=3.87) show consistent agreement that smartphone apps not only enhance the interaction with others, but also facilitate connection with native speakers, and strengthen ties to the target language community. This suggests that most students perceive smartphone apps as having a positive role in fostering connectivity. Overall, students' responses revealed that they highly agree that smartphone apps are powerful tools for facilitating social and communicative connectivity within the language learning context.

Interview data revealed that some students were very positive about using smartphone apps, such as Cambly (n=2), to communicate with native English speakers. In contrast, others

valued the authentic content presented on TikTok, Instagram, and Snapchat about English culture (n=13).

"For me the use of Cambly helped me to connect with native speakers of English and this has helped me a lot to talk freely without feeling anxious"

"I use my iPhone apps to practice English vocabulary, grammar, listening, and speaking outside class. Usually, I follow social media influencers and watch them when they do live on TikTok. I spend hours listening to English speakers talking about different topics, and this definitely developed my language"(S8)

" I like to use Snapchat a lot. It is my favorite compared with other apps. I know it is not specifically designed for learning English, but I found it really effective when I listen and watch real-life snaps, I think it develops my speaking and listening skills" (S21)

Discussion

Based on the analysis of the questionnaire and interview data, findings indicated that students held a strong positive attitude towards using mobile apps for learning the English language informally. Students frequently used apps such as YouTube, TikTok, and Snapchat alongside Google Translate for improving receptive and oral language skills, particularly listening and vocabulary. In contrast, productive skills like writing are less frequently developed through these apps. In-depth analysis also indicated that students valued the effectiveness of smartphone apps in supporting anytime-anywhere personalised learning and the development of language skills. Additionally, students believed that smartphone apps support the informal language learning process by providing unlimited authentic resources and connecting them with native speakers of English.

The findings that students hold positive attitudes towards using mobile apps for informal English language learning align with extensive prior research emphasizing the significant role of MALL in enhancing language (Bin-Hady & Al-Tamimi, 2021; Burston & Giannakou, 2022). Indeed, students' frequent utilization of apps such as YouTube, TikTok, Snapchat, and Google Translate for improving receptive skills, particularly listening and vocabulary, reflects a global trend where smartphone apps provide accessible and engaging platforms for authentic language exposure (Burston & Arispe, 2022; Kessler et al., 2025; Klimova, 2018). In line with previous studies' findings (Shanthi et al., 2025; Tran, 2023), students valued the significant role of TikTok in supporting incidental vocabulary learning. Similar to Al-Hunaiyyan et al. 's (2016) findings, Kuwaiti students who participated in the current study valued video-based social media platforms such as YouTube and Snapchat. Snapchat, while not traditionally designed for developing language skills, assists students in engaging in brief, authentic interactions in the English language, thereby contributing to the development of communicative and intercultural competence in informal settings. Consistent with previous studies' findings (Dashti & Abdulsalam, 2025; Kessler et al., 2025), writing as a productive skill tended to receive less emphasis in informal mobile use, possibly due to the interactive and multimedia-focused nature of these platforms.

Additional findings revealed that students' preference for using social interactive smartphone apps over behaviouristic apps to develop their English language skills informally. This finding reflects broader cultural values that emphasize social connection and community engagement. This aligns with Chung and Lo's (2025) findings with Hong Kong students, who were

gravitated toward apps that offer opportunities for collaboration, peer interaction and meaningful social engagement. Unlike South Korean students, who tended to use of mobile apps for practicing structured activities such as vocabulary drills and grammar exercises, Kuwaiti students use socially oriented smartphone apps such as YouTube, TikTok and Snapchat to engage with English language, mirroring previous studies finding (Dashti & Abdulsalam, 2025). These platforms provide comfortable, anxiety-reduced environment that support informal learning through real life communication. This preference suggests that implementing socially interactive apps into language teaching in Kuwait could better align with learners' natural inclination and promote sustained motivation and effective language development.

Theoretically, the findings indicated that students valued anytime-anywhere language learning via smartphone apps, which align with the theoretical framing of mobile learning as inherently flexible and situated, assisting learners to integrate language practice into their daily lives seamlessly (Chung & Lo, 2025; Duman et al., 2015; Kessler et al., 2025). Students' positive attitudes towards using smartphones to access unlimited authentic materials and to connect with native speakers through smartphone apps echo previous studies' findings that emphasize the role of smartphone apps in providing exposure to real-world language use beyond formal classroom settings (Guo, 2022). The findings also suggest a shift in students' learning behaviour, where learning is no longer restricted to formal classroom contexts but extends into their daily digital practices. This resonates with theories of informal learning and self-directed learning in the language learning field (Benson, 2001), particularly with informal online language learning theories (Soyoof et al., 2023), which argue that MALL empowers learners to take ownership of their learning outside institutional settings. The informal and non-threatening nature of smartphone app learning seems appealing to Kuwaiti students who may feel anxious or inhibited in formal language learning settings.

Conclusion

The current study is an attempt to explore Kuwaiti EFL students' attitudes toward using MALL informally to learn the English language. The main findings suggest that Kuwaiti students generally hold positive perceptions towards using smartphone apps to develop their English, particularly for developing listening and vocabulary. The findings also indicate that students valued the effectiveness of MALL, particularly socially oriented smartphone apps, in supporting the development of language skills, providing opportunities to interact informally with an unlimited number of authentic materials and native speakers of English at any time and from anywhere.

Some limitations have surfaced while conducting this study and have to be acknowledged. First, this study is limited to self-reported data, so this may cause response bias and subjectivity. Second, the small size of only 54 students constitutes another significant limitation for this study, which affects the generalizability of the findings. Third, informal language learning presents challenges in accurately measuring and identifying language development due to its spontaneous and unstructured nature. Consequently, a limitation of this study is its inability to track or quantify any observed progress in language skills over time.

Despite these limitations, there are several implications based on the current study's findings. Educators and curriculum designers should incorporate mobile apps into English-language instruction. There is a need to emphasize the development of both receptive and productive skills through guided and autonomous activities. In addition, educators should provide

explicit guidance for supporting writing skills within the MALL context, which could address the current imbalance in skills development and focus. Workshops, tutorials, or digital literacy modules could help to teach students strategies for vocabulary retention, note-taking, or practicing productive skills using specific smartphone apps. The current study findings suggest that Kuwaiti EFL students favored interacting with native speakers informally on smartphone apps; therefore, encouraging and guiding students to use social media platforms can offer rich opportunities to develop pragmatic and sociolinguistic awareness, which are critical for real-life communication and intercultural competence. Given the diversity and variability of language apps, students need training in digital literacy and critical evaluation. Educators should play a significant role in guiding students to use reliable apps that align with their learning goals and proficiency levels. Given the high smartphone penetration and increasing use of digital platforms for education, designing local apps that integrate culturally relevant content could enhance students' engagement outside the classroom setting. Collaboration between educational institutions and app developers to adapt smartphone apps reflecting Kuwaiti linguistic, cultural, and pedagogical contexts is recommended. These implications could empower Kuwaiti students to harness smartphone apps' full potential not only for English language learning but also for lifelong learning.

Drawing from the findings of the present study, several potential directions for future research emerged. First, future research could investigate the effectiveness of specific smartphone apps such as YouTube, TikTok, Snapchat, or Google Translate on the development of receptive or productive skills using longitudinal studies. Second, exploring teachers' perceptions about the effectiveness of using smartphones could offer a more comprehensive understanding of the effectiveness of MALL in this context. Third, examining the potential integration of MALL, specifically smartphone apps, into the formal English curriculum could provide further insights into the benefits and challenges that students and teachers may face when using smartphone apps.

Funding statement: The author received no financial support for the research, authorship, and /or publication of this article.
Acknowledgment: The author would like to express our gratitude to the participants of our study, as well as to the anonymous reviewers.
Conflict of interests: The authors declare no competing interests.
Authenticity: This manuscript is an original work
Declaration of Generative AI: AI and AI assisted technology, specifically Grammarly, was used in producing this manuscript for grammar and language refinement. The author takes full responsibility for the accuracy and integrity of the content tools were used in preparing this manuscript.

About the author:

Maha Alghasab is an associate professor in the English Department at the College of Basic Education, The Public Authority for Applied Education and Training (PAAET), Kuwait. Her main research interests are in computer-assisted language learning (CALL) and computer-supported collaborative learning (CSCL), with a focus on collaborative writing activities, telecollaborative activities and intercultural communicative competence. She is also interested in exploring how teachers can harness the affordances of technology including Artificial Intelligence (AI) technologies to enhance language learning. Her most recent work investigates teachers' TPACK and the use of AI in language learning.

ORCID: https://orcid.org/0000-0002-3580-0337

References

Aghajani, M., & Adloo, M. (2018). The effect of online cooperative learning on students' writing skills and attitudes through telegram application. *International Journal of Instruction.1*(3), 433–448. https://doi.org/10.12973/iji.2018.11330a

Albogami, A., & Algethami, G. (2022). Exploring the Use of WhatsApp for Teaching Speaking to English Language Learners: A Case Study. *Arab World English Journal (AWEJ) 2nd Special Issue on Covid 19 Challenges* (*2*) 183-201.DOI: https://dx.doi.org/10.24093/awej/covid2.12

Aldossari, A.T. R. (2024). *An Overview of the Utilization of TikTok in Improving Young Saudi Learners' Speaking*. Department of English Language and Literature, Imam Mohammed bin Saud Islamic University, Saudi Arabia (M.A.Thesis). Retrieved from *Arab World English Journal* (ID Number: 304) May 2025: 1-47. DOI: https://dx.doi.org/10.24093/awej/th.304

Al-Hunaiyyan , A., Alhajri, R., & Al-Sharhan, S. (2016). Perception and Challenges of mobile learning in Kuwait. *Journal of King Saud University Computer and Information Sciences*, *30* (2), 279-289.

AlSaied, N., & Akhtar, F. (2021). Media students using mobile phones in the Arabian Gulf to improve English writing and video production skills. *Learning and Teaching in Higher Education: Gulf Perspectives*, *17*(1), 4–15. https://doi.org/10.1108/LTHE-09-2020-0038

Alshammary, M. S. (2020). The impact of using Cambly on EFL University students' speaking proficiency. *English Language Teaching*, *13*(8). https://doi.org/10.5539/elt.v13n8p12

Alzougool, B., & Almansour, J. (2017). The use of Smartphone for learning activities by university students in Kuwait, *Proceedings of Teaching and Education Conferences 4907509. International institute of Social and Economic Sciences*. Retrieved on 26Th Oct, from https://ideas.repec.org/p/sek/itepro/4907508.html

Andujar, A. (2016). Benefits of mobile instant messaging to develop ESL writing. *System. [Crossref]*, *62*, 63–76. https://doi.org/10.1016/j.system.2016.07.004

Awada, G. (2016). Effect of WhatsApp on critique writing proficiency and perceptions toward learning. *Cogent Education. [Crossref]*, *3*(1), Article 1264173. https://doi.org/10.1080/2331186X.2016.1264173

Benson, P. (2001). *Teaching and researching autonomy in language learning*. Pearson Education.

Bin-Hady, W. R. A., & Al-Tamimi, N. O. M. (2021). The use of technology in informal English language learning: Evidence from Yemeni undergraduate students. *Learning and Teaching in Higher Education: Gulf Perspectives*, *17*(2), 107–120. https://doi.org/10.1108/LTHE-09-2020-0037

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Burston, J., & Arispe, K. (2022). Experimental MALL research in SLA: Past, present, and future. In N. Ziegler & M. González-Lloret (Eds.), *The Routledge handbook of technology & SLA*. Routledge. Retrieved from https://www. http://routledge.com/The-

Routledge-Handbook-of-Second-Language- Acquisition-andTechnology/Ziegler-Gonzalez-Lloret/p/book/9780815360773

Burston, J., & Giannakou, K. (2022). MALL language learning outcomes: A comprehensive meta-analysis 1994–2019. *ReCALL*, *34*(2), 147–168. https://doi.org/10.1017/S0958344021000240

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approach*. Sage Publications.

Crompton, H. (2013). A historical overview of mobile learning: Toward learner-centered education. In Z. L. Berge & L. Y. Muilenburg (Eds.), *Handbook of mobile learning* (pp. 3–14). Routledge.

Chung, B., & Lo, N. (2025). Comparative study of self-regulated English learning through mobile language-learning applications in post-pandemic Hong Kong and South Korea. *Smart Learn. Environ.* **12**, 42. https://doi.org/10.1186/s40561-025-00399-w

Dashti, F., & Aldashti, A. (2015). EFL college students' attitudes towards mobile learning. *International Education Studies, 8* (8). DOI:10.5539/ies.v8n8p13

Dashti, F., & Abdulsalam, H. (2025). The influence of social media applications on learning English as a second langauge. Heliyon, 11 (2). https://doi.org/10.1016/j.heliyon.2025.e41874

Dai, Y., & Wu, Z. (2022). Mobile-assisted peer feedback on EFL pronunciation: Outcome effects, interactional processes, and shaping factors. *System, 111*, Article 102953. https://doi.org/10.1016/j.system.2022.102953

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

Eraut, M. (2004). Informal Learning in the Workplace. *Studies in Continuing Education,* 26(2), 247-273.

Garzón, J., Lampropoulos, G., & Burgos, D. (2023). Effects of mobile learning in English language learning: A meta-analysis and research synthesis. *Electronics, 12*(7), 1595. https://doi.org/10.3390/electronics12071595

Godwin-Jones, R. (2011). Mobile apps for language learning. *Language Learning and Technology, 15*(2), 2–11. http://dx.doi.org/10125/44244

Guo, H. (2022) *Mobile assisted language learning (MALL) from a sociocultural ecological perspective: An exploratory study of overseas Chinese students in the UK learning English speaking skills with mobile apps outside the classroom.* Doctoral thesis (PhD), UCL (University College London).

Inayati, N., Sanjani, M. I., Jayanti, F. G., Gao, X. (A.), & Nguyen, H. T. M. (May 24, 2024). Online informal language learning (OILL): A systematic review of studies (2014–2023). *Innovation in Language Learning and Teaching*, 1–17. https://doi.org/10.1080/17501229.2024.2355296

Kacetl, J., & Klímová, B. (2019). Use of smartphone applications in English language learning—A challenge for foreign language education. *Education Sciences, 9*(3), 179. https://doi.org/10.3390/educsci9030179

Kessler, M. (2023). Supplementing mobile-assisted language learning with reflective journal writing: A case study of Duolingo users' metacognitive awareness. *Computer Assisted Language Learning, 36*(5–6), 1040–1063. https://doi.org/10.1080/09588221.2021.1968914

Kessler, M., Loewen, S., & Gönülal, T. (2025). Mobile-assisted language learning with Babbel and Duolingo: Comparing L2 learning gains and user experience. *Computer Assisted Language Learning*, *38*(4), 690–714. https://doi.org/10.1080/09588221.2023.2215294

Klimova, B. (2018). Mobile phones and/or smartphones and their apps for teaching English as a foreign language. *Education and Information Technologies*. *[Crossref]*, *23*, 1091–1099.

Klimova, B. (2019). Impact of mobile learning on students' achievement results. *Education Sciences*. *[Crossref]*, *9*(2), 90. https://doi.org/10.3390/educsci9020090

Knowles, M. S. (1970). *The modern practice of adult education: Andragogy vs. pedagogy*. Association Press.

Kucuk, T., & Daskan, A. (2024). Mobile Assisted Language Learning (MALL) in Vocabulary: A Study on Tishk Language Preparatory School Students. Arab World English Journal, 15(4): 255-266. DOI: https://dx.doi.org/10.24093/awej/vol15no4.16

Kukulska-Hulme, A. (2009). Will mobile learning change language learning? *ReCALL*, *21*(2), 157–165. https://doi.org/10.1017/S0958344009000202

Kusyk, M., Arndt, H. L., Schwarz, M., Yibokou, K. S., Dressman, M., Sockett, G., & Toffoli, D. (2025). A scoping review of studies in informal second language learning: Trends in research published between 2000 and 2020. *System*, *130*, Article 103541. https://doi.org/10.1016/j.system.2024.103541

Lai, C., & Zheng, D. (2018). Self-directed use of mobile devices for language learning beyond the classroom. *ReCALL: Journal of Eurocall*, *30*(3), 299–318. https://doi.org/10.1017/S0958344017000258

Larsen-Freeman, D., & Long, M. H. (1991). *An introduction to second language acquisition research*. Routledge.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press. https://doi.org/10.1017/CBO9780511815355

Li, R. (2024). Effects of mobile-assisted language learning on foreign language learners' speaking skill development. *Language Learning and Technology*, *28*(1), 1–26. https://hdl.handle.net/10125/73553

Liddicoat, A. J. (2007). *An Introduction to Conversation Analysis*. London and New York: Continuum

Lightbown, P. M., & Spada, N. (2001). *How languages are learned* (2nd ed.). Oxford University Press.

Livingstone, D. W. (2001). Adults' Informal Learning: Definitions, Findings, Gaps, and Future Research. *NALL Working Paper #21*. Toronto: Centre for the Study of Education and Work, OISE/University of Toronto

Chung, B., & Lo, N. (2025). A comparative study of self-regulated English learning through mobile language-learning applications in post-pandemic Hong Kong and South Korea. *Smart Learning Environments*, *12*(1), 42. https://doi.org/10.1186/s40561-025-00399-w

Machmud, K., & Abdulah, R. (2017). Using smartphone-integrated model of teaching to overcome students' speaking anxiety in learning English as a foreign language. *Journal of Arts and Humanities*, *6*(9). https://doi.org/10.18533/journal.v6i9.1249

Metruk, R. (2022). Smartphone English language learning challenges: A systematic literature review. *Sage Open*, *12*(1). https://doi.org/10.1177/21582440221079627 (Original work published 2022).

Naderi, S., & Akrami, A. (2018). EFL learners' reading comprehension development through MALL: Telegram groups in focus. *International Journal of Instruction, 11*(2), 339–350. https://doi.org/10.12973/iji.2018.11223a

Nami, F. (2020). Educational smartphone apps for language learning in higher education: Students' choices and perceptions. *Australasian Journal of Educational Technology, 36(4), 82–95*. https://doi.org/10.14742/ajet.5350

Peng, H., Jager, S., & Lowie, W. (2021). Narrative review and meta-analysis of MALL research on L2 skills. *ReCALL, 33*(3), 278–295. https://doi.org/10.1017/S0958344020000221

Sarı, Ö., & Ersanlı, C. (2024). A systematic review of language learning applications in language education. *ELT Research Journal*, *13*(1), 71–84.

Roy, S., & Gandhimathi, S. N. S. (2025). Self-directed learning for optimizing sustainable language learning via mobile-assisted language learning: A systematic review. *Frontiers in Education, 9*, Article 1463721. https://doi.org/10.3389/feduc.2024.1463721

Shanthi, A., Paramasivam, S., Ling, C. G., & Tong, T. J. (2025). TikTok videos as a catalyst for vocabulary acquisition in tertiary education. *International Journal of Social Science Research, 13*(1), 1-15. https://doi.org/10.5296/ijssr.v13i1.22287

Shortt, M., Tilak, S., Kuznetcova, I., Martens, B., & Akinkuolie, B. (2023). Gamification in mobile-assisted language learning: A systematic review of Duolingo literature from public release of 2012 to early 2020. *Computer Assisted Language Learning*, *36*(3), 517–554. https://doi.org/10.1080/09588221.2021.1933540

Sockett, G. (2014). *The online informal learning of English*. Palgrave Macmillan. https://doi.org/10.1057/9781137414885

Soyoof, A., Reynolds, B. L., Vazquez-Calvo, B., & McLay, K. (2023). 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

Stockwell, G. (2022). *Mobile assisted language learning: Concepts, contexts and challenges*. Cambridge University Press. https://doi.org/10.1017/9781108652087

Tran, D. T. (2023). TikTok's effect on English vocabulary learning: An empirical study of English major curriculum. *GeSec: Revista de Gestão e Secretariado, 14*(10).

Xodabande, I., Pourhassan, A., & Valizadeh, M. (2022). Self-directed learning of core vocabulary in English by EFL learners: Comparing the outcomes from paper and mobile application flashcards. *Journal of Computers in Education, 9*(1), 93–111. https://doi.org/10.1007/s40692-021-00197-6

Yang, J. (2013). Mobile assisted language learning: Review of the recent applications of emerging mobile technologies. *English Language Teaching. [Crossref], 6*(7), 19–25. https://doi.org/10.5539/elt.v6n7p19

Zhao, D., Jablonkai, R. R., & Sandoval-Hernandez, A. (2024). Enhancing willingness to communicate in English among Chinese students in the UK: The impact of MALL with Duolingo and HelloTalk. *Journal of China Computer-Assisted Language Learning*, *4*(1), 42–73. https://doi.org/10.1515/jccall-2023-0027

Appendices

Appendix A Questionnaire

Section 1 Demographic

Gender:

Male

Female

Which grade/year are you currently in college?

Second year

Third year

Fourth year

Do you have a smartphone?

Yes

No

Section 2 Smartphone language learning apps

Which of the following smartphone language learning apps do you use?

- Duolingo
- YouTube
- Instagram
- TikTok
- Snapchat
- Twitter
- Babbel
- Orai
- Rosetta Stone
- Cambly
- FluentU
- Quizlet
- Google Translate
- BBC learning English
- Learn English vocabulary
- The English Club
- Others :……

How frequently do you use these apps?

- Daily
- Weekly
- Monthly
- Occasionally
- I don't use any of them

Language learning purposes of using Mobile APPs

I use smartphone apps develop listening skills

I use smartphone apps to develop my vocabulary knowledge

I use smartphone apps to develop my grammatical knowledge

I use smartphone apps to develop my cultural awareness

I use smartphone apps to develop my reading skills

I use mobile apps to develop my writing skills

Other purposes: ……..

Section 3 Perceptions of using smartphone to develop language skills

5 Likert scale- strongly disagree, disagree, neutral, agree, strongly agree

Personalization:

1- Smartphone language learning apps give immediate support to my language learning

2- Smartphone language learning apps extend my language learning experience beyond the language class.

3- Smartphone language learning apps enable me to learn the language at any time and any place

4- Smartphone language learning apps enable me to have self-paced, personalized language learning outside the classroom.
5- Smartphone language learning apps provide opportunities to act independently in learning the target language
Development of language skills and areas of linguistic knowledge
6- Smartphone Language learning apps can be effective for learning and improving my knowledge of English vocabulary.
7- Smartphone Language learning apps can be effective for learning and improving my knowledge of English grammar and writing.
8- I don't think that Smartphone language learning apps can be effective for improving my listening skill.
9- Smartphone Language learning apps can be effective for improving my reading comprehension skill.
10- Smartphone Language learning apps cannot be effective for learning correct pronunciation of English words.
11- Smartphone language learning apps can be effective for improving my speaking skill.
12- Smartphone language learning apps enhance my interaction with others in English language.
13- Smartphone language learning apps cab be effective to develop my cultural awareness about the foreign language
Authenticity:
14- Smartphone language learning apps enhance my willingness to participate in English language discussion
15- Smartphone language learning apps enhance my willingness to write in English when participating on social media platforms.
16- Smartphone language learning apps enhance the authenticity of my language learning resources and activities.
Connectivity:
17- Smartphone language learning apps enhance my interaction with others in the target language
18- Smartphone language learning apps enhance my connection with peer learners of the target language
19-Smartphone language learning apps connect me with speakers of the language
20- Smartphone language learning apps enhance my connection with the target language community
21- Smartphone language learning apps make my language learning experience in- and outside the class more connected

How do you use educational language learning apps outside the class to improve your English language?
……
……
……
……..
……

What are the challenges that you face when using smartphone Apps to improve your English language?
……
……

Appendix B Semi-structured interview:

Themes	Questions
Personalization	• Can you describe how you personalize your learning when using smartphone apps? • Are there features in the apps that you find helpful for tailoring your learning to your needs?
Authenticity	• Do you find the tasks or content in these apps similar to what you encounter in real-world language situations? • Can you give an example of an authentic task from an app that you found helpful?
Connectivity	• How do these apps help you connect with teachers, peers, or native speakers?

	• Have online interactions through these apps contributed to your learning? In what ways?
General experience	• What motivates or discourages you when using smartphone apps for language learning? • Are there features you wish these apps had to better support your language development?

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp.90-106
DOI: https://dx.doi.org/10.24093/awej/vol16no4.5

Applying the File-Card Model in Teaching and Learning English Speaking Skills for English-Majored Students

Huynh Thi Bich Phuong
Faculty of Foreign Languages, Ho Chi Minh University of Banking, Vietnam
Email: phuonghtb@hub.edu.vn

Tran Thanh Du
Faculty of Foreign Languages, Thu Dau Mot University,
Ho Chi Minh City, Vietnam
Corresponding author: dutt@tdmu.edu.vn

Nguyen Huu The
Faculty of Foreign Languages, Ba Ria - Vung Tau College of Education,
Ho Chi Minh City, Vietnam
Email: huuthecdsp@mail.com

Received:08/24/2025 Accepted: 11/02/2025 Published:12/10/2025

Abstract

This study examines how effectively the File-Card Model (Erteschik-Shir, 2007, 2016) can be applied to teaching and learning advanced English speaking skills among third-year English-major students in Vietnam. The reserachers conducted a quasi-experimental study with 30 participants divided into an experimental and a control group over eight weeks. The experimental group received explicit instruction through five stages, including topic mapping, focus marking, topic chaining, counterargument retrieval, and closing summary. The the control group, in contrast, followed the standard Speaking four syllabus. Data from pre- and post-test pair debates were analyzed using quantitative (t-tests) and qualitative discourse analysis, focusing on two indices: Topic Chaining Index (TCI) and Focus Shift Effectiveness (FSE). Results indicated significant improvements in the experimental group (TCI +23.1%, FSE +21.4%, $p < .001$). Students displayed smoother topic transitions, more explicit focus signaling, and stronger argumentative coherence. The study contributes theoretically by supporting the topic–focus interface in information structure and pedagogically by demonstrating that the File-Card Model provides a cognitive-linguistic framework to enhance coherence, fluency, and persuasiveness in EFL speaking. It also proposes TCI and FSE as effective discourse-level measures for evaluating oral proficiency in academic speaking training contexts.

***Keywords**:* debate pedagogy, discourse coherence, English-majored students, file-Card Model, Vietnam

How to cite: Phuong, H. T.B., Du, T.T., & The, N.H. (2025). Applying the File-Card Model in Teaching and Learning English Speaking Skills for English-Majored Students. *Arab World English Journal, 16*(4): 90-106. https://dx.doi.org/10.24093/awej/vol16no4.5

Introduction

Speaking is widely recognized as one of the most challenging skills to master in second language acquisition, especially at advanced levels, where learners are expected to manage complex discourse, construct coherent arguments, and sustain meaningful interaction. These demands require learners to develop four specific abilities that make their oral communication more explicit and effective (Luoma, 2004; Thornbury, 2005). In English for Academic Purposes programs, the ability to maintain coherent, persuasive, and extended spoken interaction is a key indicator of communicative competence, especially for learners preparing for academic and professional contexts (Hyland, 2019). Despite extensive exposure to English throughout their undergraduate studies, many advanced learners still struggle to maintain topic continuity, transition smoothly between ideas, and construct persuasive arguments during extended speaking tasks.. Recent studies further confirm these challenges in contemporary EFL contexts. For example, Elisathusilawani (2023) found that EFL students face persistent difficulties in maintaining fluency, coherence, and interactive competence in extended speaking tasks. Karim et al. (2023) demonstrated that digital-assisted learning tools can enhance students' oral performance but do not fully resolve discourse management issues. Moreover, Akhter et al. (2025) highlighted that advanced learners often require structured pedagogical frameworks to improve argumentation and focus management in academic speaking.

This issue is particularly salient in Speaking four, the final speaking course in the undergraduate English Language Studies program at a Vietnamese university. The course, designed for 30 third-year English-major students with proficiency levels ranging from B2+ to C1 on the CEFR scale, requires learners to participate in 5- to 7-minute pair debates, during which they are expected to present arguments that are logically organized and persuasive. Although students at this level typically demonstrate broad lexical and grammatical control, they often lack systematic strategies for topic chaining and focus management, which can lead to occasional loss of coherence, abrupt shifts in focus, and repetitive content that weakens the overall persuasiveness of their debates.

While reserachers widely use debate and argumentation to enhance critical thinking and oral fluency, most pedagogical approaches rely on rhetorical templates or memorized argumentative outlines rather than discourse-based frameworks. As a result, advanced learners may produce valid ideas but truggle to connect them coherently across turns or sustain topic development throughout interactions. In the Vietnamese academic oral context, researchers have paid limited attention to information structure, especially to how topics and foci interact dynamically to shape discourse. Thus, a pedagogical and theoretical gap remains in applying a linguistically informed model that explicitly teaches how to manage topic continuity and shifts in focus during extended speaking tasks such as debates.

To address this gap, the present study applies the File-Card Model (Erteschik-Shir, 2007, 2016) to the teaching and learning of advanced speaking skills. The model conceptualizes discourse as a dynamic set of "file cards," each representing a topic that can be updated, chained, or replaced as the conversation unfolds. A central insight of this model is its treatment of topic-focus relations, which, as Lambrecht (1994) notes, form the backbone of discourse organization by distinguishing between what is given (topic) and what is new or contrastive (focus). By raising learners' awareness of how topics are introduced, maintained, and shifted, and as well as how focus elements highlight new or contrastive information, the model provides a theoretically grounded means of improving the coherence, flow, and persuasiveness of spoken discourse.

The pedagogical significance of this study extends beyond classroom speaking performance. Developing argumentation and debate competence equips English majors with transferable skills that are essential for academic and professional success. Recent research confirms that structured debate activities enhance students' critical thinking, oral fluency, and confidence in presenting arguments (Ban, Pang & Em, 2023). Moreover, the ability to communicate persuasively in English has been linked to improved employability and professional readiness among university graduates (Alkooheji & Al-Hittami, 2024) . In educational settings, the ability to argue logically and respond critically to counterarguments fosters higher-order thinking and deeper disciplinary engagement (Kuhn, 2019). In professional contexts, persuasive speaking underpins effective communication in negotiations, presentations, and decision-making (Andrews, 2015; Robles, 2012). Hence, strengthening students' debate competence in Speaking four contributes both to immediate academic performance and to their long-term employability and communicative confidence. Theoretically, this research also provides empirical evidence for extending the File-Card Model into applied EFL pedagogy, linking information structure theory with oral argumentation training.

Accordingly, this study pursues three interrelated objectives. First, it seeks to apply the File-Card Model systematically in advanced speaking instruction to enhance discourse management and argument development. Second, it aims to examine the model's impact on learners' ability to maintain topic continuity, manage shifts in focus, and organize arguments coherently during pair debates. Finally, it evaluates whether such training leads to measurable improvements in coherence, fluency, and persuasiveness in students' spoken discourse. Together, these objectives create a clear link between theoretical rationale, pedagogical intervention, and empirical evaluation.

To achieve these objectives, the study addresses two guiding questions:

(1) How does applying the File-Card Model influence third-year English majors' ability to maintain topic continuity and coherence in extended debates?

(2) In what ways does the model affect the organization and persuasiveness of students' arguments in Speaking four pair debates?

By exploring these questions, the research aims to demonstrate how a linguistically motivated, cognitively transparent framework can strengthen learners' discourse competence in advanced speaking skills.

Literature review

The File-Card Model and Focus Structure

Focus Structure

Focus structure is a central mechanism in discourse that governs how information is presented and interpreted. It divides an utterance into two complementary parts: the topic, which anchors the sentence to already established knowledge, and the focus, which conveys new, contrastive, or emphasized information (Krifka, 2008; Vallduví & Engdahl, 1996). This distinction allows speakers to highlight what is important at a given moment while ensuring that their message links to the ongoing discourse. Focus can be realized in multiple ways: syntactically, through cleft or pseudo-cleft constructions (e.g., "*What we need to focus on is …*"); prosodically, through stress and intonation patterns that mark emphasis; and lexically, by inserting focus-sensitive markers such as *especially* or *in particular* (Gussenhoven, 2007; Féry, 2013). Beyond simple emphasis,

focus plays a key role in information packaging, guiding listeners in distinguishing between background information and the speaker's central message.

The File-Card Model and Its Relation to Focus Structure

The File-Card Model (Erteschik-Shir, 2007, 2016) conceptualizes discourse as a dynamic network of "file cards," each representing a topic that can be activated, updated, or replaced as the conversation unfolds. Within this model, focus structure functions as the updating mechanism. Whenever a speaker introduces new information, that focus is either attached to the currently active card or initiates the creation of a new, related card. In this way, focus not only marks what is informationally salient but also drives discourse progression, ensuring that communication remains both coherent and incremental.

This relationship between topic and focus is particularly critical in extended speaking tasks such as academic debates, where speakers must continuously balance anchoring their contributions to the established topic and advancing the argument through new focal information (Büring, 2016). Successful discourse, therefore, depends on maintaining this balance, keeping the thread of continuity while introducing contrastive or elaborative information that propels the discussion forward.

Building on this conceptual foundation, we can describe the File-Card-Focus Interaction as a three-stage cyclical process that explains how coherent and connected discourse develops in real time.

First, *Topic Identification*, the speaker determines which file card the current utterance is associated with. This stage involves activating the relevant card representing the ongoing topic. For example, in a debate on climate change, when a speaker states, "One major effect of global warming is rising sea levels.", the file card labeled *"effects of climate change"* becomes active.

Second, *Focus Integration*, the speaker then updates the selected file card by incorporating the new information introduced through focus. For example, the phrase *'rising sea levels'* serves as the focus, adding new and salient information to the existing card on *'effects of climate change"*. This process ensures that discourse progresses meaningfully without breaking topical coherence.

Third, *Topic Chaining*, when a focus naturally leads to a related subtopic, a new file card is created and linked to the preceding one, forming a chain of connected ideas. For example, The mention of *"rising sea levels"* may lead to a new card labeled *"impact on coastal communities."*

This chaining mechanism allows speakers to expand their arguments logically and persuasively while staying connected to the overarching theme (Erteschik-Shir, 2007; Ariel, 2010).

Through this cyclical process, discourse unfolds as a sequence of interconnected topics and foci, providing both cohesion and progression. In academic debates, this interactional mechanism enables speakers to organize their arguments hierarchically: they anchor statements to established topics, integrate new focal elements to add evidence or contrast, and chain these elements to subtopics that deepen the argument. For instance, a debater might begin with a file card on *"economic benefits of renewable energy,"* integrate the focus *"job creation,"* and subsequently chain it to a new card on *"long-term employment sustainability."* This layered progression exemplifies how the File-Card-Focus Interaction ensures logical continuity, rhetorical persuasiveness, and cognitive clarity in advanced spoken discourse. Recent studies show that mapping prosodic cues to focus remains challenging for L2 speakers, highlighting the cognitive load of updating discourse units (Yan, Warren & Calhoun, 2022). Cross-linguistic evidence further indicates that focus-to-prosody mappings differ between Mandarin and English, suggesting the

"Focus Integration" stage in the File-Card Model must account for language-specific realization of focus (Zhang, Clayards & Wagner, 2024).

File-Card Operation in Topic-Based Argumentation

Pair debates in the Speaking Four course clearly illustrate the File-Card Operation in topic-based argumentation. .The debate begins with the opening, where the speaker introduces the main topic, thereby creating the initial file card; for example, "Today, we argue that university education should be free" activates a topic card labeled University education policy.The first argument provides a claim and a corresponding focus to update the file card, such as "One key benefit is improved social equity" (focus = improved social equity). The evidence stage adds supporting information, further enriching the same file card, as in "Studies show that countries with free tuition have higher graduation rates among low-income students" (OECD, 2021). The speaker may then proceed to chaining to a related argument, which involves shifting to a new but thematically connected file card—for instance, "Beyond fairness, free education also boosts the economy," creating a new card titled Economic impact of free education linked to the original policy card (Psacharopoulos & Patrinos, 2018). At the counterargument and refutation stage, the speaker retrieves the relevant file card and introduces a contrasting focus, as in "Some argue it would strain government budgets, but research indicates long-term tax revenues outweigh initial costs" (Barr, 2014). Finally, the closing summarizes the updates across all file cards to reinforce the central thesis. By following this structured process, speakers maintain topic coherence, ensure that each new focus logically connects to an active file card, and construct persuasive, logically progressive arguments without losing the audience's attention. Recent research shows that structured topic-focus operations in debates improve discourse coherence and persuasiveness for advanced L2 learners (Li & Chen, 2023; Ahmad, 2022).

Pedagogical Applications of the File-Card Model in Speaking Instruction

Although researchers have primarily discussed the file-card model within theoretical linguistics, it offers substantial pedagogical potential for speaking instruction, particularly in contexts that demand sustained, coherent, and persuasive oral performance. In *Speaking four*-the final speaking course for third-year English-majored students, students perform pair debates lasting 5-7 minutes, aiming at C1-level fluency and discourse control. At this stage, many learners continue to face challenges in maintaining topic coherence, achieving smooth transitions, and organizing arguments logically. The file-card model can address these challenges by providing a clear cognitive framework for topic and focus management. To operationalize the model in the classroom, instructors can integrate specific activities targeting each stage of file-card operation, from topic introduction to closing summaries. These activities not only raise learners' metalinguistic awareness of discourse structure but also provide repeated, scaffolded practice in applying the model during real-time interaction.

The table below outlines suggested classroom activities aligned with the file-card model for debate preparation and delivery in Speaking four. These activities target different stages of file-card operation, from topic introduction to closing summaries.

Table 1. *Classroom activities aligned with the File-Card Model*

Stage in File-Card Model	Classroom Activity	Objective	Example in Speaking 4 Debate
Topic Identification & Mapping	Topic mapping exercise	Help students identify main and subtopics, assign file-card labels	Students brainstorm subtopics for "Free university education" (e.g., social equity, economy, quality assurance)
Focus Highlighting	Focus marking drills	Practice signaling new or critical information syntactically, prosodically, and lexically	Rewriting sentences with cleft structures ('What we need to focus on is…') and practicing stress patterns
Topic Chaining	Chaining expansion task	Develop ability to extend a topic with related arguments before shifting	Students add 2–3 linked arguments to "social equity" before moving to "economic benefits"
Controlled Topic Shifting	Transition marker practice	Ensure smooth and precise topic shifts	Using phrases like 'Turning now to…', 'Another key point is…' to introduce a new file card
Counterargument & Refutation	Argument retrieval & contrast exercise	Train students to recall and reframe existing file cards to respond to opposing points	Retrieving the "economic benefits" card to counter "government budget strain" the argument
Closing & Summary	File-card recap activity	Consolidate all updated file cards and reinforce the thesis	Summarizing each topic card with its focus updates to form a final speech

Criteria for Coherent and Persuasive Oral Argumentation

To effectively engage in academic debate and persuasive argumentation, learners must develop a range of discourse-level competencies beyond lexical and grammatical accuracy. First, clear topic identification and logical organization are essential, as they enable speakers to structure arguments coherently around a central thesis and its supporting subtopics (Nation & Newton, 2021). The use of topic mapping in the File-Card Model helps students visualize the overall argumentative structure more effectively.

Second, the ability to highlight focus plays a crucial rhetorical role. Strategic focus marking-through cleft constructions (e.g., "What we need to focus on is…"), lexical signals (e.g., "especially", "in particular"), and prosodic emphasis enhances the clarity and salience of new or contrastive information, thus increasing the persuasive force of the speaker's message (Gussenhoven, 2007; Féry, 2013).

Third, successful debating requires smooth transitions between arguments. Topic chaining achieves this by ensuring that each new point logically builds upon the preceding one. Transition markers such as "Building on that point" or "Another key issue is" serve to maintain discourse cohesion and thematic continuity (Erteschik-Shir, 2007; Ariel, 2010).

Fourth, speakers must anticipate and respond to counterarguments. The retrieval of previously established file cards and the strategic framing of rebuttals are essential for defending one's position and demonstrating critical thinking (Barr, 2014). Finally, persuasive delivery also depends on audience awareness. Intonation, stress, and structured signaling of key points contribute to more compelling and listener-friendly communication (Chafe, 1994; Wennerstrom, 2001).

These discourse strategies, when integrated through the File-Card framework, support not only argument development and coherence but also rhetorical impact and communicative confidence in academic speaking contexts.

Although the File-Card Model has been widely discussed in theoretical linguistics and has informed studies of discourse coherence and information structure (Erteschik-Shir, 2007, 2016; Büring, 2016), its pedagogical application in second language speaking instruction remains underdeveloped. Previous studies have primarily examined written discourse and theoretical models of topic–focus relations, but few have explored ways to operationalize the model in classroom contexts. . Recent studies have further emphasized the importance of structured argumentative practice and coherence in L2 academic speaking. Tsunemoto & Trofimovich (2024) found that improved logical ordering of ideas significantly enhanced perceived coherence and comprehensibility in advanced L2 academic speaking tasks. Additionally, research on parliamentary debate pedagogy (Smith & Lee, 2024) demonstrated that participation in structured debate improves learners' argumentation skills and overall persuasive performance. In the context of English advanced speaking instruction, studies have emphasized the importance of argumentation and debate (Hyland, 2019; Kuhn, 2019), but few have provided concrete discourse-based frameworks to support learners in managing topic continuity and focus integration during extended oral tasks. Moreover, within the Vietnamese advanced speaking settings of teaching and learning, there is a notable lack of empirical research investigating how structured models such as the File-Card approach can enhance debate performance and argumentative competence among advanced learners. This gap highlights the need for studies that not only test the effectiveness of the File-Card Model in teaching speaking but also explore how it influences learners' ability to sustain coherence, manage shifts in focus, and improve persuasiveness in academic debates.

Methodology

Research design

This study employs a quasi-experimental design to investigate the effects of applying the File-Card Model on the speaking performance of third-year English-major students enrolled in the Speaking four course.. The quasi-experimental approach was selected because the intact class could not be randomly reassigned to different conditions, yet a control group and an experimental group were available for comparison (Cook & Campbell, 1979; Dörnyei, 2007). The experimental group received explicit instruction and guided practice in applying the file-card model for topic-based argumentation, whereas the control group followed the standard Speaking four curriculum without specific reference to the model. The study employed a pre-test/post-test control group design (Creswell, 2014), in which both groups completed the same debate tasks at the beginning (pre-test) and end (post-test) of a nine-week intervention period. This design enabled the measurement of students' gains in topic management and argument development.

Participants

The participants were 30 third-year English-major students enrolled in the *Speaking four* course at a Vietnamese university. Their ages ranged from 20 to 22 years old, including 9 males and 21 females. The participants represented an intact class selected for convenience and consistency in instructional context. Participation in the study was voluntary, and informed consent was obtained before data collection.

All participants had completed prerequisite courses in Academic Speaking and Discourse Analysis, ensuring comparable linguistic and analytical proficiency. Their English proficiency levels ranged from B2+ to C1 according to the Common European Framework of Reference for Languages (CEFR), as verified by the university's placement testing system.

To ensure the suitability of task materials for this learner profile, the debate topics and task prompts were reviewed and pilot-tested before implementation to confirm appropriate linguistic and cognitive demands.

Research instruments

The instructional intervention for the experimental group lasted eight weeks, integrating the file-card model into both debate preparation and delivery. Each week consisted of one 100-minute session of instruction and one 30-minute practice session. The instructor provided 16 hours of instruction and 8 hours of practice in total. The intervention focused on five key components: 1) Topic Mapping (Week 1-2): 2) Students identified and organized main and subtopics for debates; 3) Focus Marking (Week 3-4): 4) They practiced using syntactic, prosodic, and lexical signals to highlight new or essential information; 5) Topic Chaining (Week 5-6): Students extended ideas before introducing new subtopics, ensuring logical flow. Counterargument Retrieval (Week 7): They learned to retrieve and refute arguments using established file cards. Closing Summary (Week 8): Students summarized and consolidated all updated file cards, reinforcing key arguments. These activities were adapted from speaking pedagogy frameworks (Nation & Newton, 2021; Richards, 2015) and the file-card-based discourse structure outlined by Erteschik-Shir (2007, 2016). This ensured that students developed an awareness of how topics and foci interact dynamically in extended discourse.

The instructional intervention for the experimental group was delivered over an eight-week period, integrating the file-card model into both debate preparation and delivery. The training sequence consisted of five key components: (1) topic mapping to identify main and subtopics for debates; (2) focus marking using syntactic, prosodic, and lexical signals; (3) topic chaining drills to extend ideas before introducing a new topic; (4) counterargument retrieval activities using previously established file cards; and (5) closing summary tasks consolidating all updated file cards. These activities were adapted from previous speaking pedagogy frameworks (Nation & Newton, 2021; Richards, 2015) and from the file-card-based discourse structure outlined by Erteschik-Shir (2007, 2016), ensuring that students developed an awareness of how topics and foci interact dynamically in extended discourse.

Speaking tasks (Pre- and Post-test)

Both the pre-test and post-test employed a pair debate format lasting between five and seven minutes per pair, following the principles of task-based speaking assessment (Fulcher, 2015; Brown & Abeywickrama, 2019). Each debate followed a fixed structure, consisting of an opening statement and topic introduction, a first argument supported by evidence, a second argument incorporating topic chaining, a counterargument with refutation, and a closing statement. Topics for the pre- and post-tests were matched in terms of familiarity, complexity, and potential for balanced argumentation, with examples including 'University education should be free' and 'Social media does more harm than good'. This parallel design ensured that improvements in performance could be attributed to the instructional intervention rather than to topic-specific advantages (Bachman & Palmer, 2010).

Data collection and coding

Recording and transcription

All debates were video-recorded and transcribed verbatim to facilitate detailed discourse analysis, in line with best practices in spoken language research (Markee, 2015).

Measurement of topic chaining

Topic chaining was operationalized as the number of coherent topic transitions linking a newly introduced topic card to an existing one in a semantically and logically connected manner, following definitions in discourse coherence studies (Halliday & Hasan, 1976; McCarthy, 1991). A valid chain was coded only when the new topic was explicitly linked to the preceding focus and signaled by a clear transition marker such as 'Building on that point' or 'Another benefit is'. The Topic Chaining Index (TCI) was calculated as: TCI = (Number of valid topic chains / Total number of topic shifts) × 100 (adapted from McCarthy & Carter, 1994).

Measurement of focus shifts

Focus shifts were identified whenever new information relevant to the active file card was introduced. Each shift was coded for its type—syntactic (e.g., cleft sentences), prosodic (intonation and stress), or lexical (focus markers)—and for clarity on a scale from 0 to 2 (0 = unclear, 1 = partially clear, 2 = fully clear), drawing on criteria from discourse-pragmatic research (Gussenhoven, 2007; Féry, 2013). The Focus Shift Effectiveness (FSE) score was computed as: FSE = (Sum of clarity scores for all focus shifts / Maximum possible clarity score) × 100.

Data analysis

The study adopted a mixed-methods approach (Creswell & Plano Clark, 2018). Quantitatively, paired-sample t-tests were used to compare pre-test and post-test scores within each group. In constrast, independent-sample t-tests compared the gains between the experimental and control groups for both TCI and FSE (Field, 2018). Qualitatively, discourse analysis was conducted on selected transcripts to illustrate how file-card operations were employed in real-time debates (Gee, 2014; Schiffrin et al., 2015). These excerpts provided evidence of how topic chaining and focus shifts contributed to sustained, coherent, and persuasive argumentation.

Ethical considerations

During the study period, ethical principles were carefully implemented. First, the researchers ensured the participants' welfare and treated them with humanity not only while deciding on the research question but also throughout the study. Second, participants' privacy and confidentiality were protected. Although the researchers knew who the participants were, they removed all identifying information from the report. Participants' personal data, such as names, phone numbers, email addresses, physical characteristics, and photos, were protected because all participants had a right to privacy. Finally, this article has never been submitted to or approved for consideration for another degree at another institution.

Results

Quantitative findings

The analysis revealed significant improvements in both Topic Chaining Index (TCI) and Focus Shift Effectiveness (FSE) for the experimental group compared to the control group. After eight weeks of instruction, the experimental group achieved a mean TCI of 78.4% (SD = 6.2), an increase of 23.1 percentage points from the pre-test. In contrast, the control group reached a mean TCI of 62.7% (SD = 7.5), showing only a 7.8-point increase.

For FSE, the experimental group recorded a mean score of 82.6% (SD = 5.9), up by 21.4 percentage points, whereas the control group averaged 68.9% (SD = 6.4), with a gain of just 8.2 points. Paired-sample t-tests indicated that both improvements in the experimental group were statistically significant ($p < .001$). In constract, independent-sample t-tests confirmed that the post-test differences between groups were also important ($p < .01$).

These results indicate that the experimental group showed significantly greater improvement in TCI than the control group, suggesting that the File-Card Model effectively enhanced their ability to maintain coherent topic shifts and logical progression during debates.. Example from Transcript (Topic Chaining): Experimental Group: Building on that point, free university education can reduce income inequality by offering more opportunities for people from disadvantaged backgrounds. Control Group: "Free education is a right... uh, people should have the right to education no matter their background".

Table 2. *Pre-test and post-test results with p-value*

Measurement	Group	Pre-Test Mean (%)	Post-Test Mean (%)	Gain (%)	p-value for within-group comparisons (paired-sample t-test)
TCI	Experimental	55.3 (SD=7.1)	78.4 (SD=6.2)	+23.1	< .001
TCI	Control	54.9 (SD=6.8)	62.7 (SD=7.5)	+7.8	.072
FSE	Experimental	61.2 (SD=6.5)	82.6 (SD=5.9)	+21.4	< .001
FSE	Control	60.7 (SD=6.3)	68.9 (SD=6.4)	+8.2	.054

Note: p-values reported for within-group comparisons (paired-sample t-test).

Qualitative observations

Discourse analysis of the debate transcripts provided further insights into how the file-card model influenced students' speaking performance. In the experimental group, topic chaining was more frequently accompanied by explicit transition markers, ensuring smooth thematic progression. For example, a student debating free university education transitioned from "It promotes equal opportunities" to "Moreover, it helps the economy grow" using the linking phrase "Building on that point", thereby maintaining logical cohesion.

Similarly, focus shifts in the experimental group were frequently signaled syntactically (e.g., through cleft sentences such as "What we must consider is…"), prosodically (with clear stress on new information), and lexically (using markers such as *especially* or *in particular*). In contrast, the control group showed more abrupt topic changes and less consistent signaling of focus, which leading to occasional caused lapses in coherence.

Analysis of the debate transcripts further highlighted how the file-card model influenced students' speaking behavior. In the experimental group, topic shifts were more frequently accompanied by explicit transition markers, ensuring smooth thematic progression. For example, during a debate on free university education, a student transitioned from "It promotes equal

opportunities" to "Moreover, it helps the economy grow" using the phrase "Building on that point", maintaining logical coherence. Example from Transcript (Focus Shifting): Experimental Group: "What we must focus on now is the social equity of free education. Especially for low-income families, it can be a game-changer. Control Group: "Uh, free education could be good... but what about the costs? It's hard to make it work.

Student interview findings

To enrich the qualitative analysis, semi-structured interviews were conducted with 10 students from the experimental group. The objective was to explore their perceptions regarding how the file-card model influenced their speaking performance in Speaking 4. The interviews were approximately 20 minutes in length and focused on three dimensions: (1) discourse organization, (2) focus marking, and (3) learner confidence.

A dominant theme was improved discourse coherence and strategic planning. Nine out of ten students reported that file cards helped them "see the whole argument" and organize their speech more effectively. One student stated, *"Before, I just said what came to my mind. But with file cards, I could arrange points in advance. For example, in the debate about social media, I had a card for privacy, a card for addiction, and a card for mental health. I just followed the cards, and it made my speaking more logical."*

Another participant elaborated on how the cards helped avoid repetition: *"I used to repeat the same idea because I forgot what I said before. With file cards, I had a clear list, so I didn't go in circles anymore."* This reflects the theoretical function of topic chaining and supports the observed gains in the Topic Chaining Index (TCI).

In terms of transitions and focus highlighting, students noted the helpfulness of having practiced specific discourse markers and focus structures. As one explained: *"Using phrases like 'Building on that point…' or 'Another key issue is…' helped me connect my ideas more smoothly. I used to pause or say 'uh… so… next', but now I can shift topics more confidently."*

Six participants specifically cited the use of cleft constructions (e.g., *"What we need to focus on is…"*) and lexical focus markers (e.g., *"especially"*, *"in particular"*) during their post-test performances. These elements contributed to the improved Focus Shift Effectiveness (FSE), as one student reflected: *"In the debate about free education, I said, 'What really matters is equal access.' That structure made my point sound stronger, more serious. My partner nodded right away, so I felt heard."*

Another student emphasized the persuasive effect on the audience: *"When I used phrases like 'The key issue we cannot ignore is…', I noticed my classmates listened more carefully. It gave me more authority."* Affective benefits were also reported. Seven students stated that the file-card method reduced anxiety and increased their speaking confidence. One noted, *"I used to get lost in the middle of my speech, and my mind went blank. But with file cards, I knew what came next, so I didn't panic."*

Another added that file cards supported collaborative performance: *"Because my partner and I used similar cards, we could follow each other better. When she moved to 'solutions', I knew my card was 'limitations', so I was ready. It felt more like teamwork."*

Collectively, these findings demonstrate that the file-card model supported not only linguistic gains but also metacognitive and affective development. By providing a structured and visualizable discourse plan, the method enhanced students' ability to manage extended speaking

tasks, thereby reinforcing both the observed statistical improvements and the theoretical rationale underpinning the study.

Findings related to recording and transcription

The video and audio recordings provided a rich dataset for analyzing both linguistic and paralinguistic features of the debates. Verbatim transcriptions captured not only the lexical content but also key discourse markers, pauses, self-repairs, and overlapping speech, which were crucial for identifying topic chaining and focus shifts.

Analysis of the transcripts revealed the following:

►Frequency of explicit transition markers: In the experimental group, 87% of topic shifts were accompanied by explicit transition markers (e.g., "Building on that point", "Another key issue is"), compared to 54% in the control group.

► Coherence in topic development: The recorded performances showed that the experimental group maintained coherent topic threads for an average of 4.2 turns before shifting topics. In constrast, the control group averaged 2.7 turns, indicating a stronger ability to sustain topic-relevant contributions.

► Focus marking signals: Prosodic features such as stress and intonation patterns, observable in the recordings, aligned with lexical focus markers in 76% of the experimental group's focus shifts, but only 48% in the control group, suggesting a more integrated use of verbal and prosodic cues.

►Disfluency patterns: Fillers (e.g., um, uh) and false starts occurred less frequently in the experimental group (mean rate: 3.1 per minute) compared to the control group (5.4 per minute), reflecting improved fluency and planning efficiency after file-card training.

These findings confirmed that the recording and transcription process was critical for capturing nuanced speaking behaviors that written scores alone could not reveal, thereby allowing for a more accurate assessment of the impact of the file-card model on speaking performance.

The video and audio recordings provided a rich dataset for analyzing both linguistic and paralinguistic features of the debates. Verbatim transcriptions captured not only lexical content but also discourse markers, pauses, self-repairs, and overlapping speech—key factors in identifying topic chaining and focus shifts. Example from Transcript (Explicit Transition Markers): Experimental Group: "To continue from where I left off, we can also see that free university education will lead to a more competitive workforce, which will benefit the economy." Control Group: "Okay, so, now I'll talk about the economic side of things... uh, it's costly, you know?"

Table 3. *Comparison of transcript-based indicators between experimental and control groups*

Indicator	Experimental Group	Control Group	Difference	Interpretation
% Topic shifts with explicit transition Markers	87%	54%	+33%	Greater discourse cohesion in experimental group
Average turns sustaining topic	4.2	2.7	+1.5	Improved topic development
% Focus shifts with prosodic and lexical marking	76%	48%	+28%	Better integration of verbal and prosodic cues
Fillers/False starts (per minute)	3.1	5.4	-2.3	Higher fluency and planning efficiency

Pre-Test and Post-Test Results

This subsection presents the detailed pre-test and post-test results for both the experimental and control groups (Table 4). Measures include Topic Chaining Index (TCI) and Focus Shift Effectiveness (FSE), with corresponding mean scores, gains, and standard deviations (SD). These data offer a more precise picture of the changes in performance following the intervention.

Note:

- TCI (Topic Chaining Index): Percentage of topic shifts coherently linked based on the file-card model.

- FSE (Focus Shift Effectiveness): Percentage of focus shifts clearly and appropriately marked using syntactic, prosodic, or lexical cues.

The experimental group demonstrated significant improvements in both TCI (+23.1%) and FSE (+21.4%), reflecting the positive impact of file-card-based training. The control group showed only modest gains, approximately one-third of those in the experimental group. Additionally, the reduced standard deviations in the experimental group from pre- to post-test indicate more consistent performance among participants.

Table 4. *Pre-test and post-test results with standard deviation*

Measure	Group	Pre-test Mean (%)	Post-test Mean (%)	Gain (%)	SD Pre	SD Post
TCI	Experimental	55.3	78.4	+23.1	7.1	6.2
TCI	Control	54.9	62.7	+7.8	6.8	7.5
FSE	Experimental	61.2	82.6	+21.4	6.5	5.9
FSE	Control	60.7	68.9	+8.2	6.3	6.4

Example from Transcript (Topic Chaining): Experimental Group: To continue from where I left off, free university education not only reduces inequality, but it also stimulates long-term economic growth. Control Group: “Uh, and yeah... free education is good... and people want it”. Example from Transcript (Focus Shift): Experimental Group: “What we really need to highlight now is the financial sustainability of such a policy. Especially in countries with limited budgets, it becomes a crucial issue.” Control Group: “Free education sounds great, but like... I don’t know, maybe it’s too expensive?” These examples illustrate the clear improvements in both thematic coherence and focus management in the experimental group, as evidenced by their use of explicit markers and structured transitions.

Discussion

The findings of this study demonstrate that applying the File-Card Model in *Speaking 4* for third-year English-majored students has a significant positive effect on both the Topic Chaining Index (TCI) and Focus Shift Effectiveness (FSE). The substantial gains in the experimental group compared to the control group - more than double in both measures - highlight the potential of file-card-based training as a pedagogically effective approach for advanced-level speaking instruction. These results explicitly support Erteschik-Shir’s (2007, 2016) theoretical proposition that structured topic management enhances discourse coherence by maintaining active “file cards” and updating them with relevant focus information. The model’s function as a dynamic topic–focus interface was empirically validated here: learners who internalized the file-card logic exhibited smoother topic chaining and more precise focus signaling, suggesting that discourse coherence emerges from the continuous updating of mental “cards.”

The quantitative improvements in TCI indicate that students became more adept at linking ideas within a coherent thematic progression, suggesting that explicit training in topic chaining

directly translated into better argumentative organization. This finding aligns with Erteschik-Shir's claim that topic continuity is a core mechanism in maintaining discourse accessibility and supports the model's extension into applied linguistics. It also resonates with previous findings in debate pedagogy, such as Kuhn (2019) and Hyland (2019), who emphasize dialogic reasoning and genre-based scaffolding. However, while Kuhn focuses on cognitive reasoning and Hyland on rhetorical organization, the File-Card Model provides a *micro-discourse mechanism* that operationalizes these principles: it shows *how* speakers retrieve, update, and link ideas through topic–focus operations during real-time debate.

Similarly, the gains in FSE suggest that students were better able to introduce, highlight, and transition focus points within their arguments. The use of syntactic, prosodic, and lexical markers observed in the experimental group reflects a more deliberate and audience-oriented discourse strategy. . hese improvements likely result from several underlying cognitive mechanisms. First, *attentional control*: focus marking heightens the salience of key propositions, guiding message planning and comprehension. Second, *cue-based retrieval*: labeled topic cards facilitate memory recall of earlier claims during rebuttals. Third, *working-memory offloading*: grouping arguments into topic-linked cards reduces planning load, allowing smoother formulation. These mechanisms together explain why the model effectively enhanced coherence and persuasiveness—learners could track discourse states cognitively while maintaining linguistic precision.

The qualitative findings further support the quantitative outcomes. Discourse analysis showed that participants in the experimental group frequently employed cohesive transition markers (e.g., "Building on that point") and maintained coherent and smoothly progressing topic development throughout their discourse. In contrast, the control group often shifted issues abruptly, leading to occasional breakdowns in coherence. These differences suggest that the File-Card Model's structured mental representation of issues can help learners maintain argument continuity over extended turns, even in cognitively demanding debate tasks.

Moreover, the reduction in standard deviations for both TCI and FSE in the experimental group points to more uniform performance across participants. This may indicate that the File-Card Model not only benefits high-achieving students but also supports weaker speakers by providing a concrete cognitive framework for organizing speech. Such scaffolding appears particularly relevant for C1-level target outcomes, where communicative success depends on precise discourse management rather than basic grammatical accuracy alone.

In summary, the study's results confirm that integrating the File-Card Model into advanced speaking instruction fosters measurable improvements in discourse organization and focus control. Theoretically, these results strengthen the explanatory power of Erteschik-Shir's model by showing its pedagogical adaptability; empirically, they extend the scope of debate pedagogy through discourse-cognitive training.

Nevertheless, researchers should note several limitations. The study involved a relatively small sample (n = 30) from a single Vietnamese university, which restricts the generalizability of the findings. Future studies should replicate this approach with larger and more diverse cohorts across institutions to validate the observed effects and explore long-term retention and transferability.

Conclusion

The current study examined the application of the File-Card Model in advanced English-speaking instruction for third-year English-majored students. The quasi-experimental results

revealed significant gains in both Topic Chaining Index (TCI) and Focus Shift Effectiveness (FSE), supported by qualitative evidence of smoother transitions, more explicit focus marking, and greater coherence in student debates. These outcomes confirm that integrating the File-Card Model fosters structured discourse organization and persuasive communication at the C1 level. Pedagogically, the model offers a practical cognitive framework that helps learners manage topic continuity and focus more systematically. Classroom activities designed around topic mapping, chaining, and focus updating can enhance discourse coherence and confidence in debate contexts. The study also introduces TCI and FSE as reliable indicators for assessing discourse-level proficiency, bridging theory and pedagogy through measurable outcomes. Theoretically, the findings provide empirical support for Erteschik-Shir's topic–focus framework, demonstrating its applicability in EFL speaking classrooms. Nonetheless, the study's small sample size and single-institution setting limit the generalizability of the results. Future research should test the model across diverse contexts and examine its long-term impact on oral performance and transferability to other genres of academic communication.

Funding: The authors did not receive any funding for this research

Acknowledgment: Not applicable.

Conflicts of interest: The author declares no conflicts of interest

Authenticity: This manuscript is an original work

Artificial intelligence statement: During the preparation of this work, the authors used ChatGPT of OpenAI to check the grammatical corrections. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the content of the publication. AI and AI-assisted technologies were also used partially to check the clarity of language.

About the authors:

Dr. Huynh Thi Bich Phuong is currently a lecturer at the Faculty of Foreign Languages, Ho Chi Minh University of Banking (HUB), Vietnam. She holds her M.A. in TESOL and her Ph.D. in Contrastive Linguistics, both from the University of Social Sciences and Humanities, Vietnam National University of Ho Chi Minh City. She specializes in teaching methodology, contrastive linguistics, and applied linguistics. She has also been a reviewer for some national journals. Her research interests include language skills, English teaching methodology, language testing and assessment, pragmatics, and the interplay between language, culture, and society. Her published works focus on both linguistic studies and issues related to TESOL Email: phuonghtb@hub.edu.vn; ORCID: https://orcid.org/0009-0006-5142-629X

Dr Tran Thanh Du specializes in both TESOL methodologies and Contrastive Linguistics. He is currently a lecturer at the Faculty of Foreign Languages at Thu Dau Mot University in Vietnam. He takes his interests in studies of linguistics, sociolinguistics, pragmatics, learner motivation, and English language teaching methodologies. His publications concentrate on both linguistic research and TESOL-related issues.. He endeavors to improve language use in general, and English language teaching and learning in global contexts. Email: dutt@tdmu.edu.vn; ORCID: https://orcid.org/0009-0002-7802-9949

Nguyen Huu The is a Doctor of Contrastive Linguistics. He is currently a lecturer at the Faculty of Foreign Languages, Ba Ria - Vung Tau College of Education, Vietnam. His primary areas of expertise are TESOL and Contrastive linguistics. His research interests are linguistics, semantics, stylistics, pragmatics, and English language teaching methodologies. All of his published works focus on linguistic research as well as topics related to TESOL. Email: huuthecdsp@gmail.com; ORCID: https://orcid.org/0009-0004-2075-6793

References

Alkooheji, B., & Al-Hittami, A. (2024). *University Graduates' Perception about the Effect of EFL Speaking Skills on Employability.* International Education and Social Studies, 5(2), 78-89. https://www.scholink.org/ojs/index.php/iess/article/view/3723

Ahmad, S. (2022). *The impact of structured debate tasks on English L2 learners' argumentative skills. Language Teaching Research*, 26(5), 745–763. https://doi.org/10.1177/13621688221091015

Andrews, R. (2015). *Argumentation in higher education: Improving practice through theory and research*. Routledge.

Akhter, S., Ahmed, K., Tambi, F., Siwok, S. R., Lin, C., & Shaheen, M. (2025). *Enhancing EFL Learners' Communication Skills: A CEFR-Inspired Approach for High School Students.* Forum for Linguistic Studies, 7(5), 213-226. https://doi.org/10.30564/fls.v7i5.918

Ariel, M. (2010). *Defining pragmatics*. Cambridge University Press.

Ban, P., Pang, M., & Em, S. (2023). *Debate: One of the Key Factors to Improving Students' English Language Speaking Skills.* Global Education and Humanities University Journal, 2(1), 45-56. https://journal-gehu.com/index.php/gehu/article/view/69

Barr, R. (2014). Argumentation and critical thinking: A guide for classroom debate. *Journal of Education and Learning, 3*(2), 12–21. https://doi.org/10.5539/jel.v3n2p12

Büring, D. (2016). *Intonation and meaning*. Oxford University Press.

Chafe, W. (1994). *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. University of Chicago Press.

Elisathusilawani, E. (2023). *Exploring EFL Students' Speaking Challenges in Their Speaking for Social Purposes.* Journal of English and Education, 9(1), 10-21. https://doi.org/10.20885/jee.v9i1.28010

Erteschik-Shir, N. (2007). *Information structure: The syntax–discourse interface*. Oxford University Press.

Erteschik-Shir, N. (2016). *The grammar of focus*. Cambridge University Press.

Féry, C. (2013). *Focus as prosodic prominence*. In M. Krifka (Ed.), *The expression of information structure* (pp. 234–258). De Gruyter.

Gussenhoven, C. (2007). Types of focus in English. In C. Féry, G. Fanselow, & M. Krifka (Eds.), *The notions of information structure* (pp. 83–100). De Gruyter.

Hyland, K. (2019). *Second language writing*. Cambridge University Press.

Karim, S. A., Hamzah, A. Q. S., Anjani, N. M., Prianti, J., & Sihole, I. G. (2023). *Promoting EFL Students' Speaking Performance through ELSA Speak: An Artificial Intelligence in English Language Learning. Journal of Languages and Language Teaching*, 11(4), 655-668. https://doi.org/10.33394/jollt.v11i4.8958

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, *55*(3–4), 243–276. https://doi.org/10.1556/ALing.55.2008.3-4.2

Kuhn, D. (2019). A role for reasoning in a dialogic approach to critical thinking. *Topoi, 38*(1), 121–129. https://doi.org/10.1007/s11245-018-9561-1

Lambrecht, K. (1994). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge University Press.

Li, X., & Chen, Y. (2023). *Enhancing discourse coherence in L2 speaking through structured argumentation. System,* 115, 102797. https://doi.org/10.1016/j.system.2023.102797

Luoma, S. (2004). *Assessing speaking*. Cambridge University Press.

Nation, I. S. P., & Newton, J. (2021). *Teaching ESL/EFL listening and speaking*. Routledge.

Robles, M. M. (2012). Executive perceptions of the top 10 soft skills needed in today's workplace. *Business Communication Quarterly, 75*(4), 453–465. https://doi.org/10.1177/1080569912460400

Smith, J., & Lee, K. (2024). *The effects of parliamentary debate as a pedagogy for argumentation in L1 and L2 contexts.* Argumentation, 38(2), 123-145. https://link.springer.com/article/10.1007/s10503-024-09645-w

Thornbury, S. (2005). *How to teach speaking*. Longman.

Tsunemoto, T., & Trofimovich, P. (2024). *Coherence and comprehensibility in second language speakers' academic speaking performance.* Studies in Second Language Acquisition. https://www.cambridge.org/core/journals/studies-in-second-language-acquisition/article/coherence-and-comprehensibility-in-second-language-speakers-academic-speaking-performance/85221E116322A1A7ACD84242C454712A

Vallduví, E., & Engdahl, E. (1996). The linguistic realization of information packaging. *Linguistics, 34*(3), 459–519.

Wennerstrom, A. (2001). *The music of everyday speech: Prosody and discourse analysis*. Oxford University Press.

Yan, M., Warren, P., & Calhoun, S. (2022). *Focus interpretation in L1 and L2: The role of prosodic prominence and clefting.* Applied Psycholinguistics, 43(6), 1275-1303. https://doi.org/10.1017/S0142716422000376

Zhang, W., Clayards, M., & Wagner, M. (2024). *Prosodic Effects of Focus and Constituency in Mandarin and in English.* Laboratory Phonology, 15(1). https://doi.org/10.16995/labphon.970

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp.107-123
DOI: https://dx.doi.org/10.24093/awej/vol16no4.6

An Investigation of Writing Errors Made by Saudi EFL Learners

Rawan Ahmed Alsoweed
Department of English, College of Education, Majmaah University
Al-Majmaah, 11952, Saudi Arabia
Email: rawan.alsoweed@gmail.com

Abeer Shujaa Alharbi
Department of English, College of Education, Majmaah University
Al-Majmaah,11952, Saudi Arabia
Corresponding Author: as.alharbi@mu.edu.sa

Fahad Abdulaziz Alqusumi
Department of Strategic Planning and Transformation, Qassim University Medical City
Burydah 52571 Saudi Arabia
Email: fahad.alqusumi@gmail.com,

Received:08/15/2025 Accepted:11/4/2025 Published: 12/10/2025

Abstract

Grammatical errors remain a persistent obstacle for English language learners, often impeding their progress and causing frustration. Addressing these recurring issues requires a systematic understanding of their root causes, especially within specific linguistic contexts. This study investigates the grammatical errors made by Saudi learners of English as a Foreign Language (EFL) enrolled in the preparatory year program. Drawing on students' written compositions, the analysis aims to deepen learners' understanding of English while acknowledging that errors are a natural part of language acquisition. Developing accurate writing skills in English remains a challenge, particularly in the use of grammar and punctuation, both of which are critical for conveying clear and effective meaning. Using Byrd and Benson's (1994) comprehensive error taxonomy alongside Ellis's (1997) error analysis framework, the study categorizes the observed errors into three main types: most serious errors, intermediate errors, and punctuation and mechanical problems. The findings reveal widespread and repeated patterns of deviation, suggesting that many students struggle with specific grammatical rules due to insufficient comprehension or incomplete learning. Contributing factors such as omission, addition, substitution, carelessness, and first language (L1) interference are identified as key reasons behind these challenges. The influence of Arabic as the learners' native language is found to be particularly significant, often leading to structural and syntactical transfer that manifests in common grammatical and punctuation mistakes.

Keywords: Error analysis, error taxonomy, grammatical errors, linguistic challenges, L1 interference, Saudi English as a foreign language students, second language acquisition, writing
How to cite: Alsoweed, R. A., Alharbi, A.S., & Alqusumi, F.A. (2025). An Investigation of Writing Errors Made by Saudi EFL Learners. *Arab World English Journal, 16*(4): 107-123. https://dx.doi.org/10.24093/awej/vol16no4.6

Introduction

Mastering writing skills remains one of the most demanding tasks for English as a Foreign Language (EFL) learners, particularly among Saudi students enrolled in university preparatory programs. Despite years of English study, these learners frequently commit errors in grammar, punctuation, and capitalization, which hinder the development of effective written communication. Writing demands not only linguistic competence but also the ability to organize and articulate ideas clearly—a challenge intensified when learners' first language (L1) differs significantly from the target language. As noted by Greenbaum and Nelson (2009), grammar involves a set of rules for constructing meaningful expressions, yet many EFL learners struggle to apply these rules accurately. Hermini (2014) emphasizes that while writing instruction helps develop grammar and vocabulary, many students still exhibit persistent inaccuracies in syntax and structure.

Saudi universities, offer preparatory year programs across various tracks—Medical, Nursing, and Business—with curricula tailored to students' language proficiency levels. Despite this instructional support, Saudi learners continue to face difficulties, especially in written compositions. Common issues include misuse of verb tenses, articles, prepositions, word order, punctuation, and capitalization. Moreover, writing errors are not random but frequently influenced by learners' native language. According to Ellis (1997, 2008), language learners often rely on L1 structures when acquiring a second language, leading to interference in various linguistic areas such as grammar and syntax.

Errors in language learning should not be viewed merely as shortcomings. Corder (1967) and James (1998) argue that they offer valuable insights into the learner's developmental process and can help educators identify learning gaps. Writing errors, in this context, become diagnostic tools for understanding how students internalize and apply English language rules. According to Moqimipour and Shahrokhi (2015), many mistakes stem from inadequate knowledge of language structures, which results in incorrect expressions that hinder comprehension. Effective teaching thus requires a systematic approach to identifying, classifying, and addressing these errors.

Error Analysis provides a structured method for uncovering the root causes of learners' mistakes. Several models exist, including Corder's (1967) steps for identifying and explaining errors, Ellis's (1997) four-stage process, and Gass and Selinker's (1992) six-step model. Byrd and Benson's (1994) taxonomy, which categorizes errors into three areas—most serious, intermediate, and punctuation/mechanics—offers a practical framework for instructional analysis and was adopted in this study. These approaches reveal that common causes of writing errors include omission, addition, substitution, carelessness, and L1 interference (Ellis, 2010; Hasyim, 2002).

This study aims to identify the most frequent grammatical, punctuation, and capitalization errors made by Saudi EFL students, explore the underlying causes of these mistakes—including the influence of the Arabic language—and offer practical recommendations for improving students' writing proficiency through targeted instructional strategies. This study seeks to answer two key research questions: What are the most common grammatical, punctuation, and capitalization errors made by Saudi EFL learners in preparatory year programs, and to what extent does the influence of Arabic as a L1 contribute to these recurring writing errors?

Literature Review

The study of writing errors among Arab learners of EFL, particularly Saudi students, has attracted growing attention in applied linguistics due to the persistent challenges learners face in mastering English grammar and mechanics. Writing, as a productive skill, often reveals deep-

seated difficulties that are not always visible in receptive skills. Numerous studies have emphasized that the difficulties EFL learners encounter in writing are often rooted in L1 interference, insufficient knowledge of English grammatical structures, or pedagogical limitations in EFL instruction. One major theme across the literature is the role of L1 interference in causing grammatical and mechanical errors. Al-Zoubi and Abu-Eid (2014) investigated the writing of 266 Jordanian university students through translation-based assessments and found a high frequency of grammatical errors caused by direct translation from Arabic to English (p. 364). This study aligns with Ridha (2012), who analyzed the errors of 80 Iraqi university students and categorized them into grammatical, lexical/semantic, mechanical, and word order errors (p. 31). These findings underscore the significance of interlanguage theory, where learners construct a transitional linguistic system influenced by their L1.

The Saudi context further supports these findings. Alsahafi (2017) conducted a study involving 70 Saudi female students and identified orthographic, lexical, and grammatical errors (p. 127). Among the orthographic errors, substitution was most frequent, while lexical errors often stemmed from literal translation and word confusion. Smith (2001) provided a broader perspective on common grammatical errors made by Arabic-speaking EFL students, noting that challenges often arose with consonant clusters, negation, modal verbs, word order, and article use. Smith (2001) particularly emphasized the omission of the indefinite article a, a recurring issue across different types of nouns.

Alhamadi (2015) further emphasized that many of the errors Arab learners made were rooted in intralingual causes rather than interlingual ones. Abdul Rahman (2021) offered insights from a different angle by surveying 100 English language instructors across Saudi universities. The majority of respondents identified verb tense, article use, and preposition errors as the most problematic areas for students. Similarly, Al-Shujairi and Tan (2017) applied Corder's (1967) and James's (1998) error analysis frameworks to Iraqi pre-university students' writing. Their results highlighted frequent omissions, substitutions, and additions—particularly involving articles and verb tenses—and suggested that interlanguage development was the main source of these errors.

AlTameemy and Daradkeh (2019) investigated the writing of 80 Saudi EFL students and classified 1580 errors, of which 42.15% occurred at the sentence level (p. 178). The most common errors included grammar, punctuation, spelling, and capitalization issues. The frequency and diversity of these mistakes revealed a lack of familiarity with standard written English conventions. Raba (2012) also examined punctuation errors among TEFL students and found that misuse of periods and commas was highly prevalent. For example, the Arabic comma is visually different and used differently than the English comma. Macarthur (1999) pointed out that transcription skills, including accurate punctuation and spelling, remain a challenge even for advanced learners, and that these aspects should be explicitly taught rather than assumed. Siddiqui (2015) conducted a focused study on capitalization errors among Saudi female students and found high frequencies of incorrect use.

Together, these studies provide a comprehensive picture of the error landscape for Arab EFL learners. The evidence consistently shows that both interlingual (L1 interference) and intralingual (learning-related) factors contribute to the errors found in students' writing. While grammatical errors—especially those involving verbs, articles, and sentence structure—are the most discussed in the literature, mechanical issues such as punctuation, spelling, and capitalization are also widespread and equally detrimental to effective communication. This study contributes to the existing body of research by exploring the writing errors of Saudi EFL learners. By employing

Byrd and Benson's (1994) taxonomy and Ellis's (1997) error analysis model, the study aims to classify the types of errors, investigate their sources, and offer pedagogical recommendations for grammar and writing instruction.

Methodology

This study investigates the most frequent grammatical errors made by undergraduate students enrolled in the preparatory year program in Al-Bukayriyah, Qassim, Saudi Arabia. The key objective is to identify, describe, and analyze the most commonly occurring grammatical errors in students' written English compositions and explore the underlying causes of these errors within the context of EFL instruction. To achieve this, the study employed Byrd and Benson's (1994) twelve-type error taxonomy as the primary analytical framework. These errors were grouped into three main categories: major grammatical errors, intermediate-level errors, and errors related to punctuation and mechanics. Due to the notable recurrence of prepositional errors, an additional subcategory under intermediate errors was introduced. For the analysis procedure, Ellis's (1997) four-step model of error analysis—comprising identification, description, explanation, and evaluation—was applied. A qualitative descriptive research design was chosen for this investigation. The analysis focused on understanding the specific types of errors and their frequency, and deriving pedagogical insights that can inform instructional strategies. The study was conducted during the 2023–2024 academic year and relied on authentic writing samples collected from regular coursework and examinations.

The participant pool originally included 80 students; however, only the writing samples of 56 undergraduate students were selected for final analysis, as only those who made grammatical errors were included in the qualitative assessment. These 56 participants were enrolled in Medical, Nursing, and Business programs. The selected sample consisted exclusively of female students aged between 18 and 22. Most of them had either recently graduated from Saudi high schools or completed a preparatory year before entering their respective tracks. The students' proficiency levels ranged from lower-intermediate to intermediate. They were taught by three different EFL instructors, each responsible for delivering instruction in their designated academic tracks. The writing tasks varied depending on course objectives and skill levels, offering a diverse data set across the three disciplines.

Data collection encompassed a variety of writing samples derived from different tasks, including pieces composed in class as assignments and those generated during the midterm examination. As indicated below, in table 1, each academic track submitted two writing samples representing the initiation of the course (task 1), the midpoint (midterm task) were students had an option of two tasks. The differentiation in tasks aimed to discern patterns and variations in error occurrence. Grammatical sentence errors were analyzed per conventional grammatical rules, utilizing the framework outlined by Byrd and Benson (1994), and the error analysis procedure delineated by Ellis (1997); the collected data underwent identification, description, explanation, and evaluation stages. Qualitative data analysis followed the three-step process of data reduction, data display, and verification of conclusions, as proposed by Miles and Huberman (2014). The errors were identified and calculated by the three instructors, then the types of errors were systematically categorized into omission, addition, substitution, carelessness, and L1 interference, with the classification of errors analyzed by instructors providing insights into the causal factors of the classification of errors.

Table 1. *A description of data collection from the tasks*

Track	Midterm Task	Classroom Task 1
Nursing	Write a paragraph about how you can study or describe a vacation trip	Write a paragraph about how you study
Medical	Write a descriptive paragraph: describe a painting or a famous person from the last hundred years	Write a letter inviting a friend
Business	Write an email for a job opening	Recommend an employee to join your team

Results

Grammatical Error Classification based on Byrd and Benson (1994)

A detailed analysis of the written compositions from 56 Saudi undergraduate students revealed three primary categories of grammatical errors: Serious Errors, Immediate Errors, and Punctuation and Mechanics Problems, as adapted from Byrd and Benson's (1994) classification. The study underscored that while all students exhibited grammatical mistakes, there were notable differences in the frequency and type of errors across the sampled cohorts.

The first category, Serious Errors, includes issues such as incorrect verb tense usage, subject-verb agreement failures, and sentence boundary problems, including run-ons, comma splices, and sentence fragments. For instance, in the sentence "My mother wanting me study in university," the verb form is erroneous due to incorrect tense and missing infinitive construction. Similarly, agreement issues appear in sentences like "She study every day," where the verb does not align with the singular subject. Run-on sentences such as "I want to take my IELTS exam I want to score of 5" indicate an absence of proper punctuation, while comma splices, exemplified by "I will plan a business event, the company is big, they want to invite many guests," further reflect sentence boundary confusion. Additionally, sentence fragments like "I student studying nursing" lack auxiliary verbs necessary to form a complete sentence.

The second category, Immediate Errors, encompasses a broader array of grammatical issues, including incorrect word order, misuse of word forms, article omissions, prepositional errors, and violations of parallel structure. A representative example is the sentence "My project business is solve problem company American," which illustrates disordered word placement leading to incoherence. Errors in word forms also surface in expressions like "I like reading about helping sick person, I like curing patient," where the plural forms "people" and "patients" are required. Article misuse appears frequently, as in "I am student", which omits the necessary indefinite article "a." Prepositional confusion, often influenced by first-language interference, is reflected in "The girl standing on the picture, is standing on the river," where the prepositions "in" and "on" are misapplied. Problems with parallelism are also evident, as seen in "I love study Nurse and helped patients," which requires consistency in verb forms, i.e., "studying and helping."

The third and final category pertains to Punctuation and Mechanics Problems, which includes errors in punctuation, capitalization, and spelling. The absence of proper punctuation is shown in sentences such as "I want to be a doctor I want to specialize in cardiology," where the lack of full stops undermines sentence clarity. Improper capitalization appears in "a nurse has 3 Job Responsibilities," violating sentence-initial and proper noun conventions. Spelling errors are rampant and include both vowel and consonant inaccuracies, as in "The desrptive picture has a beautful bakgrond with a mother and her childrn," where the words "descriptive," "beautiful," "background," and "children" are all misspelled.

Table 2 provides a taxonomy of all identified errors, including descriptive explanations and authentic examples from the student writing samples.

Table 2. *Grammatical errors classification (Bryd & Benson, 1994)*

Group of Errors	Error Types	Description	Examples	Error Explanation
Serious errors	**1.** Verbs	Making errors in verb tense usage or constructing passive sentences inappropriately.	My mother wanting me study in university.	Two errors, verb tense and infinitive from.
	2. Agreement	Making an incorrect form of the subject or the verb and pronoun usage.	-She study every day. -The IT companies didn't like the offer, it was expensive.	The subject she is singular. Correction: the verb: study, ends with s, studies.
	3.Run-on sentence	Combining two or more sentences without using the proper punctuation.	I want to take my IELTS exam I want to score of 5	Two clauses are combined in one sentence without any punctuation.
	4. Comma Splice	A comma is used to join two or more complete sentences incorrectly.	I will plan a business event, the company is big, they want to invite many guests.	The sentence should be separated by a full stop not a comma.
	5. Fragment	Writing a fragment (run-on sentence) instead of a complete one.	I student studying nursing.	In this example, there should be a verb- am in this sentence.
Immediate errors	**6.** Word order	Arranging a sentence's order in incorrect positions, including adjectives.	My project business is solve problem company American. Student has to hard study.	Words are ordered in a wrong way and resulting in other errors. Correction: My business project is to solve the problems for the American company.
	7. Word from	Misusing a word's form, including its tenses or single and plural forms.	I like reading about helping sick person, I like curing patient.	Word from should be plural: people, patients.
	8. Article	Misusing the incorrect article to convey the idea or omitting an article that is necessary to convey the meaning.	I am student.	The sentence requires a, the indefinite article.
	9. Prepositions	Misusing prepositions to convey movement, location, simple preposition and in phrases.	The girl standing on the picture, is standing on the river.	The preposition is misused, correction: in, on.
	10. Parallelism	Composing/writing words in different grammatical forms.	I love study Nurse and helped patients.	Its common to use the same from in the sentence, correction: studying/helping

ISSN: 2229-9327

Punctuation and Mechanic problems	**11**. Punctuation	Making a punctuation error.	I want to be a doctor I want to specialize in cardiology.	A full stop is need to end the first and second part of the sentence.
	12.Capitalization	Inappropriate capital or small letter forms.	a nurse has 3 Job Responsibilities.	The letter case in the word from is incorrect, errors at the beginning, middle and ending of the sentence.
	13. Spelling	Misspelling words	The descrptive picture has a beautful bakgrond with a mother and her childrn.	The misspelling of vowels and consonants. Misspelled words: descriptive, beautiful, background children.

Frequency of Errors

This section presents an analysis of the grammatical errors made by Saudi undergraduate students in their English writing, as shown in Table 3. A systematic process was employed to identify and categorize these errors, enabling the calculation of both the frequency and percentage for each error type. The error identification and quantification were carried out by English language instructors, who assessed the students' written compositions and classified the errors based on their occurrences within each category. Following Ellis's (1997) error analysis framework, the process involved four key stages: identification, description, explanation, and evaluation, which will be elaborated in subsequent sections.

The results in Table 3 reveal significant patterns in student errors, indicating the need for targeted instructional strategies. A major finding is the frequent influence of Arabic linguistic structures on students' English writing, particularly among lower-intermediate learners, who demonstrated a higher number of errors due to L1 transfer.

Table 3. *The frequencies of errors committed by the participants*

The Most Serious Problem						Immediate errors					Punctuation and Mechanic problems		
Verb		Agr.	Ro. sentence	Comma S.	Frag.	Word order	Word Form	Art.	Prep.	Parall.	Pun.	Cap.	Sp.
No. of students	45	30	32	48	42	33	34	23	43	20	41	27	48
%	80.36 %	53.57 %	57.14 %	85.71 %	75 %	58.93 %	60.71 %	41 %	76.79 %	35.71 %	73.21 %	48.21 %	85.71 %
% of student errors	50 students= 81.86%					36 students= 64.29%					48 students= 85.71%		
Table abbreviation: Agr: Agreement, Run-on: RO, Comma Splice: Comma S., Parallelism: Paral., Punctuation: Pun., Capitalization: Cap, Spelling: Sp, Article: Art.													

The Most Serious Problems category, encompassing 81.86% of students, reveals errors that significantly impair reader comprehension. Chief among these are verb-related mistakes (80.36%), which reflect the challenges Saudi learners face with verb conjugation and tense usage due to

ISSN: 2229-9327

fundamental structural differences between Arabic and English. High occurrences of run-on sentences (53.57%) and agreement issues (53.57%) further underscore students' struggles with maintaining subject-verb and pronoun-antecedent consistency. The prevalence of comma splice errors (85.71%) points to difficulties in punctuating compound sentences effectively, possibly influenced by Arabic writing patterns. Additionally, a substantial number of phrase fragments (75%) suggests an inability to consistently construct complete, cohesive sentences.

Intermediate-level issues were also frequent, with 62% of students committing errors that impacted individual words. These include word order mistakes (58.93%) and incorrect word forms (60.71%), both of which illustrate challenges stemming from differences in Arabic and English syntax and morphology. Article errors (41.7%) highlight the difficulty of mastering definite and indefinite articles in English, which function differently than in Arabic. Preposition misuse was also widespread (76.79%), reflecting the influence of native language interference. Errors in parallel structure (35.71%) further indicate challenges in maintaining grammatical symmetry within sentences, affecting the clarity and rhythm of students' writing.

The Punctuation and Mechanics Problems category affected 85.71% of students, compromising both readability and coherence. Frequent punctuation errors (73.21%) demonstrate uncertainty in applying English punctuation rules, likely resulting from contrasts with Arabic conventions. Capitalization mistakes (48.21%) show inconsistent adherence to English standards, again pointing to cross-linguistic influence. Spelling errors, present in 85.71% of student work, indicate the phonetic discrepancies between Arabic and English, which complicate accurate word spelling for learners.

Taken together, the findings suggest that while Saudi students possess foundational knowledge and are capable of constructing both simple and complex sentence structures, they continue to encounter substantial grammatical difficulties across both proficiency levels. Errors in subject-verb agreement, verb tense, and sentence cohesion hinder clarity, while spelling, punctuation, and capitalization mistakes affect the overall readability and formality of writing. Such issues can negatively influence how learners' written work is perceived in academic and professional settings.

Moreover, these results offer key insights into areas requiring targeted instructional attention. The recurring errors—especially in verb usage, sentence construction, and mechanics—highlight the importance of instructional designs that not only focus on correcting surface-level mistakes but also emphasize deeper grammatical understanding. Addressing these areas will enhance learners' ability to communicate clearly and effectively, a skill critical for academic success and future employability in Saudi Arabia's increasingly competitive labor market.

Ultimately, these findings underscore the need for strategic educational responses that account for the structural and cultural differences between Arabic and English. Integrating explicit grammar instruction, error correction strategies, and comparative linguistic analysis can help learners overcome these persistent challenges. The upcoming analysis will delve deeper into the causes of the most commonly observed errors and illustrate each with authentic examples drawn from students' classroom writing.

Grammatical Error Analysis among Saudi EFL Learners

English grammar comprises various parts of speech—nouns, pronouns, adjectives, verbs, adverbs, prepositions, conjunctions, and others—whereas Arabic grammar operates primarily through three main categories: nouns, verbs, and prepositions. This structural disparity often

causes confusion when Arabic-speaking learners attempt to translate or apply their native language patterns to English, particularly when dealing with morphological rules. Errors in morphology, especially the omission of critical morphemes in nouns and verbs, are widespread among EFL learners. As Atashian and Al-Bahri (2018) highlight, grammatical errors remain prominent in second-language writing, often due to negative transfer or insufficient morphological understanding.

The Verb "To Be"

A major grammatical challenge for Saudi learners involves the verb "to be," particularly in the present tense. In Arabic, the present tense is less morphologically marked than in English, which leads many students to omit "to be" in present constructions. This omission is more frequent in complex sentences, especially those containing relative clauses or multiple tenses, where learners focus more on selecting the correct tense and neglect auxiliary verbs. Students frequently confuse or switch between tenses and are more likely to drop "to be" than other modals like "can" or "will" (Chomsky, 1995; Pinker, 1995). Stromswold (1990) also noted frequent auxiliary-related errors among Arab learners, such as incorrect question formation or unnecessary duplications.

Examples:

- Omission: *She studying today in class.*
- Addition: *She is studied in class yesterday.*
- Substitution: *I is hungry* / *Norah are my friend.*

Singular and Plural Nouns

The singular/plural distinction in English presents difficulties due to the morphological complexity of pluralization. Irregular plurals, in particular, lack consistent rules, making them harder to internalize (dos Santos Gonçalves et al., 2020). Arab learners often struggle with derivational and inflectional morphemes in plurals (Altakhaineh & Zibin, 2017). They tend to either omit plural markers or incorrectly apply them to uncountable nouns, which reflects the influence of their native linguistic patterns (Ridha, 2012). Errors such as switching singular and plural forms or misapplying gendered and dual noun structures suggest an intralingual influence, often overlooked due to a focus on content over form.

Examples:

- Omission: *Teacher gave me many quizz.*
- Addition: *I enjoy going to many swimming pools.*
- Substitution: *I have many book* / *A Muslim pray 5 time a day.*

Adjectives

Adjective placement errors are common due to syntactic differences between Arabic and English. Arabic typically places adjectives after nouns (N+Adj), whereas English follows the Adj+N structure (Abushihab et al., 2014). This reversal results in frequent word order mistakes. Additionally, Arabic adjectives must agree with nouns in gender and number (Ryding, 2005), adding another layer of complexity. Many of these errors stem from literal translation or a lack of understanding of comparative and superlative forms (Jakobovits, 1969).

Examples:

- Misordered adjectives: *House blue* / *The house smallest.*
- Number errors: *I saw houses 3 blue.*
- Comparative/superlative issues: *She is taller than her sister, she is the tallest.*

Prepositions

English prepositions are notoriously difficult for Saudi learners, primarily due to differences in function and meaning between the two languages. Unlike Arabic, where prepositions may be more semantically flexible, English prepositions often require contextual understanding and syntactic precision (Scott & Tucker, 1974). Common errors include omission, redundancy, and substitution (Noor, 1996), often caused by direct translation or misinterpretation (Albanawi, 2018). While locative prepositions are typically used correctly, others—especially abstract or idiomatic ones—pose challenges.

Examples:

- Omission: *I'm waiting (…) my teacher (…) her office.*
- Addition: *We returned to home.*
- Substitution: *She is sitting in at my chair.*
- Interference: *She laughed on me.*

Articles

Articles are another problematic area for Saudi EFL learners. While Arabic includes a definite article, it lacks an indefinite one, which leads to both overuse and omission in English (Swan & Smith, 2001). Ridha (2012) and Rass (2015) note that article usage remains one of the most persistent sources of error, even at advanced levels. Shukri (2014) found frequent omissions of indefinite articles among Saudi learners, highlighting negative transfer from Arabic. Article errors are particularly prevalent when students attempt to translate ideas directly from Arabic, where such linguistic elements are either absent or differently applied (Alhaysony, 2012).

Examples:

- Addition: *I like to eat many (a) burgers.*
- Omission: *I come from (the) Kingdom of Saudi Arabia.*
- Substitution: *My classmate started (the) college in September.*

Capitalization

Arabic writing does not differentiate between uppercase and lowercase letters, which contributes to widespread capitalization errors in English. These include inconsistent capitalization of titles, proper nouns, or the pronoun "I." According to Almurshedi (2014) and Ababneh (2017), such errors are due to unfamiliarity with the English orthographic system and limited practice with standardized capitalization rules.

Examples:

- *i am a student.*
- *burydah, unayzah.*
- *My Life In Saudi Arabia.*

Punctuation

Differences between English and Arabic punctuation systems result in frequent punctuation errors. Arabic typically uses fewer punctuation marks, with commas replacing full stops in many cases (Abdelmohsen, 2022). This leads to run-on sentences and improper use of commas and periods in English writing. As Abisamra (2003) noted, comma misuse is particularly common due to cross-linguistic influence.

Examples:

- Run-on: *I enjoy studying nursing nursing has 3 job responsibilities being patient and helpful to patients.*
- Series: *I love to drink coffee and tea and juice.*
- Coordinating conjunctions: *I don't want to study but my mother is forcing me to do so.*

A detailed analysis of grammatical, morphological, and mechanical errors made by Saudi EFL learners reveals the significant influence of Arabic linguistic structures on English writing. These errors range from misuse of the verb "to be" to challenges with pluralization, article usage, adjective placement, prepositions, punctuation, and capitalization. Most of these difficulties are rooted in negative transfer and intralingual interference, often compounded by literal translation strategies and limited morphological understanding. Understanding the nature of these challenges is critical not only for curriculum design but also for enabling learners to communicate effectively in academic and professional English contexts.

Factors Contributing to Errors

An analysis of the written work of 56 Saudi EFL students led to the classification of grammatical difficulties into three primary categories: Most Serious Problems, Immediate Errors, and Punctuation and Mechanical Issues. Using the analytical framework proposed by Ellis (1997)—comprising data collection, error identification, classification, analysis, and explanation—we systematically examined the nature of errors related to grammar, capitalization, and punctuation, as illustrated in the earlier examples.

The findings, summarized in Table 4, present a quantitative overview of the specific types and frequencies of errors. The most frequent error type was Addition, accounting for 55.36% of all errors, followed by L1 Interference at 51.79%. Omission errors comprised 32.29%, Substitution errors 21.43%, and Carelessness errors 10.71%. This distribution highlights the dominant linguistic challenges encountered by Saudi learners, with addition and first language interference emerging as particularly significant.

Saudi/Arab learners frequently experience difficulties with English grammar, capitalization, and punctuation due to several underlying factors. These include structural differences between Arabic and English, syntactic variation, and contrastive writing conventions. To effectively address these challenges and enhance writing proficiency, it is essential to implement targeted instructional strategies that directly respond to these root causes and promote greater linguistic awareness in the EFL context.

Table 4. *Distribution of factors causing errors*

Type of error	Percentage	Number of errors
Omission	32.29%	22

Addition	55.36%	31
Substitution	21.43%	12
Carelessness	10.71%	6
L1 interference	51.79%	29

Pedagogical Recommendations

Language acquisition is a gradual process shaped by trial and error. For Saudi and Arab learners, grammatical errors are not failures but necessary steps toward proficiency. Training teachers to recognize and address language errors, especially those that hinder communication, is essential (Alsahafi, 2017; Abdul Rahman, 2021). Systematic error analysis, such as that suggested by Divsar and Heydari (2017), helps educators identify common challenges and adapt instruction accordingly. Recognizing that errors often follow patterns (Smith, 2001) allows for more focused teaching. Traditional grammar-based teaching should be supplemented with authentic writing activities and self-editing exercises. These promote learner autonomy and help prevent fossilization of common mistakes while limiting L1 transfer. Given the structural differences between Arabic and English, targeted instruction on frequent problem areas—such as subject-verb agreement, verb forms, and punctuation—is crucial. Teachers should create structured, context-rich exercises that address these challenges.

Technology also plays a valuable role in enhancing learning. Language apps, grammar tools, and online platforms support independent practice and give immediate feedback (Shahbaz & Khan, 2017). These tools complement classroom instruction and cater to diverse learner needs. Cultural relevance is equally important. Drawing comparisons between Arabic and English structures enhances comprehension and engagement. When learners connect new language forms to familiar concepts, content becomes more meaningful. A supportive classroom environment that encourages questions and welcomes mistakes fosters greater confidence and motivation. Since Saudi and Arab learners face context-specific challenges shaped by L1 interference, teaching strategies must be tailored accordingly. Teachers can turn errors into opportunities for learning by providing clear explanations and consistent feedback. Finally, institutions may adopt mentorship or peer-assisted learning models to enhance student support. Pairing learners with more proficient peers builds confidence, reduces anxiety, and encourages continuous improvement. A balanced instructional approach—combining grammar instruction, digital tools, cultural awareness, and active engagement—can significantly improve writing proficiency among Saudi and Arab EFL learners.

Discussion

This study examined the grammatical challenges faced by 56 Saudi EFL learner, revealing persistent difficulties across three categories: Most Serious Problems, Immediate Errors, and Punctuation and Mechanic Problems, following Byrd and Benson's (1994) classification. The analysis showed that all participants exhibited errors, indicating the widespread nature of these issues among Saudi learners. Table 2 offered a detailed taxonomy of errors, providing practical guidance for instructors and curriculum developers.

Table 3 presents the Most Serious Problems were the most frequent, affecting 92.86% of the participants. Among these, verb-related issues, including tense misuse and incorrect conjugation, were particularly prevalent (80.36%). These results align with the structural contrast between English and Arabic, often contributing to tense confusion. Additional challenges such as

run-on sentences (53.57%), subject-verb agreement errors (53.57%), and comma splices (85.71%) further emphasize the syntactic and rhetorical differences learners face due to L1 influence.

Immediate Errors affected 62% of the learners and typically involved incorrect word forms and awkward sentence construction. Preposition misuse was especially frequent (76.79%), likely resulting from direct translation practices from Arabic to English. In the category of Punctuation and Mechanic Problems, high incidences were noted in spelling (85.71%), punctuation (73.21%), and capitalization (48.21%). These findings suggest not only mechanical challenges but also a limited awareness of English writing conventions.

The data strongly point to the need for focused instruction that addresses these specific problem areas. Teaching strategies should emphasize verb usage, sentence boundaries, subject-verb agreement, and punctuation accuracy. These findings are consistent with previous studies. Al-Zoubi and Abu-Eid (2014) reported similar difficulties among Arab learners, especially regarding verb tenses and L1 interference. Abdul Rahman (2021) and Alsahafi (2017) confirmed frequent issues in preposition use, articles, and verb forms. Smith (2001) also provided a broader perspective on the grammatical challenges Arabic speakers face when learning English.

However, contrasting viewpoints must be considered. Scholars such as Mahrous and Ahmed (2010) and O'Sullivan (2004) warned against overgeneralizing across Arab contexts, while Lado (1957) suggested that not all errors stem from L1 interference. Moskovsky and Picard (2018) even reported proficiency among Saudi learners in tense usage, while Alangari (2019) argued that effective instruction can reduce L1 transfer. Despite such claims, this study shows that grammatical errors persist even after direct instruction, indicating deeper linguistic or cognitive factors at play.

Conclusions

To conclude, this research offers valuable insights for improving EFL instruction. Understanding learners' error patterns can help develop more responsive teaching strategies. Future studies should incorporate a broader participant pool, diverse teaching styles, and linguistic software tools to enhance accuracy. Longitudinal tracking of students beyond the foundation level would provide a clearer picture of grammar development over time. This study's limitations include its small, homogeneous sample, reliance on two drafts for progress assessment, and absence of corpus-based analysis. Addressing these in future research will contribute to more comprehensive understandings of Saudi EFL learners' grammatical development.

Funding: This research is not funded.

Acknowledgments: The corresponding author extends the appreciation to the Deanship of Postgraduate Studies and Scientific Research at Majmaah University for funding this research work through the project number R-2025-2056.

Conflicts of Interest: The authors declare no conflict of interest.

Authenticity: This manuscript is an original work.

Artificial Intelligence Statement: The authors acknowledge the use of Grammarly to support grammar correction and structural refinement during the preparation of this manuscript. Prompts

were not used; rather, Grammarly's automated feedback was applied to enhance sentence clarity, punctuation accuracy, and overall linguistic cohesion. The output from this tool was reviewed and selectively incorporated at the authors' discretion. While the authors acknowledge the usage of AI, they affirm that they are the sole authors of this article and take full responsibility for the content herein, in accordance with COPE recommendations.

About the Author:

Rawan Ahmed Alsoweed is an English Lecturer at Sulaiman AlRajhi University in Saudi Arabia, where she teaches undergraduate English language courses in the Preparatory Year Program. She holds a Master of Arts in Applied Linguistics. Her research interests include writing pedagogy, teacher and learner autonomy, and the integration of AI tools in English language teaching and assessment. ORCID: https://orcid.org/0009-0001-5559-4235

Dr. Abeer Shujaa Alharbi is an Assistant Professor in the English Language Department at Majmaah University, Saudi Arabia. She earned her Ph.D. in Applied Linguistics and ELT from the University of Southampton, England. She teaches undergraduate and postgraduate students, supervises theses, and has published widely in her fields of interest. Her research focuses on sociolinguistics, semantics, CALL, teacher and learner autonomy. ORCID: https://orcid.org/0000-0003-3985-0785

Fahad Abdulaziz Alqusumi is an Analysis and Strategy Planning Expert at Qassim University Medical City, Saudi Arabia. His expertise includes statistical data analysis for academic and applied research, healthcare systems evaluation, strategic foresight, and organizational transformation. His research interests focus on the practical application of strategic planning to support medical city objectives and national healthcare goals. ORCID: https://orcid.org/0009-0008-6840-1202

References

Ababneh, I. (2017). Analysis of written English: The case of female university students in Saudi Arabia. *International Journal of Social Science Studies, 5*(4), 1. https://doi.org/10.11114/ijsss.v5i4.2264

Abdelmohsen, M. M. (2022). Arab EFL learners' writing errors: A contrastive error analysis study. *British Journal of English Linguistics, 10*(4), 1–17. https://doi.org/10.37745/bjel.2013/vol10n4117

Abdul Rahman, M. A. (2021). Investigating teachers' views on writing errors made by Saudi university students. *International Journal of Linguistics, Literature and Translation, 4*(11), 113–118. https://doi.org/10.32996/ijllt.2021.4.11.12

AbiSamra, N. (2003). An analysis of errors in Arabic speakers' English writings: Investigating writing problems among Palestinian students studying English as a foreign language [Doctoral dissertation, P5]. https://abisamra03.tripod.com/nada/languageacq-erroranalysis.html

Abushihab, E., Ramadan, S., Ababneh, T., & Abushihab, I. (2014). An analysis of spoken grammatical errors of Arab learners of English as a foreign language at Al-Zaytoonah University of Jordan. *Journal of Al-Ma'moon College, 24.* https://www.iasj.net/iasj/article/96110

Al-Zoubi, D. M., & Abu-Eid, M. A. (2014). The influence of the first language (Arabic) on learning English as a second language in Jordanian schools, and its relation to educational policy: Structural errors. *Sino-US English Teaching, 11*(5), 355–372. https://doi.org/10.17265/1539-8072/2014.05.006

Alangari, M. A. (2019). *A corpus-based study of verb-noun collocations and verb complementation clause structures in the writing of advanced Saudi learners of English* (Doctoral dissertation, University of Reading).

Albanawi, M. F. (2018). Error analysis of using prepositions: A case study of 3rd intermediate school at Almakarim school. https://doi.org/10.33850/1803-000-003-011

Alhamadi, N. S. (2015). *Language transfer errors in speaking among Saudi Arabia students: A comparative study between students in Saudi Arabia and students in the UK* (Doctoral dissertation, University of Leeds). https://etheses.whiterose.ac.uk/10744/1/Thesis%20N.%20S.%20Alhamadi.%20Ph.D.pdf

Alhaysony, M. (2012). An analysis of article errors among Saudi female EFL students: A case study. *Asian Social Science, 8*(12), 55–66. https://doi.org/10.5539/ass.v8n12p55

Alsahafi, N. (2017). An investigation of written errors made by Saudi EFL foundation year students. *UNSWorks*. https://doi.org/10.26190/unsworks/19853

Altakhaineh, A. R., & Zibin, A. (2017). The effect of incidental learning on the comprehension of English affixes by Arabic-speaking EFL learners: Acquisition and application. *Research in Language, 15*(4), 405–423. https://doi.org/10.1515/rela-2017-0023

AlTameemy, F., & Daradkeh, A. (2019). Common paragraph writing errors made by Saudi EFL students: Error analysis. *Theory and Practice in Language Studies, 9*(2), 178–187. https://doi.org/10.17507/tpls.0902.07

Atashian, S., & Al-Bahri, K. (2018). Towards Arab students' grammatical errors in academic writing & their perceptions. *Arab World English Journal, 1*, 140–145. https://doi.org/10.24093/awej/mec1.10

Byrd, P., & Benson, B. (1994). *Problem/solution: A reference for ESL writers*. Heinle & Heinle.

Chomsky, N. (1995). Language and nature. *Mind, 104*(413), 1–61. http://www.jstor.org/stable/2254605

Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics in Language Teaching, 5*(1–4), 161–170. https://doi.org/10.1515/iral.1967.5.1-4.161

Divsar, H., & Heydari, R. (2017). A corpus-based study of EFL learners' errors in IELTS essay writing. *International Journal of Applied Linguistics and English Literature, 6*(3), 143–150. https://doi.org/10.7575/aiac.ijalel.v.6n.3p.143

Ellis, R. (1997). *Second language acquisition*. Oxford University Press.

Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford University Press.

Ellis, R. (2010). Second language acquisition, teacher education and language pedagogy. *Language Teaching, 43*(2), 182–201. https://doi.org/10.1017/S0261444809990139

Gass, S. M., & Selinker, L. (Eds.). (1992). *Language transfer in language learning* (Rev. ed., Vol. 5). John Benjamins Publishing.

dos Santos Gonçalves, A., Guterres, C. F., & de Castro Guterres, L. D. N. (2020). A study on mastery of using plural nouns. *ISCE: Journal of Innovative Studies on Character and*

Education, 4(1), 1-11. https://www.iscjournal.com/index.php/isce/article/download/72/65
Greenbaum, S., & Nelson, G. (2009). *An introduction to English grammar* (3rd ed.). Pearson Education.
Hasyim, S. (2002). Error analysis in the teaching of English. *Kata, 4*(1), 62–74. https://doi.org/10.9744/kata.4.1.62-74
Hermini, H. (2014). Grammatical error analysis in recount text made by the students of Cokroaminoto University of Palopo. *Ethical Lingua: Journal of Language Teaching and Literature, 2*(1), 87–101.
Jakobovits, L. A. (1969). Research findings and foreign language requirements in colleges and universities. *Foreign Language Annals, 2*(4), 436–447. https://doi.org/10.1111/j.1944-9720.1969.tb00321.x
James, C. (1998). *Errors in language learning and use: Exploring error analysis*. Longman. https://doi.org/10.4324/9781315842912
Al-Shujairi, Y. B. J., & Tan, H. (2017). Grammar errors in the writing of Iraqi English language learners. *International Journal of Education and Literacy Studies, 5*(4), 122–129. https://doi.org/10.7575/aiac.ijels.v.5n.4p.122
Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. University of Michigan Press.
MacArthur, C. A. (1999). Overcoming barriers to writing: Computer support for basic writing skills. *Reading & Writing Quarterly, 15*(2), 169–192. https://doi.org/10.1080/105735699278251
Mahrous, A. A., & Ahmed, A. A. (2010). A cross-cultural investigation of students' perceptions of the effectiveness of pedagogical tools: The Middle East, the United Kingdom, and the United States. *Journal of Studies in International Education, 14*(3), 289–306. https://doi.org/10.1177/1028315309334738
Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Sage.
Moqimipour, K., & Shahrokhi, M. (2015). The impact of text genre on Iranian intermediate EFL students' writing errors: An error analysis perspective. *International Education Studies, 8*(3), 122–135. https://doi.org/10.5539/ies.v8n3p122
Moskovsky, C., & Picard, M. (Eds.). (2018). *English as a foreign language in Saudi Arabia: New insights into teaching and learning English*. Routledge. https://doi.org/10.4324/9781315688466
Noor, H. H. (1996). English syntactic errors by Arabic speaking learners: Reviewed.
O'Sullivan, A. (2004). Reading and Arab college students: Issues in the United Arab Emirates higher colleges of technology. In *Proceedings of the First International Online Conference on Second and Foreign Language Teaching and Research: Beyond Borders*.
Pinker, S. (1995). Why the child holded the baby rabbits: A case study in language acquisition. In *An invitation to cognitive science* (Vol. 1, pp. 107–133). MIT Press. https://www.academia.edu/download/31942238/pinker_why_the_child_holded_the_baby_rabbits.pdf

Raba, A. A. (2012). The most common punctuation errors made by the English and the TEFL majors at An-Najah National University. *An-Najah University Journal for Research-B (Humanities), 10*(1), 211–233. https://doi.org/10.35552/0247-010-001-010

Rass, R. A. (2015). Challenges facing Arab students in writing well-developed paragraphs in English. *English Language Teaching, 8*(10), 49–59. https://doi.org/10.5539/elt.v8n10p49

Ridha, N. (2012). The effect of EFL learners' mother tongue on their writings in English: An error analysis study. *Journal of the College of Arts, 60*, 22–45. https://www.iasj.net/iasj/download/e6fd0762cb4916a6

Ryding, K. C. (2005). *A reference grammar of modern standard Arabic*. Cambridge University Press. https://doi.org/10.1017/CBO9780511486975

Scott, M. S., & Tucker, G. R. (1974). Error analysis and English-language strategies of Arab students. *Language Learning, 24*(1), 69–97. https://doi.org/10.1111/j.1467-1770.1974.tb00236.x

Shahbaz, M., & Khan, R. M. I. (2017). Use of mobile immersion in foreign language teaching to enhance target language vocabulary learning. *MIER Journal of Educational Studies Trends and Practices*, 66–82.

Shukri, N. A. (2014). Second language writing and culture: Issues and challenges from the Saudi learners' perspective. *Arab World English Journal, 5*(3), 190–207. https://awej.org/images/AllIssues/Volume5/Volume5number3September/15.pdf

Smith, B. (2001). Arabic speakers. In M. Swan & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems* (pp. 195–213). Cambridge University Press. https://doi.org/10.1017/CBO9780511667121.014

Stromswold, K. J. (1990). *Learnability and the acquisition of auxiliaries* (Doctoral dissertation, Massachusetts Institute of Technology). https://cir.nii.ac.jp/crid/1130282270223693568

Swan, M., & Smith, B. (2001). *Learner English: A teacher's guide to interference and other problems* (Vol. 1, 2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511667121

Siddiqui, M. A. (2015). Evaluating capitalization errors in Saudi female students' EFL writing at Bisha University. *Arab World English Journal, 6*(1), 232–250. https://doi.org/10.24093/awej/vol6no1.19

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp. 124-147
DOI: https://dx.doi.org/10.24093/awej/vol16no4.7

On the Integration of Excellent Traditional Chinese Culture into Public English Teaching in Higher Vocational Colleges in China

Xia Peng
Jiangxi College of Foreign Studies, Nanchang City, China
Corresponding Author: 754736339@qq.com

Xinjian Hu
Foreign Languages College, Jiangxi Normal University, Nanchang City, China
Email: huxinjian168@sina.com

Caiyun Cheng
The Yaohu Campus of the Affiliated Middle School of Jiangxi Normal University
Email: 2466788255@qq.com

Zhengjun Yang
Foreign Languages College, Jiangxi Normal University, Nanchang City, China
Email: 939353194@qq.com

Received: 08/13/2025 Accepted:10/17/2025 Published:12/10/2025

Abstract
This research adopts a mixed-methods approach (structured questionnaire survey of public English teachers from some key vocational colleges in Jiangxi province in China, semi-structured interviews with public English department directors, literature review, and inductive analysis) to explore effective pathways for integrating excellent traditional Chinese culture into public English teaching in higher vocational colleges and address implementation dilemmas. It finds out that the integration is essential for fostering students' cultural confidence, safeguarding national cultural identity, and meeting the talent development objectives outlined in *the English Curriculum Standards for Higher Vocational Education (2021 Edition)*. However, current challenges include limited awareness and competence among public English teachers regarding cultural education, fragmented and superficial inclusion of traditional cultural elements in teaching materials, and teaching practices that are often formalistic and ineffective. To address these issues, it is crucial to promote collaboration among governments, colleges , teaching staff to explore cultural resources and develop diversified teaching materials, strengthen teacher training to improve their capacity for cultural integration, and innovate classroom instruction through the expansion of extracurricular and "cloud classroom" learning platforms. Given their effectiveness, these measures are recommended to enhance cultural integration efficacy and ultimately achieve the talent cultivation goals outlined in the *New Curriculum Standards*.

Keywords: Cultural education, excellent traditional Chinese culture, higher vocational colleges, public English teaching
How to cite: Peng, X., Hu, X., Cheng, C., & Yang, Z. (2025). On the Integration of Excellent Traditional Chinese Culture into Public English Teaching in Higher Vocational Colleges in China. *Arab World English Journal, 16*(4):124-147. https://dx.doi.org/10.24093/awej/vol16no4.7

Introduction

Xi Jinping, general secretary of the Communist Party of China (CPC) Central Committee, pointed out (2024,p.11): "The best of China's traditional culture is the lifeblood of the Chinese nation. It provides fertile soil for the core socialist values and the foundation for the nation to gain a firm foothold in global cultural interaction. Strengthening cultural consciousness and confidence is integral to strengthening confidence in the path, theory, and system of Chinese socialism." This statement profoundly clarifies the strategic significance of inheriting excellent traditional Chinese culture (hereinafter referred to as ETCC) for national development and cultural identity.

As a core public foundation course in higher vocational education, public English teaching directly influences the construction of cultural identity among vocational students through the integration of ETCC in its pedagogical practices. However, traditional public English teaching has long been dominated by an overemphasis on instrumental functions, demonstrating a tendency to prioritize instrumental value over humanistic cultivation, which significantly weakens the effectiveness of fostering holistic development through cultural education in such courses.

Against this backdrop, the *English Curriculum Standards for Higher Vocational Education (2021 Edition;* hereinafter referred to as the *New Curriculum Standards*) explicitly identifies "cultural literacy" as a core educational objective, requiring the integration of ETCC into English teaching. To respond to this requirement and address the instrumentalization dilemma, this study aims to construct a "trinity teaching model" that integrates English instruction, cultural inheritance, and value cultivation. By systematically exploring the selection of ETCC content, innovation of teaching methods, and optimization of evaluation mechanisms, this study intends to supplement the theoretical framework of ideological and political education in vocational public English courses, and provide practical pathways for cultivating high-quality technical talents with both ETCC grounding and global perspectives.

Research Purpose:

Aligned with ETCC inheritance strategies and the *New Curriculum Standards*, this study tackles vocational public English teaching's overemphasis on instrumentality over humanism, develops the "English instruction-cultural inheritance-value cultivation" trinity model, and offers pathways to nurture technical talents with Chinese cultural roots and global perspectives.

Research Questions:

What effective content selection, teaching method innovation, and evaluation mechanisms apply to the integration of ETCC into vocational college public English teaching?

How can the trinity model resolve the "instrumentalization" dilemma and boost the cultural educational effectiveness of public English courses?

Literature review

In recent years, Chinese scholars have conducted in-depth research on the *Three Teachings Reform* (covering teachers, teaching materials, and teaching methods) in higher vocational colleges. As of now, more than 600 articles indexed by CNKI, including over 40 articles from the *Peking University Core Journals List*. These research findings have reached a considerable level of depth. For example, Zhou & Chen (2019) and Liang & Zhang (2022) conducted comprehensive analyses of the background, core connotations, and practical pathways of the *Three Teachings Reform.* Wu & Zhang (2021) and Gong (2021) focused on practical challenges in the reform

process, such as the adaptability of teachers' competencies and textbook content, and proposed targeted solutions. Huang & Yao (2023)explored the new requirements and implementation paths of the *Three Teachings Reform* in higher vocational education.

However, research specifically addressing the integration of ETCC into public English teaching under the *Three Teachings Reform* in higher vocational colleges remains relatively limited: only more than 50 articles indexed by CNKI, and none of them are from core journals. Existing studies in this area primarily concentrate on two aspects: teaching practices and the integration of ideological and political education. For instance, Yu (2022) and Xu (2022) analyzed the implementation pathways of ideological and political education in public English courses within the context of the *Three Teachings Reform*. Ge (2020) identified key problems in current higher vocational English teaching. Liu (2025) systematically examined the significance, feasibility, and specific strategies for integrating traditional culture. Y. Zhang (2024) and C. Zhang (2025) conducted in-depth studies on the challenges and countermeasures of implementing the *Three Teachings Reform* in public English courses of higher vocational colleges under the context of industry-education integration.

Research by overseas scholars primarily centered on cultural issues in second language teaching. For example, Miller (1972) argued that many cultural challenges in second language instruction stemmed from conceptual misunderstandings, and that teachers should possess a background in philosophical analysis to enhance language courses. Osmancık (2021) investigated teachers' perceptions of culture and its integration in English as a foreign language classrooms. Chao (1991) studied how students from different Chinese normal universities acquired social and cultural knowledge about the United States through foreign language instruction. Yu (2018) explored the teaching and learning of culture in Chinese high school English classrooms. Lee (2024) pointed out that the imbalance between the emphasis on local culture and the cultures of the *United States* and the *United Kingdom* in English textbooks hinders the effective development of cross-cultural communication skills. However, no relevant foreign literature was found regarding the integration of ETCC into public English teaching in vocational colleges.

In summary, existing research has three primary deficiencies. First, foreign studies primarily emphasize the transmission of Western culture in second language teaching, with limited discussion on the integration of local cultures. Second, domestic research lacks systematic strategies for integrating traditional culture within the framework of the *Three Teachings Reform*. Third, scholars have given insufficient attention to external influencing factors, such as the roles of the government and enterprises.

Methodology

The researchers conducted a structured questionnaire survey (administered via "Sojump") and semi-structured interviews to collect data for this study. A total of 131 public English teachers from over 20 provincial-level key vocational colleges in Jiangxi province completed the questionnaires. In the meanwhile, the researchers conducted telephone interviews with public English department directors from 10 of these colleges, aiming to examine the current status of integrating ETCC into public English teaching practices at these local institutions.

The researchers conducted a comprehensive literature review to establish a theoretical foundation, which included scholarly works focusing on the integration of ETCC and the *Three Teachings Reform* in the context of vocational English education.

Inductive analysis was applied to the collected data to identify core challenges across the dimensions of "teachers, textbooks, and teaching methods." This analytical approach ensured that the proposed countermeasures were grounded, rigorous, and feasible.

Results

This study identifies three key dilemmas—related to teachers, textbooks, and teaching—that constrain the integration of ETCC into public English teaching in higher vocational colleges:

First, public English teachers demonstrate insufficient awareness of "cultural education," limited ETCC literacy, and inadequate competence in delivering cultural content. Survey results indicate that nearly 30% of teachers continue to prioritize language skill development; 57.25% have not received systematic training in ETCC literacy, and over 47% report insufficient knowledge or teaching ability in this domain.

Second, the textbook dimension is hindered by delayed compilation processes (with few textbooks specifically designed for higher vocational undergraduate programs), fragmented ETCC content, and low levels of digitalization.

Third, the teaching dimension exhibits three key issues: an underdeveloped instructional environment, ineffective pedagogical activities, and unsatisfactory outcomes. Students' limited English proficiency, combined with reduced class hours, imposes dual constraints on effective teaching. Moreover, 48.85% of surveyed teachers rate students' ability to express ETCC in English as average. Additionally, perfunctory second classroom activities and superficial engagement in online learning—referred to as "fake online learning"—result in a disconnection between cultural learning and its practical application.

The Necessity of the Integration

Language is not only a fundamental component of culture but also a vital medium for its transmission and development. Within the context of traditional Chinese culture, the notion that literature conveys moral principles reflects the unity of language's instrumental and intrinsic values, which aligns with the contemporary English education philosophy of holistic education. Integrating ETCC into public English teaching in higher vocational colleges is an essential requirement for fostering students' cultural literacy and their ability to inherit traditional heritage, reinforcing the subjectivity of Chinese culture, and aligning with the developmental needs of cultural inheritance in the new era. It also represents a necessary step toward achieving the talent cultivation goals outlined in the *New Curriculum Standards* for higher vocational public English.

An Essential Component of Cultural Identity Maintenance

"Language constitutes the concrete content and vivid expression of foreign language education, while culture represents its ultimate objective and intrinsic essence." (Zhang, 2017,p.7). Currently, English has become a global lingua franca and dominates foreign language education in China. Western languages and cultures have exerted a profound influence on Chinese English teachers and students, presenting a significant challenge to the inheritance and preservation of Chinese linguistic and cultural traditions. Cultural awareness within English curriculum instruction remains underdeveloped, and there exists a noticeable imbalance between Chinese and Western cultural representations. The *New Curriculum Standards* clearly state: "English courses in higher vocational education are an integral part of the curriculum system at vocational colleges and serve as compulsory or limited elective public basic courses for students across all majors,

embodying both instrumental and humanistic attributes." (MOE PRC Doc., 2021,p.1). This policy emphasizes that public English teaching in higher vocational education must incorporate cultural identity considerations and maintain a balanced representation of Chinese and Western cultures within language skill development. On the one hand, integrating ETCC into teaching helps students acquire the ability to "tell Chinese stories in English"; On the other hand, through the design of teaching materials and classroom activities, teachers should emphasize a comparative cultural perspective to guide students in critically understanding cultural differences and building a defense line for cultural identity.

The Intrinsic Requirement for Enhancing Cultural Confidence

Xi (2013) has pointed out successively that cultural confidence is "a more fundamental, pervasive, and enduring source of confidence" and that "a nation without spiritual strength cannot stand firm among the nations of the world." "Excellent traditional Chinese culture represents the wisdom and essence of Chinese civilization. It is the root and soul of the Chinese nation and serves as the foundation upon which we steadfastly uphold ourselves amidst the tides of global cultures." *The Recommendations for Implementing the Initiative on Inheriting and Developing China's Fine Traditional Culture* explicitly stipulates that "Educational institutions and educators should fully integrate excellent traditional Chinese culture into all aspects of ideological and moral education, cultural knowledge education, art and physical education, and social practice education, and permeate it through all stages of early childhood education, basic education, vocational education, higher education, and continuing education." (CPC & State Council Doc., 2017,p.7) .

The discussions and policy provisions require higher vocational public English courses to actively participate in cultural inheritance and innovation by incorporating ETCC into English teaching and cultivating students' cultural consciousness. By comparing cultural differences between China and the West (such as collectivism versus individualism, and harmony versus competition), teachers can guide students to understand the unique value of Chinese culture deeply. Through the English interpretation of traditional Chinese concepts such as "harmony between man and nature" and "harmony without uniformity," students' ability to articulate Chinese wisdom in English can be enhanced. This process of cultural comparison and bidirectional interpretation enables students to rediscover the contemporary relevance of Chinese culture within the context of globalization, and thereby strengthen their sense of Chinese cultural identity and national pride.

Measures to Achieve the Talent Cultivation Objectives

The *New Curriculum Standards* explicitly state that teachers should integrate the improvement of students' four core competencies--cross-professional communication in the workplace, multicultural exchange, language thinking enhancement, and autonomous learning improvemen--throughout the entire teaching process (MOE PRC Doc., 2021). This core competency requirement represents a specific and comprehensive demand for higher vocational English courses in the new era, reflecting the organic integration of professional knowledge transmission, work capability cultivation, and values-oriented education.

In the context of the deepening implementation of the Belt and Road Initiative and China's increasing role on the global stage, international economic cooperation and cultural exchange activities occur with growing frequency. Consequently, society has an escalating expectation for high-quality technical and skilled talents who are not only proficient in professional foreign language skills but also possess a deep understanding of the essence of Chinese culture.

To achieve this new goal of talent development, it is essential to deeply integrate ETCC into higher vocational public English teaching. In teaching practice, teachers should employ innovative teaching methods to guide students in completing tasks centered on language application and cultural dissemination, thereby promoting the coordinated development of language proficiency and cultural literacy. Only when students are capable of expressing themselves fluently in English and accurately conveying Chinese values can the curriculum goal of "cultivating technical and skilled talents with Chinese sentiments and international vision," as outlined in the *New Curriculum Standards*, be genuinely realized. These students can then become active and effective disseminators of Chinese culture in future international exchanges.

The Challenges of the Integration

In 2000, Professor Cong of Nanjing University pointed out that "there remains in English teaching in our country a significant issue –the aphasia phenomenon of Chinese culture." (Cong, 2000). In recent years, this issue has drawn extensive attention from the academic community. However, it remains unresolved within the context of higher vocational education. To gain a comprehensive understanding of the current situation, the authors conducted a survey across more than 20 representative higher vocational colleges, including those in the G10 Alliance in Jiangxi province, in July 2024, collecting a total of 131 valid questionnaires and conducting 10 in-depth interviews.

The results reveal that public English teachers in higher vocational colleges have exhibited a heightened awareness of "cultural education." Consequently, the aphasia phenomenon of Chinese culture in public English teaching has gradually improved. Nevertheless, an imbalance persists in the representation of Chinese and Western cultures. ETCC remains underrepresented in public English teaching, and its integration faces multidimensional challenges related to "teachers, teaching materials, and teaching methods."

Teachers Exhibit Limited Awareness, Knowledge Reserve, and Weak Ability

According to the survey data, 96.95% of the respondents agreed that educators should cultivate students' awareness of promoting ETCC during English instruction, and 54.96% agreed that "the classroom is the main arena for moral education, and we should place emphasis on enhancing students' English language application ability and cross-cultural communication competence." These findings suggest that the Concept–"English teaching is not merely language instruction but also talent cultivation through culture"–has gained a certain level of consensus among public English teachers.

However, some teachers still adhere to the pedagogical philosophy of "application as the goal," perceiving English primarily as a means of communication. They place excessive emphasis on the instrumental function of language while neglecting its cultural dimensions. The survey also revealed that nearly 30% of the respondents believed that public English teaching should focus on developing students' English language skills, 14.5% emphasized helping students absorb the essence of English culture and develop new ways of thinking, and 34.35% admitted to having insufficient awareness of communication. Under such cognitive biases, classroom teaching practice prioritizes language skill training and knowledge transmission, while the educational orientation of the English curriculum remains underemphasized (Wen & Zhang, 2021).

Most public English teachers in higher vocational colleges, who graduated from English majors, have internalized the phenomenon of "Chinese cultural aphasia." The Concept of "cultural education" is notably absent from their professional development, resulting in an inherent

deficiency in cultural literacy. In the prior English major curriculum, courses related to Chinese culture accounted for a low proportion—which were insufficient to meet teachers' needs for in-depth study of traditional cultural knowledge. The current survey reveals that 57.25% of respondents acknowledged inadequate instruction in core traditional Chinese cultural literacy during their undergraduate studies. After entering the workforce, public English teachers have minimal access to opportunities for improving their cultural literacy, and there is a severe lack of continuing education in this area. None of the teachers surveyed had participated in any state-level training programs related to cultivating talents through culture in the past five years; only two had attended provincial-level training sessions, and just 46.31% had participated in school-based or online training. The scarcity of such training resources makes it difficult for teachers to update their cultural knowledge or master effective cultural teaching methods. From the perspective of teachers' self-assessment, they also exhibit a significant lack of confidence in this aspect. Specifically, 63.36% of the surveyed teachers reported having received insufficient education in ETCC, and 42.75% rated their own literacy in this area as merely average. This situation not only constrains teachers' ability to instruct in the classroom, but also directly impacts the effectiveness of students' cultural literacy development.

Although 64.89% of the surveyed teachers reported that the content of Chinese and English culture was approximately equivalent in public English teaching, and 96.95% indicated that they assigned out-of-class activities—such as watching movies, listening to the radio, and viewing English videos —to help students learn ETCC, many teachers still prioritize improving students' English language application skills and boosting their pass rates on English proficiency tests, often at the expense of cultivating students' humanistic qualities (Wang & Tang, 2019).

In terms of teaching methods, most public English teachers continue to rely on the traditional "translation+explanation" model, which lacks innovation, resulting in a weak ability to cultivate students through culture. Such weak cultural cultivation capacity makes it difficult to effectively stimulate students' interest in learning traditional Chinese culture. According to 47.33% and 48.85% of the surveyed teachers, the primary reasons for the challenges in teaching ETCC in English classes are the insufficient knowledge reserve and the weak teaching ability of teachers in delivering ETCC. These shortcomings hinder teachers from flexibly applying diverse teaching methods and achieving the deep integration of language instruction and cultural education.

Textbooks Exhibit Deficiency in Integration, Content Incorporation, and Digitalization

Currently, due to the absence of a multiparty collaborative textbook development system involving education authorities, universities, industry enterprises, and cultural institutions, there is a lack of specialized policy guidance and financial support—and these parties have yet to establish an effective collaborative mechanism for textbook compilation. Compared with other disciplines, the development of public English textbooks for higher vocational education lags considerably behind. In 2023, the Ministry of Education announced the first batch of 7,251 national planning textbooks for vocational education during the 14th Five-Year Plan period (hereinafter referred to as the "14th Five-Year Plan"). Among these, only 45 were public English textbooks for higher vocational education, accounting for just 0.62% of the total selected, with only 24 being newly submitted titles.

In terms of the composition of public English textbook compilation teams, over 75% of the members are English language experts. At the same time, the participation of scholars in ETCC and vocational education remains significantly inadequate. This imbalance has led to limitations in the selection of cultural content and the representation of vocational characteristics in textbooks,

making it challenging to meet the evolving demands of public English teaching in higher vocational colleges.

The development of localized English textbooks also remains underdeveloped. The survey found that only a few provinces, such as Jiangxi, Shandong, and Guangdong, have developed English textbooks incorporating local cultural elements. However, these localized textbooks generally face issues such as limited quantity, low quality, and minimal usage.

Moreover, the development of public English textbooks specifically designed for higher vocational undergraduate programs is severely lacking. Among the national planning textbooks selected under the 14th Five-Year Plan, only one series–*"New Century College Applied English – Integrated Course1-3*," edited by Fu Yu from Shanghai International Studies University and published by Shanghai Foreign Language Education Press–was included. Currently, many higher vocational undergraduate colleges still directly adopt college English textbooks designed for regular undergraduate institutions. These textbooks fail to fully address the career development needs and learning characteristics of higher vocational undergraduate students and lack targeted integration of ETCC. Consequently, they are unable to effectively support the achievement of public English teaching objectives in higher vocational undergraduate education.

Although the national standard for public English textbooks in higher vocational education under the "14th Five-Year Plan" places greater emphasis on ideological and political education than previous editions and incorporates new content—such as fostering students' ability to tell Chinese stories and promote Chinese culture in English—there remain notable shortcomings in practical application. According to survey data, 49.62 percent of the teachers believe that the current textbooks provide insufficient coverage of ETCC. Furthermore, researchers widely view the integration of cultural knowledge as inadequate, superficial, and unsystematics, which limits its cultural influence on students. As a result, the incorporation of traditional culture in existing textbooks exhibits a "fragmented" feature. Take the *Fifth Edition* of the "14th Five-Year Plan" national standard textbook, *New Practical English Comprehensive Course 2* (Rev. ed.), as an example. Although the "Section V: Appreciating Chinese Culture" in each unit is specifically dedicated to presenting the English expression of ETCC, over 70% of this content is presented solely as after-class exercises. This fragmented arrangement undermines the coherence and integrity of cultural instruction, making it difficult to guide students in conducting an in-depth exploration of traditional cultural connotations. In terms of content distribution, there exists a problem of "singleness" in the integration of ETCC within current textbooks.

Currently, elements of ETCC predominantly concentrate in poetry, with insufficient representation of national characteristics and deeper cultural dimensions such as history and philosophy (Shao, 2024). For instance, in the *Fifth Edition* of *New Practical English Comprehensive Course 2* (Rev. ed.), content related to poetry translation accounts for over 50% of ETCC elements, and content focusing on food or festivals makes up nearly 70% of such elements. In contrast, current textbooks minimally incorporate institutional culture, spiritual culture, and revolutionary culture. The content structure fails to fully demonstrate the richness and diversity of ETCC, thereby hindering the cultivation of students' comprehensive understanding and deep appreciation of Chinese culture, as well as failing to meet the diversified cultural literacy requirements for technical and skilled talents in the new era.

Digital college English textbooks refer to a teaching system that is based on advanced information infrastructure and deeply integrates digital technologies—such as cloud computing, the Internet, and big data—with college English content. These systems serve teachers, students,

textbook developers, and researchers through interactive and personalized teaching platforms that provide extensive resources and comprehensive digital courses (Bao & Li, 2017). However, the development of digital textbooks for higher vocational English significantly lags behind current educational demands. Among the national standard textbooks for higher vocational public English during the 14th Five-Year Plan period, only two titles-*"Practical English 1, 2 Digital Course"* and *"New Practical English Digital Course 1, 2,"* published by Higher Education Electronic Audio-Visual Press Co., Ltd., were selected, accounting for a tiny proportion.

In terms of the format of existing digital textbooks, over 50% are remain static e-books (PDF, EPUB, etc.), which are essentially digitized replicas of printed textbooks that lack intelligent interactive functionalities. Such static digital formats only provide basic operational functions, such as text browsing and audio playback, and do not support meaningful interactive engagement between students and the textbook content. Consequently, they are unable to meet the growing demand for autonomous and personalized learning.

Moreover, most digital resources provided by mainstream teaching platforms (e.g., The Smart Education of China platform, Rain Classroom, and SuperStarLearn) remain at the "audio-visual" level, such as playing English listening materials and displaying cultural videos. These platforms lack pedagogically oriented advanced interactive functionalities, such as task-based immersive virtual simulation, and therefore remain unable to dynamically adjust teaching content in line with students' current proficiency levels in cultural comprehension.

In terms of resource content, data from the National Higher Vocational English Teaching Resource Library indicate that Western cultural resources account for two-thirds of the total, while digital English textbooks and supplementary materials related to local traditional culture and revolutionary culture are severely insufficient. Moreover, the update cycle for these resources is excessively long.

This lagging state of digital textbook development not only hinders the innovation of public English teaching models but also negatively impacts students' engagement with and dissemination of ETCC, making it challenging to meet the demands of education in the new era.

Teaching Exhibit Underdeveloped Environment, Superficial Participation, Low Effectiveness

Under the impetus of English teaching reform and the *New Curriculum Standards* higher vocational colleges have preliminarily established a "cultural education" framework. However, they face multiple constraints in policy implementation, which significantly hinder the deep integration of ETCC with public English teaching.These challenges primarily manifest themselves in two key dimensions: the structural imbalance within the sociocultural ecology, and the dual constraints arising from the learning context and class hours.

The current social cultural ecology is under dual pressures from historical cultural discontinuity and the utilitarian orientation of the market economy. Modernization and digitalization have impeded the intergenerational transmission of traditional culture, leading to a weakened sense of cultural identity among the younger generation, which in turn undermines the foundation for its continued inheritance. Although the *Annual Report on the Quality of Vocational Education in China (2023)* indicates that since 2022, vocational colleges have consistently implemented a series of initiatives, such as "Contributing to a strong country through skills and success"–including skill-based inheritance of ETCC, which could serve as crucial measures to address existing gaps and foster a supportive cultural atmosphere. However, survey data still reveal significant shortcomings: 13.74% of surveyed teachers perceive low social recognition, and 35.11% believe the policy environment remains imperfect. These issues directly limit students'

opportunities to engage with traditional culture through English expression. The frequency of such cultural encounters in daily and campus contexts has notably declined, making it difficult to cultivate an integrated learning environment that combines traditional culture with English, not to mention establishing a continuously immersive cultural atmosphere on campus.

Under the dominance of the "certificate + examination" evaluation system, public English teaching in higher vocational colleges continues to emphasize language proficiency, with CET-4 and CET-6 pass rates serving as the primary performance indicators. In contrast, educational practices generally overlook the integration of ETCC into English instruction. Many higher vocational colleges directly link English proficiency test scores to students' graduation eligibility and the evaluation of teaching staff. This assessment orientation compels teachers to prioritize exam-oriented instruction, thereby reducing the time and opportunities for designing and implementing content related to traditional culture. Consequently, students tend to focus on obtaining certificates rather than engaging with cultural content, fostering a prevailing mindset of "emphasizing exams over culture." This tendency further exacerbates the challenges associated with integrating traditional culture into English education.

Vocational college students' learning characteristics, such as a weak English ability and low motivation for self-study, when combined with the current reduction in public English class hours, form a dual constraint. The generally low English proficiency among higher vocational students directly hinders their ability to comprehend English expressions about traditional culture. According to research data, 69.13% of surveyed teachers identified students' poor English foundation as the primary barrier to integrating ETCC into English instruction. At the same time, many vocational colleges have generally cut public English class hours from four to two sessions per week. With this significant reduction in instructional time, teachers struggle to design and implement teaching activities systematically, while students lack sufficient opportunities to engage in the cultural activities deeply. These overlapping challenges make it difficult to achieve the goals of culture cultivation effectively and undermine the overall teaching outcomes significantly.

Although many public English teachers in higher vocational colleges have actively explored the integration of ETCC into English instruction, some problems, including rigid teaching methods, monotonous instructional formats, and a lack of substantive engagement between teachers and students in teaching activities, still deserve attention.

Many public English teachers continue to rely on traditional lecture-based approaches that emphasize rote knowledge transmission. The incorporation of ETCC often appears "occasional" and "fragmented," lacking a systematic design, which hinders students from fully and deeply grasping the core values of ETCC and impedes the realization of effective cultural education.

Many higher vocational colleges organize activities such as "English corners" with themes centered on ETCC, and also host English speech contests and ETCC-themed cultural performances during traditional festivals like the Spring Festival and Mid-Autumn Festival. However, some of these activities demonstrate superficial engagement at the expense of cultural depth and pedagogical quality. Consequently, these activities often devolve into symbolic "performances" that fail to enhance students' cultural literacy or cross-cultural communication competence substantially. Although many higher vocational colleges have adopted the "online + offline" teaching model through platforms such as SuperStar Learn and Rain Classroom, these platforms are often underutilized. In practice, students primarily use them as tools for out-of-class self-study, with minimal supervision or guidance provided during the learning process. A widespread phenomenon of "Online course completion through automated methods" exists, where students

mechanically complete tasks such as watching videos, without genuinely engaging with the cultural content. Furthermore, relevant parties have not effectively integrated online teaching and classroom instruction, resulting in the problem of disconnection. For example, the learning tasks related to ETCC assigned on online platforms are rarely reviewed or discussed in classroom settings. This disconnection prevents students from consolidating and applying the knowledge they have acquired online, significantly undermining the effectiveness of the integration efforts.

Currently, the integration of ETCC into public English teaching in higher vocational colleges has not achieved the desired outcomes. Many vocational college students demonstrate only a superficial understanding of traditional culture, with limited comprehension of its more profound connotations and values. According to survey results, 48.85% of the teachers surveyed assessed students' ability to express ETCC in English as merely average, while 48% rated it as "poor" or "very poor." That means many students are still struggling to use English to introduce and explain ETCC.

This lack of teaching effectiveness directly hinders the achievement of core literacy goals outlined in the *New Curriculum Standards* such as "cross-professional communication in the workplace, multicultural exchange." As a result, it not only constrains students' career development potential but also fails to meet the demands for high-quality technical and skilled talents in the new era, highlighting the urgent need to enhance the integration of ETCC into public English teaching in higher vocational education.

Practical Approaches to the Integration

The integration of ETCC into public English teaching in higher vocational colleges is a systematic endeavor that requires coordinated efforts from multiple stakeholders, including the government, colleges and universities, teachers, and students. It should be advanced through comprehensive measures in textbook development, teacher training, pedagogical reform, and resource support, while fully leveraging students as central participants, teachers as facilitators, governments and schools as supervisors, textbooks as carriers, and teaching methods as mediators （Peng, X., Peng, C., & Zhang, B., 2024）. Only in this way can effective integration be realized.

Textbook Selection Path: Development and Selection of High-Quality Teaching Materials

English textbooks serve as crucial carriers of teaching content and form the foundation for effective English instruction in higher vocational education. According to the *New Curriculum Standards*, higher vocational English textbooks should reflect professional characteristics, integrate ETCC, and align with the language learning patterns of vocational students. (MOE PRC Doc., 2021). Therefore, the development and selection of high-quality textbooks is essential. Strict quality control over textbooks and the effective integration of ETCC into public English textbooks are key prerequisites for achieving the talent cultivation objectives. To ensure this, education authorities should invite experts and scholars in English education and traditional culture, along with representatives of teachers from higher vocational colleges, to jointly participate in English textbook reforms(Li, 2024).

Government departments and higher vocational colleges should actively mobilize diverse forces and adhere to the principles of classicism and readability to thoroughly examine the core ideas, traditional virtues, and life philosophies embedded in ETCC. For example, they should focus on classic cultural elements centered around fundamental concepts such as "benevolence, people-oriented governance, integrity, justice, harmony, and the pursuit of the Great Unity," as well as traditional virtues such as "self-improvement, dedication to one's profession and social

responsibility, helping those in need, acting bravely for justice, and respecting the elderly and cherishing family tie" (*Document of the CPC Central Committee and the State Council*, 2017). Furthermore, educators should pay attention to the ideological and moral education themes present in Chinese cultural classics, such as the Confucian emphasis on "the moral character of the noble person " and "the unity of family, nation, and the world", as well as the teachings in *The Great Learning* on "governing the state and bringing peace to the realm" and "the learning of the great person."(Yang & Li, 2021).

In the process of compiling specific teaching materials, it is essential to accurately identify the intersection between cultural content and language teaching by establishing sections dedicated to the comparison of Chinese and Western cultures. For instance, in the *Fifth Edition* of N*ew Practical English Comprehensive Course 2* (Rev. ed.), Section V, "Appreciating Chinese Culture," in each unit could be revised to "Comparison of Chinese and Western Excellent Traditional Cultures." Taking Unit 2 as an example, food culture in the section can serve as an entry point for comparing cultural differences.

Regional ethnic culture and revolutionary culture are essential components of ETCC, embodying its unique historical memory and spiritual connotations of Chinese civilization. Integrating these cultural elements into higher vocational public English textbooks can not only realize the localization of teaching content but also stimulate students' learning interest, thereby fulfilling the dual objectives of cultural inheritance and the development of cross-cultural communication competence. To fulfill this integration, education authorities and higher vocational colleges should establish a multiparty collaborative mechanism for cultural resource exploration, aligned with the talent cultivation goals that emphasize "workplace international communication and multicultural exchange." This mechanism should take into account both the characteristics of English language teaching and the demands of cultural exchange and inheritance.

From the perspective of language instruction, English experts should assess the linguistic and pedagogical appropriateness of cultural content. At the same time, regional traditional culture specialists should identify and summarize cultural symbols and core values.

For example, in the unit about "Travel" of the *Fifth Edition* of *New Practical English Comprehensive Course 2* (Rev. ed.), Jiangxi culture can be effectively integrated by exploring Jingdezhen's ceramic culture and the region’s revolutionary culture heritage. English audio, video, and image materials about the historical development of Jingdezhen ceramics can be incorporated into the textbook, guiding students to describe the artistic features of ceramic craftsmanship. Additionally, Curriculum designers should integrate English-language historical texts and images of the Nanchang Uprising into instructional materials. Moreover, with Jiangxi (the Long March's starting point) as a focal point, they can develop interactive modules to guide students in narrating the Long March story in English. Through systematic exploration and integration of local cultural elements, such efforts can significantly elevate the educational value of teaching resources, thereby fostering students’ cultural confidence and deepening their identification with regional cultural heritage.

To meet the requirements of high-quality development in vocational education and effectively enhance the quality of public English teaching in higher vocational colleges, it is imperative to accelerate the development and adoption of new forms of English textbooks. The Ministry of Education of the People's Republic of China (2022) proposed precise requirements in its *Implementation Plan for the Construction of Planning Textbooks for Vocational Education* during the 14th Five-Year Plan Period, including advancing the development of textbook-related

resources and digital textbooks, exploring the digital transformation of printed textbooks, and developing more audio-visual, interactive, and practice-oriented digital textbooks. This policy highlights the nation's strong emphasis on the innovation of textbook formats and provides clear guidance for the reform of public English textbooks in higher vocational education.

Higher vocational colleges should fully leverage AIGC (Artificial Intelligence Generated Content) technology to expedite the development and enrichment of digital textbook resources, thereby constructing an English teaching system that integrates online self-study with offline, teacher-led instruction. For example, in Unit 3 "Travel" of the *Fifth Edition* of *New Practical English Comprehensive Course 2* (Rev. ed.), digital resources related to the grandeur of the Great Wall, the historical depth of the Imperial Palace, the exquisite craftsmanship of the Terracotta Warriors, and the travel narratives of Xu Xiake can be incorporated and presented through diverse digital formats such as videos, audio clips, texts, images, and virtual tours. Students can gain historical and cultural insights by watching high-definition documentaries, experience scenic spots through immersive virtual tours, and participate in interactive Q&A sessions and role-playing activities to deepen their comprehension and application of the content.

The application of these new digital teaching materials facilitates interactive and data-driven teaching models between teachers and students. Based on students' online learning data, teachers can monitor learning progress and provide timely, targeted offline guidance. The intelligent and personalized features of digital textbooks cater to the diverse learning needs of students, support teachers in implementing intelligent classroom management, enhance students' understanding of ETCC, and ultimately improve their cross-cultural communication competence.

Government departments and higher vocational colleges should collaborate in coordinating the selection and compilation of public English textbooks, ensuring that the content meets high academic standards and upholds the correct ideological orientation.

During the textbook compilation process, textbook compilers should maintain the principle of combining diversified development with unified standards. On the one hand, local regions should be encouraged to develop localized teaching materials that reflect regional characteristics, tailored to the demands of local industrial development and cultural resources, thereby expanding the sources of English teaching content. For example, higher vocational colleges in coastal cities can focus on international trade and port English, while those in regions rich in tourism resources can create materials that integrate local scenic culture with tour guide English.

On the other hand, relevant authorities should coordinate national resources and utilize the expertise of scholars in linguistics and education to develop high-quality national textbooks.

These textbooks should address the fundamental English learning needs of students across different regions and majors, and contribute to the standardization and enhancement of English teaching nationwide.

In teaching material selection, it is imperative to reject the practices of "borrowing without adaptation" and regional protectionism. The "uncritical adoption" approach often involves the use of unadapted foreign textbooks centered on Western culture, emphasizing foreign narratives while being seriously disconnected from Chinese society and values. This approach tends to cause misalignment in students' cultural identity. Regional protectionism, in contrast, results in the circulation of uneven-quality local materials into classrooms, undermining teaching quality.

Therefore, relevant authorities must strictly enforce the principle of "reviewing all selected materials" and establish a scientific and rigorous textbook review mechanism. These two measures will strengthen the significance of textbook selection, ensuring content accuracy, structural

coherence, and appropriate difficulty levels. Only in this way can high-quality teaching materials be guaranteed to enter higher vocational public English classrooms, thereby supporting students in improving both their English proficiency and cultural confidence.

Teacher Training Pathways: Enhance Teachers' Awareness and Competence through Multifaceted Approaches

In higher vocational public English teaching, public English teachers serve as the leaders of classroom instruction and the core force in achieving educational objectives. Their awareness and competence in "cultural education" are therefore of critical importance. Public English teachers with a strong foundation in ETCC not only contribute to their own professional growth and career development, but also promote the overall quality of education, acting as key agents in the dissemination and inheritance of ETCC. These dual contributions directly influence the formation of students' cultural identity and the effective transmission of traditional cultural values. (Cao, 2019). Therefore, higher vocational colleges must innovate training formats and implement multidimensional and systematic approaches to comprehensively enhance public English teachers' cultural literacy and strengthen their capacity.

Enhancing teachers' cultural literacy is a fundamental prerequisite for achieving the goal of "cultural education." Education authorities should fulfill their administrative responsibilities by formulating policy guidelines and industry standards, coordinating training resources, and establishing communication platforms. Colleges and universities should take primary responsibility by designing structured training programs that align with their institutional development and faculty planning. Teacher training design should prioritize developing teachers' cultural literacy and refining their pedagogical skills, covering both the traditional cultural knowledge system and effective strategies for integrating cultural elements into classroom instruction.

Training designers should develop innovative "Internet +" training models for the information age. They can build online platforms, integrate premium educational resources (exemplary courses, expert lectures, cultural materials) to enable teachers' flexible, anytime, anywhere self-directed learning, and leverage online tools for real-time discussion and Q&A sessions. They should also organize regular offline sessions (training, workshops, seminars) and invite cultural and educational experts for on-site guidance. Such systematic training can continuously enhance public English teachers' cultural literacy and inheritance capacity.

Experienced senior teachers and experts in traditional culture represent invaluable resources for the development of the public English teaching staff in higher vocational colleges. In contrast, young teachers often lack sufficient traditional cultural literacy and practical teaching experience. By fully leveraging the guiding and mentoring role of veteran teachers and experts, higher vocational colleges can significantly shorten the professional growth cycle of young teachers, enabling them to adapt more quickly to the demands of teaching positions.

Colleges and universities should maximize the organizational and coordinating role of the Public English Teaching and Research Section by actively implementing a mentorship model that integrates "senior teachers + traditional culture experts + young teachers." This approach fosters the formation of collaborative learning teams and promotes complementary strengths and coordinated development among different teacher groups. Within this collaborative framework, experienced senior teachers can guide young teachers in classroom management, instructional design, and the application of effective teaching strategies. Meanwhile, traditional culture experts,

with their deep academic knowledge, can assist young teachers in strengthening their understanding of ETCC and support the organic integration of cultural elements into English teaching content. This collaborative model also contributes to the cultivation of a positive learning environment and an effective mechanism for cultural transmission, thereby facilitating the rapid professional growth and comprehensive development of young teachers.

Collaborative lesson planning serves as an effective strategy for integrating teaching resources and promoting teacher development. By drawing upon successful models of collective and cloud-based lesson planning from ideological and political courses, and by fully utilizing both internal and external school resources, higher vocational colleges can overcome institutional barriers, facilitate the sharing of teaching experience and pedagogical insights, and better support higher vocational public English teachers in integrating ETCC into their instruction.

Given challenges in higher vocational public English teaching, including fragmented resources and uneven teacher development, education authorities or industry associations should coordinate by regularly organizing cross-institutional, cross-regional collaborative lesson planning for public English teachers. This coordination and organization will help build professional development platforms transcending geographical and institutional boundaries. Mechanism implementation requires integrating diverse educational resources: leading vocational colleges' public English teachers' practical teaching strengths, key universities' public English teachers' academic research capabilities, and ETCC education experts. This multi-layered, diversified resource system covers teaching practice, academic research, and cultural education.

In terms of specific implementation, we can adopt an innovative approach that combines "in-school + out-of-school" collaboration with "offline + online" integration. Within the school, the Public English Teaching and Research Section serves as the foundational unit, where teachers hold regular collaborative lesson planning sessions to conduct in-depth discussions on curriculum reconstruction, pedagogical innovation, and the integration of cultural elements. Externally, teaching teams can make use of online Cloud-based lesson planning platforms (e.g., Smart Education of China) to overcome geographical limitations, enable the sharing of high-quality teaching resources, and promote the exchange of diverse perspectives. Teachers can upload and share materials related to ETCC, engage in collaborative discussions on teaching strategies, and invite experts for remote guidance. Offline sessions, on the other hand, focus on practical application, promoting the deep integration of theory and practice through simulated teaching, classroom observation, and peer review.

Through multi-agent collaboration and multidimensional support, we can ultimately establish a multilevel and diversified lesson planning resource system that comprehensively covers teaching practice, academic research, and cultural education. This system will systematically enhance public English teachers’ awareness of cultural inheritance and their instructional capabilities, promoting the transformation of teachers from a single "professional" or "technical" role to a more holistic "teaching-research-cultivation" role. It will provide strong support for cultivating high-quality technical and skilled talents who possess both language proficiency and cultural confidence.

Teaching Innovation Pathways: Multidimensional Organization and Innovative Classroom Instructional Formats

Teaching activities serve as a vital medium for achieving educational objectives. In an era marked by continuously evolving pedagogical concepts and rapidly advancing technological tools,

innovating classroom instructional formats and enriching students' learning experiences have become critical to improving the quality of public English teaching in higher vocational colleges. Traditional teaching models are no longer adequate to meet students' diverse needs or effectively integrate traditional culture into instruction. Therefore, it is imperative to explore multidimensional pathways for teaching innovation, fully utilizing the classroom as the "main arena," expanding the educational role of the second classroom, and establishing "cloud classrooms" through new media platforms. This integrated approach will facilitate the organic integration of cultural inheritance and the development of language proficiency.

As the core component of public English instruction in higher vocational colleges, classroom teaching plays a central role in the effective integration of ETCC. Innovating teaching methods and optimizing teaching content are crucial to achieving this integration. Through the adoption of innovative approaches, teachers can not only effectively stimulate students' learning interests but also significantly enhance the overall pedagogical effectiveness of classroom instruction.

At the level of teaching method innovation, teachers should transcend the constraints of the traditional lecture-based model and integrate diversified strategies such as task-driven learning, situational teaching, and collaborative group work. Task-driven learning centers on traditional cultural themes, guiding students to enhance their language proficiency through practical assignments—such as creating promotional brochures for traditional Chinese festivals. Situational teaching establishes realistic contexts, such as international cultural exchange exhibitions, where students can practice presenting intangible cultural heritage in English, thereby strengthening their cross-cultural communication competence. Group collaboration engages students in debates on traditional cultural topics, promoting critical thinking while reinforcing teamwork abilities.

In terms of content optimization, the focus lies on integrating elements of ETCC into the English language curriculum. In vocabulary instruction, culture-loaded words such as "Yin-Yang," "Tai Chi," and "Calligraphy" can be explained in terms of their cultural meanings and historical origins. During grammar instruction, teachers can enhance students' understanding of grammatical rules by comparing how Chinese and Western languages convey traditional cultural concepts. Furthermore, teachers can design thematic classroom activities (such as "Telling Chinese Stories Well" and "Translating Chinese Culture") to enable students to experience cultural differences through language practice and engagement. These activities promote a deeper understanding and appreciation of Chinese culture, thereby fostering cultural confidence and reinforcing national identity (Chen, 2021).

The second classroom serves as a distinctive and complementary platform for deepening cultural education and improving cross-cultural communication competence within the context of public English teaching in higher vocational colleges. Through the innovation of activity formats and the integration of diverse resources, teachers can effectively establish an immersive and multidimensional environment for traditional culture learning.

In terms of activity innovation, colleges and universities can organize a diverse range of second classroom activities, such as lectures on ETCC, the establishment of cultural theme clubs, and the initiation of "Campus English Corners."—all designed to foster students' confidence in publicly communicating culture in English（Abdullah, A.T.H., Netra, I.M., & Hassan, I.H., 2024）. Furthermore, special events like "Festival Poetry English Reading Sessions" and "Traditional Food English Cooking Workshops" can be conducted in alignment with traditional Chinese festivals. To further enhance student engagement, schools can also host competitions such as the

"Chinese Traditional Culture English Microfilm Creation Contest" and "Traditional Handicrafts English Workshop" (Ding, 2021).

Resource integration is a crucial factor in optimizing the second classroom. Colleges and universities should enhance collaboration with external cultural institutions and enterprises to access high-quality educational resources. Partnerships with museums, art galleries, and other cultural venues can enable English-guided tours, allowing students to engage with the richness of traditional culture in authentic environments. In cooperation with foreign-related enterprises, "Cultural Exchange Internship Programs" can be implemented, providing students with practical platforms to promote Chinese culture in English. Furthermore, vocational colleges should integrate internal educational resources to facilitate the development of interdisciplinary courses. For instance, integrating English studies with art design can give rise to courses such as "Visual Representation of Traditional Culture and English Expression," thereby promoting interdisciplinary collaboration.

In the current era of digital transformation, reshaping education, new media technology provides unprecedented opportunities for innovation in teaching and learning.

Teachers can fully utilize the functional advantages of various new media platforms. Professional online education platforms such as SuperStarLearn are highly effective for constructing a structured curriculum framework. Teachers can upload courseware and video materials that integrate traditional cultural content, monitor students' learning progress through built-in tracking tools, and facilitate topic-based discussions via integrated forums. Video-sharing platforms like Bilibili, with their substantial reach and creative user ecosystem, are well-suited for producing short educational videos such as "One Minute to Understand the English Expression of Chinese Intangible Cultural Heritage." These can stimulate students' interest in learning and encourage them to independently create English content centered on cultural themes, thereby fostering an interactive learning environment.

In actual teaching practice, it is crucial to emphasize the coordination between online and offline instruction. Offline classes should concentrate on delivering core knowledge and facilitating in-depth interaction, guiding students in exploring the integration of traditional culture with English expression. Meanwhile, the online "cloud classroom" functions as an essential platform for previewing, reviewing, and extending learning. Students can access pre-class materials to acquire background knowledge of traditional culture and utilize the platform for post-class reinforcement and extended learning activities, thereby maximizing overall teaching effectiveness.

Conditional Assurance Path: Optimize the Teaching Support Mechanism

Enhancing teaching conditions and institutional frameworks is an essential prerequisite for promoting the deep integration of public English teaching in higher vocational colleges with ETCC. Under the dual pressures of reduced class hours and the new moral education requirements introduced by the *New Curriculum Standards* establishing a systematic teaching guarantee mechanism holds significant practical value for achieving the goal of "cultural education." This mechanism should be advanced in a coordinated manner across three dimensions: strengthening ideological identity, improving teaching conditions, and optimizing the assessment system. By clarifying the core responsibilities of governments, colleges and universities, as well as teachers and students, a comprehensive teaching guarantee system can be effectively constructed.

Teachers' understanding and acceptance of the Concept of cultural education directly influence the depth and effectiveness of its implementation in teaching practice. Transforming ideological cognition serves as the foundational starting point for the integration. Therefore, educational administrators should prioritize ideological guidance for both teachers and students. Through activities like case sharing and academic discussions, they can strengthen value consensus on the role of "cultural education" and activate teachers' awareness of cultural inheritance and students' intrinsic learning motivation. Governments and colleges should, based on documents such as the *Opinions on Implementing the Project for the Inheritance and Development of Excellent Traditional Chinese Culture* and the *New Curriculum Standards*, deepen policy interpretation and dissemination through multiple channels. Specifically, they can utilize special training sessions, policy interpretation meetings, and online learning platforms to systematically explain the role of cultural integration in enhancing cultural literacy. Such systematic explanation will guide society to fully recognize the importance of inheriting ETCC, which is crucial for improving the cultural literacy of the population, safeguarding national cultural identity, strengthening cultural soft power, and advancing the modernization of the national governance system and governance capacity (CPC & State Council Doc., 2017). It will also enhance the sense of identity and consensus among teachers and students regarding the "cultural education" function of public English courses and inspire their ideological and behavioral commitment to learning and inheriting ETCC.

The quality of public English teaching directly influences the effectiveness of integrating ETCC into textbooks, classrooms, and school environments. Survey results indicate that 36.64% of the surveyed teachers identify limitations in school teaching plans as a significant obstacle to incorporating ETCC into English instruction. Currently, public English courses in higher vocational colleges face dual pressures: reduced class hours and the new moral education requirements. To address these challenges, higher vocational colleges should proactively optimize the teaching guarantee system in two core dimensions: class hour planning and resource construction.

In terms of class hour arrangement, higher vocational colleges should reconstruct the curriculum system based on the demand for "cultural education," appropriately increasing the total number of public English class hours. Additionally, higher vocational colleges should incorporate "second classroom" activities such as English culture salons and traditional culture-themed speech contests to advance the integration. Concerning teaching resource development, they should allocate increased funding to promote the intelligent and digital transformation of public English instruction. For instance, institutions can establish digital cultural resource libraries including videos, audio clips, and e-books that integrate English-language materials related to the teaching of ETCC. Furthermore, higher vocational colleges can introduce emerging technologies such as virtual reality and augmented reality to create immersive cultural learning environments, thereby providing a solid material and technological foundation for the effective integration.

As a core component of educational evaluation, the assessment mechanism serves as a "guiding baton" in teaching practice, exerting a substantial influence on teaching direction, the allocation of educational resources, and the teaching and learning behaviors of both teachers and students.

Education authorities should establish a multidimensional evaluation system by incorporating the implementation process and outcomes of integrating ETCC into the assessment indicators. Educational supervision departments should enforce the primary responsibilities of colleges and

universities through regular inspections, targeted supervision, and other supervision mechanisms. At the institutional level, colleges and universities should develop a teacher evaluation mechanism integrating key cultural integration aspects (curriculum design innovation, accuracy of cultural explanations, educational effectiveness) into daily, annual, and tenure-based assessments, with clear, quantifiable scoring criteria.

About student assessment, it is necessary to reform evaluation content by appropriately increasing the proportion of knowledge related to ETCC and local culture in public English examinations and English proficiency tests. For instance, teachers can incorporate questions related to traditional culture in sections such as reading comprehension, translation, and writing. This approach encourages students to engage with traditional cultural knowledge actively, fully stimulates their enthusiasm and initiative in learning, and ultimately achieves the goal of promoting both teaching and learning through assessment (Jiang, 2022).

In the context of the deepening process of globalization and the deep-rooted concept of a community with a shared future for humanity, Chinese culture is becoming more and more welcomed in the world. As key institutions for cultivating high-level skilled talents that serve national strategies and local economic development, higher vocational colleges bear the significant responsibility of cultural inheritance and innovation. As a compulsory component of the higher vocational education system, public English courses effectively promote the development of four core competencies—cross-cultural communication for work, multicultural exchange, language thinking enhancement, and autonomous learning improvement—through the coordinated integration of teaching materials, instructional methods, and teacher guidance. Only in this way can the fundamental educational objective of "cultivating virtue and nurturing talent" in higher vocational public English courses be fully achieved.

Discussion

This study focuses on integrating ETCC into public English teaching in vocational colleges. Through an analysis of its necessity, identifying existing challenges, and proposing practical pathways, this research produces findings that hold both theoretical significance and practical value.

From the perspective of necessity, integrating ETCC into public English teaching in vocational colleges is an essential requirement for safeguarding cultural identity, strengthening cultural confidence, and fulfilling the training objectives outlined in the *New Curriculum Standards* for higher vocational public English. This direction aligns closely with the national strategy that emphasizes cultural confidence and promotes the inheritance and development of ETCC. However, in practice, a key challenge lies in balancing language instruction with cultural transmission within the skill-oriented framework of vocational education.

The identified practical challenges manifest as systematic obstacles across three dimensions. First, teachers often exhibit limited awareness of "cultural education" and inadequate cross-cultural teaching competence. These deficiencies stem from the lack of cultural education during pre-service training and inadequate in-service professional development. Second, textbook content is frequently outdated, featuring superficial integration of traditional culture and a lack of digital resources. Such deficiencies in content reflects an insufficient emphasis on integrating vocational characteristics with cultural depth in textbook design. Third, higher vocational colleges face teaching environment constraints from underdeveloped cultural activities and their limited effectiveness, both influenced by external factors including social-cultural ecology and assessment

mechanisms. Additionally, outdated teaching methods further impede progress. The interplay of these three factors constitutes structural barriers to the effective integration of traditional culture into public English teaching.

To address these challenges, this study proposes a collaborative mechanism involving "government, universities, teachers, and students" and a systematic promotion strategy focusing on "textbooks, teachers, and teaching methods." This integrated approach provides a multidimensional solution to the current predicament. Within the collaborative framework, the roles of "students as the main participants, teachers as the guides, and government and institutions as the supervisors" are clearly defined and effectively implemented. In terms of implementation, textbook development emphasizes a three-dimensional integration of "classic selections + local characteristics + vocational scenarios." Teacher training establishes a closed-loop system encompassing "pre-service education, in-service research, and enterprise practice." Meanwhile, teaching methods incorporate "virtual simulation technology" to create immersive cultural experience scenarios, thereby enhancing learner engagement and improving practical learning outcomes.

The successful implementation of these strategies requires addressing key issues such as effective coordination among stakeholders and the long-term sustainability of resources. Further in-depth exploration through case studies in vocational colleges is essential to refine and validate these approaches in future research.

Conclusion

This study demonstrates that the integration of ETCC into public English teaching in higher vocational colleges is not only an inevitable response to national strategies and the requirements of the *New Curriculum Standards* but also a crucial component in building a collaborative education system aligned with contemporary educational goals.

The existing challenges—such as teachers' limited cultural literacy, lagging textbook development, and diminished teaching effectiveness —originate from multiple sources: the absence of a cultural dimension in teacher training systems, the inadequate integration of vocational characteristics with cultural depth in textbook design, and the rigidity of evaluation systems and teaching methods.

To address these practical issues, this study proposes a "four-party collaboration" mechanism and a "three-dimensional advancement" pathway. Specifically, textbook development emphasizes the integration of "classics, local culture, and vocational scenarios"; teacher cultivation follows a "pre-service, in-service, and enterprise-based training" model; and teaching innovation focuses on "technology-enhanced instruction and scenario-based learning." Through these strategies, a tripartite framework of "English instruction, cultural transmission, and value cultivation" is established. This model effectively mitigates the prevalent overemphasis on technical skills at the expense of humanities in higher vocational English education, while facilitating the cultivation of technically competent talents equipped with both a strong Chinese cultural identity and a global perspective.

Higher vocational colleges are encouraged to advance educational reforms through strengthened multiparty collaboration and continuous optimization of teaching practices. Future research could focus on specific vocational contexts and investigate discipline-specific pathways for cultural integration. The insights gained from this study may also provide reference value for curriculum reform in other stages of education.

Funding: This study was supported by the 2023 Jiangxi Province Social Sciences "14th Five-Year Plan" Research Project (Grant No. 23YY11); the 2023 Jiangxi Province Social Sciences "14th Five-Year Plan" Research Project (Grant No. 23YY02); the 2024 Jiangxi Higher Education Society Project (Grant No. WY-D-006).
Acknowledgments: Not applicable.
Conflicts of Interest: The authors declare no conflict of interest.
Authenticity: This manuscript is an original work
Artificial Intelligence Statement: In this paper the authors only utilized Doubao and NetEase Youdao Translation Tools to assist with formatting adjustments during the revision process. The authors reviewed and edited the content as needed and they took full responsibility for the content of the publication.

About the Authors:

Xia Peng is an associate professor at Jiangxi College of Foreign Studies and a visiting scholar at Jiangxi Normal University. Her research focuses on public English teaching.
ORCiD: https://orcid.org/0000-0003-2952-6711

Xinjian Hu is a professor at Jiangxi Normal University and holds a Ph.D. His research interests include foreign language education, teacher education, computer-assisted language learning, and educational linguistics. ORCiD: https://orcid.org/0000-0002-9045-8132

Caiyun Cheng is a Senior High School Teacher of the Yaohu Campus of the Affiliated Middle School of JXNU, who has been teaching English in high school for more than 20 years. ORCiD: https://orcid.org/0009-0002-9981-9720

Zhengjun Yang is an associate professor at Jiangxi Normal University. His research focuses on translation studies and English language teaching theory.
ORCiD: https://orcid.org/0009-0004-9886-9552

References:

Abdullah, A.T.H., Netra, I.M., & Hassan, I.H.(2024). Difficulties Faced by Undergraduate Students in English Public Speaking at a Malaysian University. *Arab World English Journal*, *15*(1): 269-282. DOI: https://dx.doi.org/10.24093/awej/vol15no1.17

Bao, M., & Li, X. (2017). A Research on the Development of Digital College English Textbooks in the Information Environment. *Technology Enhanced Foreign Language Education,* (03), 80-84 + 96. DOI: https://doi.org/10.20139/j.issn.1001-5795.2017.03.013

Cao, G. (2019). University teachers' excellent traditional cultural Literacy Cultivation. *Heilongjiang Higher Education Research*, *37*(10), 102–105. https://doi.org/10.19903/j.cnki.cn23-1074/g.2019.10.021.

Chao, H. (1991). The teaching of culture in foreign language education: A Chinese perspective (Order No. 9212270). (304007901). Retrieved from https://www.proquest.com/dissertations-theses/teaching-culture-foreign-language-education/docview/304007901/se-2

Chen, S. (2021). Research on Strategies for Integrating Excellent Traditional Chinese Culture into Higher Vocational English Teaching. *English on campus*, (11), 60-61.

Cong, C. (2000-10-19). "Chinese cultural aphasia": Deficiencies in English Teaching in China. *Guangming Daily*，001.

Ding, Y. (2021). An Exploration of Key Approaches to Integrating Excellent Traditional Chinese Culture into Higher Vocational English Teaching. *English on campus*, (35), 57-58.

Editorial Committee of New Practical English. (2023). *New Practical English Integrated Course 2* (Rev. ed.) (5th ed.). Higher Education Press.

Ge, M. (2020). Exploring the Practical Path of Integrating Excellent Traditional Chinese Culture into College English Teaching in Higher Vocational Colleges from the Perspective of Cultural Consciousness. *Modern English*, (02), 97–100 + 107.

Gong, W. (2021). Problems and Paths Faced by the *Three Teachings Reform* in Higher Vocational Colleges. *Hebei Vocational Educatio*n, *5*(06), 9–13.

Huang, Y., & Yao, J. (2023). The New Requirements and Realistic Paths of the *Three Teachings Reform* in Higher Vocational Education from the Perspective of Science-education Integration. *Education Research Monthly*, (10), 36–42 + 50. https://doi.org/10.16477/j.cnki.issn1674-2311.2023.10.014.

Jiang, L. (2022). Research on the Necessity and Innovative Approaches of Integrating excellent traditional Chinese culture into Public English Teaching in Higher Vocational Colleges. *Overseas English*, (18), 196-197+214.

Lee, J. F. K. (2024). Integrating culture in English language instruction: A comparative study of Hong Kong and Shanghai ELT textbooks. *The Journal of Asia TEFL, 21*(4), 839-857. http://dx.doi.org/10.18823/asiatefl.2024.21.4.5.839

Li, X. (2024-04-26). Integrating Chinese Fine Traditional Culture into Higher Vocational English Teaching. *Xinhua Daily*, *015*.https://doi.org/10.28872/n.cnki.nxhrb.2024.012107.

Liang, N., & Zhang, X. (2022). Analysis of the Value, Mechanism, and Path of the *Three Teachings Reform* in Higher Vocational Colleges. *China Higher Education*, (11), 59-61.

Liu, L. (2025). Research on the Practice of Integrating excellent traditional Chinese culture into Public English Teaching in Higher Vocational Colleges. *English Journal for Middle School Students*, (04), 145–146.

MILLER, H. N. (1972). *The Concept Of "culture" In The Curriculum Of Second Language Teaching* (Order No. 7302076). (302714889). Retrieved from https://www.proquest.com/dissertations-theses/concept-culture-curriculum-second-language/docview/302714889/se-2

Ministry of Education of the People's Republic of China. (2021). College English Curriculum Standards for Higher Vocational Education (2021 edition). *Bulletin of the Ministry of Education of the People's Republic of China*, 6, 131 + 211-306.

Notice of the General Office of the Ministry of Education of the People’s Republic of China. (2022). On the Issuance of the Implementation Plan for the Construction of Vocational Education Planning Textbooks during the 14th Five-Year Plan Period. *Gazette of the Ministry of Education of the People’s Republic of China*, (Z1), 29-33.

Osmancık, Gülin. (2021). *English Teachers' Perceptions of Teaching Culture in the EFL Classroom.* Master’s thesis. Bursa Uludag University (Turkey), Retrieved from https://www.pqdtcn.com/thesisDetails/FAD15D587E570B0A48B0B1CDDFD90550

Peng, X., Peng, C., & Zhang, B. (2024). Development of *Public* English Teachers in Chinese Vocational Colleges via the Mechanism of “One Body, Two Wings, and Three Parties

Collaborations." Arab World English Journal, 15(3): 379-394. DOI: https://dx.doi.org/10.24093/awej/vol15no3.23

Shao, D. (2024). Analysis on the Phenomenon of "Chinese cultural Aphasia" in Higher Vocational Public English Teaching. *Journal of Hubei Open Vocational College*, *37*(06), 174-176.

The General Office of the State Council. (2019). Recommendations for Implementing the Initiative on Inheriting and Developing China's Fine Traditional Culture. *Studies on Excellent Traditional Chinese Culture*, (00), 3-13.

Wang, F., & Tang, J. (2019). Reflection and Practice on the Harmonious Integration of The Instrumentality and Humanistic Nature in Higher Vocational Public English Teaching *Heilongjiang Higher Education Research*, *37*(11), 96-99. https://doi.org/10.19903/j.cnki.cn23-1074/g.2019.11.020.

Wen, Q., & Zhang, H. (2021). Interpreting English Curriculum Standards for Higher Vocational Education (2021 Edition): The Rationale, Development Process and Distinctive Features. Foreign Language World, *18*(05), 4-11. https://doi.org/10.13564/j.cnki.issn.1672-9382.2021.05.013.

Wu, Q., Yan, Z., & Hu F., et al. (2021). The Mission and Path of the *Three Teachings Reform* under the Background of High-Quality Development of Vocational Education. Chinese Vocational and Technical Education, (23), 58–63.

Xi, J. (2013-05-05). Speech at a Symposium with Outstanding Young Representatives from All Walks of Life. *People's Daily*, 002.

Xi, J. (2024). Speech at the Forum on Literature and Art(qstheory.cn). DOI: http://en.qstheory.cn/2025-01/13/c_1063923.htm

Xu, M. (2022). Exploration of Moral Education Paths for the Public English Curriculum in Higher Vocational Colleges under the Reform of "Three Education Elements". *The Guide of Science & Education*, (16), 95–97. https://doi.org/10.16400/j.cnki.kjdk.2022.16.030.

Yang, Z., & Li, Y. (2021). On the Construction of Translation Courses with Ideological - Political Elements. *Foreign Languages in China*, *18*(02), 104-109. https://doi.org/10.13564/j.cnki.issn.1672-9382.2021.02.020.

Yu, L. (2022). On the Strategies for Ideological and Political Theories Teaching in College English in Higher Vocational Colleges under the "Three Reforms of Teachers, Teaching Materials and Teaching Methods". *Journal of Hubei Open Vocational College*, *35*(10), 115–116 + 119.

Yu, Y. (2018). A Study on Culture Teaching and Learning in EFL Classrooms of Chinese Higher Vocational Colleges (Order No. 10976345). (2169098338). Retrieved from https://www.proquest.com/dissertations-theses/teaching-learning-culture-english-as-foreign/docview/2169098338/se-2.

Zhang, C. (2025). The *Three Teachings Reform* of Public English in Higher Vocational Colleges from the Perspective of Industry - Education Integration: A Case Study of Pingliang Polytechnic College. *English Square*, (17), 73–76. https://doi.org/10.16723/j.cnki.yygc.2025.17.018.

Zhang, S. (2017). Cultural Awareness in Foreign Language Education in China. *Foreign Language Education*, *38*(02), 7-11. https://doi.org/10.16362/j.cnki.cn61-1023/h.2017.02.002.

Zhang, Y. (2024). Difficulties and Countermeasures for the *Three Teachings Reform* n Higher Vocational Public English Under Industry-Education Integration. *The Guide of Science & Education*, (08), 41–43. https://doi.org/10.16400/j.cnki.kjdk.2024.8.014.

Zhou, J., & Chen Z. (2019). The *Three Teachings Reform* in Higher Vocational Colleges: Background, Connotation, and Path. *China University Teaching*, (09), 86-91. https://doi.org/CNKI:SUN:JXCY.0.2019-09-019

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp. 148-164
DOI: https://dx.doi.org/10.24093/awej/vol16no4.8

Engagement in Thesis Writing: An Appraisal Analysis of EFL Student Writers' Evaluation of Prior Academic Works

Lien-Huong Vo
Faculty of English, HUFLIS, Hue University, Hue City, Vietnam
Email: vtlhuong@hueuni.edu.vn

Long Viet Le
Faculty of English, HUFLIS, Hue University, Hue City, Vietnam
Email: longvietle0510@gmail.com

Mai-Duyen Pham
School of Foreign Languages, Can Tho University, Can Tho, Vietnam
Email: ptmduyen@ctu.edu.vn

Received: 08/09/2025 Accepted:10/17/2025 Published:12/10/2025

Abstract
Writing a thesis in a foreign language is challenging. In this process, evaluation of prior academic works in the Literature Review has always been a painstaking task that poses challenges for writers. Due to a paucity of studies on how writers evaluate other scholars' works in their Master of Arts theses, this study aims to explore how Vietnamese postgraduate students in English Language Studies and English Language Teaching programs utilize the *Engagement* system within the *Appraisal* framework to navigate through scholarly discourse in their Literature Reviews. A corpus of extracts from the Literature Review sections of 20 Master theses written in English was accordingly analyzed, using the UAM Corpus Tool. The findings show that the linguistic resources of *Heterogloss* outnumbered those of *Monogloss*. Strategies of *Expand* category such as *Attribute*, particularly *Acknowledge*, were widely used to engage with external perspectives. Meanwhile, the remaining linguistic resources were employed with significantly lower frequency, implying either hesitancy or limited familiarity with these rhetorical tools. These insights illuminate the dialogic engagement practices of Vietnamese novice academic writers and simultaneously point towards the need for explicit instruction in the use of *Engagement* to foster critical evaluation and rhetorical effectiveness. Finally, the study comes up with pedagogical implications and substantive conclusions regarding the use of these resources.

Keywords: Appraisal theory, dialogic discourse, engagement, English as a foreign language, thesis writing

How to cite: Vo, L., Le, L.V., & Pham, M. (2025). Engagement in Thesis Writing: An Appraisal Analysis of EFL Student Writers' Evaluation of Prior Academic Works. *Arab World English Journal, 16*(4): 148-164. https://dx.doi.org/10.24093/awej/vol16no4.8

Introduction

Writing research reports in a foreign language is never an easy task for writers, and evaluating other writers' arguments is even more challenging. It requires the writer not only a good command of the foreign language, but also an effective set of linguistic resources to express their position and persuade the audience. Research has proved that the writer's persuasiveness in making judgements is indicative of how they negotiate their opinions and attitudes with the reader, that is, the ability to make their position affective and reachable to the audience (Hyland, 2005; Lee & Deakin, 2016).

Xie (2016) argued that the *Appraisal* theory (Martin & White, 2005), an extension of Systemic Functional Linguistics, has been instrumental in analyzing how writers express their perspectives, opinions, and positions regarding various ideas or research subjects in their interaction with the audience. Associated with a system of evaluative language resources, the *Appraisal* theory provides an overarching framework to assess the writer's involvement in the text. Central to this framework is the concept of *Engagement*, which describes the linguistic mechanisms through which writers open or restrict the dialogic space to side with or counter opposing views. The framework, in this regard, offers a critical lens to investigate how writers navigate the dynamics between their arguments and existing academic discourse.

In a research report, the Literature Review (LR) section is seen as a rich source of evaluative language, where the writer takes account of the pros and cons of several related studies and justifies their work (Kwan, 2006; Kwan et al., 2012). Therefore, the LR has been investigated in much research in academic writing (Gil-Salom & Soler-Monreal, 2014; Naghizadeh & Afzali, 2018; Nguyen & Nguyen, 2024). There is, however, a paucity of research on how students of English as a Foreign Language (EFL), as novice academic writers, manage to judge others' views and persuade the audience of the arguments in their graduate theses. On the premise that a thorough understanding of the linguistic resources used by EFL writers can inform and improve academic writing instruction, the present study is conducted to examine how Vietnamese EFL students use the *Engagement* system within the *Appraisal* framework to evaluate prior academic works in LR sections of their Master of Arts (MA) theses. In so doing, it aims to shed more light on how novice academic EFL writers engage with scholarly discourse and thereby offers pedagogical implications for enhancing instruction in academic writing. The study, therefore, seeks answers to the following questions:

1. How do Vietnamese EFL students use Engagement resources in the evaluation of prior academic works?
2. What pedagogical implications can be drawn from the patterns of Engagement resource usage?

Literature Review

The Appraisal Theory

The *Appraisal* theory has been a powerful tool for the analysis of evaluative language. It constitutes part of a broader ecological understanding of language as a semiotic system that interactively engages with other systems, namely social, biological, and physical dimensions (Matthiessen & Teruya, 2023). Such arguments imply that the language being analyzed is not only as a tool for meaning-making but also as part of a dynamic system that interacts with various dimensions of human experience. Firmly rooted in interpersonal meta-functions posited in the Systemic Functional Linguistics, the *Appraisal* theory focuses on how language forges social

relationships and conveys attitudes. It comprises three systems of *Attitude*, *Engagement*, and *Graduation*.

Attitude is concerned with the lexical expressions of feelings, judgments, and values, including the categories that express emotions (affect), evaluations of behavior (judgment), and evaluations of objects or phenomena (appreciation). *Engagement* manages dialogic space, drawing on both lexical and grammatical resources to position the writer or speaker relative to other voices and viewpoints. *Graduation* focuses on scaling meanings, either by intensifying or softening evaluations (force) or sharpening and softening categorical distinctions (focus).

The Engagement System

Fundamentally, the *Engagement* system concerns itself with the interpersonal negotiation of sources, in close line with a dialogic perspective that examines how authors acknowledge, reject, or engage with alternative positions. This system categorizes utterances into two main subsystems: *Monogloss*, where authors present ideas without acknowledging other perspectives, and *Heterogloss*, where they introduce or consider alternative viewpoints. *Heteroglossic* utterances are further classified into two types: the *Dialogic Expand* type (extra-vocalization) or the *Dialogic Contract* type (intra-vocalization) (Martin & White, 2005). Figure 1 illustrates how the textual voice interacts with alternative voices and positions across these two dialogic types.

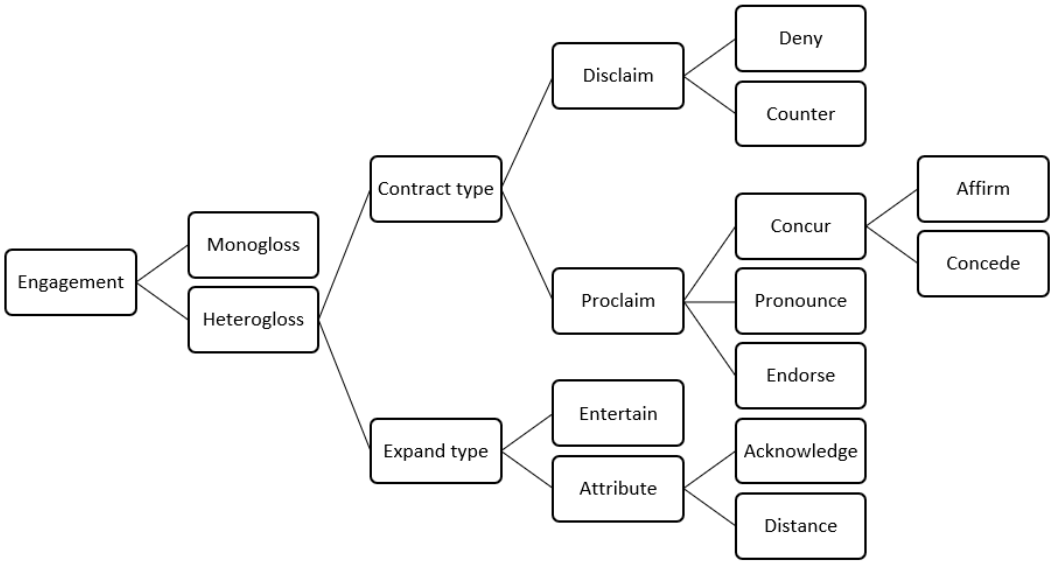


Figure 1. The Engagement system (Adapted from Martin & White, 2005)

In White's (2003) conceptualization, *Expand* pertains to the extent to which a locution allows for alternative viewpoints. Statements of the *Expand* type, such as *"some believe that"* or *"theorists propose,"* encourage audiences to entertain alternative perspectives. In contrast, *Contract* relates to how much the scope of alternative voices is constrained. Utterances of *Contract* type, such as *"everyone knows* or *it is a fact,"* work to restrict opportunities for valid dissent (Mori, 2017).

The subsystem *Heterogloss* comprises four key strategies for writers to position themselves dialogically in a text: *Disclaim* and *Proclaim* within the Contract type, and *Entertain* and *Attribute* within the *Expand* type. *Disclaim* is used to position the author's viewpoint as opposed to contrasting ones through *Deny* or *Counter*, as in the following examples from the corpus.

> (1) There is **no** [deny] universally accepted definition of lexical sophistication, as researchers defined "sophisticated" words in different ways.
>
> (2) **However** [counter], the analysis lacks the evidence to the researcher's finding which contributes a lot to his subjectivity of each character's representation.

Proclaim, including *Concur*, *Pronounce*, and *Endorse*, presents propositions as highly credible or valid and suppresses alternative perspectives.

> (3) The use of corpus is **undeniably** [concur – affirm] useful in terms of we can make generalizations about spoken and written discourse as a whole.
>
> (4) In general, almost every pupil expressed a wish to increase the variety of activities, […]. **Nevertheless** [concur – concede], one shortcoming remained in the study.
>
> (5) **It is clear** [pronounce] that the strategy is quite successful in enhancing students' speaking abilities.
>
> (6) Comoglu and Dikilitas (2020) **unveiled** [endorse] that Turkish EFL PSTs could perform a deeper critical analysis […].

Entertain creates dialogic openness by presenting propositions as subjective or contingent. *Attribute* invokes external voices, either neutrally (*Acknowledge*) or skeptically (*Distance*).

> (7) **According to** [acknowledge] Cheung (2010), listening should be the first and most important skill to learn when learning a new language.
>
> (8) He also **claimed** [distance] that a teacher who effectively integrates technology would be able to draw on extensive content knowledge […].

As this study examines Vietnamese EFL students' use of *Engagement* in the evaluation of prior academic works, the linguistic resources of *Engagement* are central to the analysis of how they construct dialogic interactions with audience in their academic writing, specifically within the LR sections of their MA theses.

Studies of Evaluative Language in Academic Texts

There has been much research in evaluative language in non-academic domains of discourse such as journalistic discourse (Luo et al., 2022), political discourse (Goudong & Afzaal, 2023; Lian, 2018), and social media interactions (Ross & Caldwell, 2020; Yuliyanti, 2023). In academic contexts, its use has also been the subject of scholarly interest. Studies have been exploring evaluative language in undergraduate essays (McKinley, 2018; Mori, 2017), graduate theses (Geng & Wharton, 2016; Gil-Salom & Soler-Monreal, 2014), and research articles (Kwan et al., 2012; Naghizadeh & Afzali, 2018; Nguyen & Nguyen, 2024; Zhang & Cheung, 2018). Given the scope of the present study, this review of literature concerns itself specifically with research on evaluation in academic texts.

Although several studies into the evaluative strategies in academic texts have shed light on how writers establish a research niche (Gil-Salom & Soler-Monreal, 2014; Kwan et al., 2012), they did not specifically explore linguistic resources within the *Appraisal* framework. Others that adopted the *Appraisal* framework often explored the systems of *Attitude* and *Graduation*. For example, Zhang and Cheung (2018) analyzed voice construction in the LR of research articles in computer networks and communications and second language (L2) writing through the lens of *Attitude* and *Graduation*. Pham (2024) later examined abstracts of research articles in economic fields based on the *Attitude* system of the *Appraisal* framework. Notwithstanding some insights into understanding evaluation across disciplinary writing, these studies did not address the *Engagement*. A few studies looked into the *Engagement,* but they steered their examination towards many different sections of academic texts, including introductions (Alotaibi, 2019; Chang & Schleppegrell, 2011), LRs (Amornrattanasirichok & Jaroongkhongdach, 2017; Naghizadeh & Afzali, 2018; Xie, 2016), discussions (Geng & Wharton, 2016; Loghmani et al., 2020), and conclusions (Loi et al., 2016).

Using Martin and White's (2005) *Engagement* system, Alotaibi (2019) investigated how authors in ranked and non-ranked journals in the Social Sciences Citation Index (SSCI) used linguistic resources to construct authorial voice in research article introductions. They found that both groups predominantly employed *Heterogloss*, but non-SSCI journals employed more *Monogloss*. Employing the same frameworks, Chang and Schleppegrell (2011) investigated the linguistic resources used in research article introductions in the social sciences, qualitatively analyzing patterns of *Expand* and *Contract* options in educational research. They identified two patterns for each type: Expanding dialogic space, which uses resources like *Entertain* and *Attribute* to explore multiple perspectives and establish a research territory, and contracting dialogic space, which employs *Disclaim* and *Proclaim* to assert strong authorial stances, dismiss alternatives, and emphasize the study's significance.

Regarding the LRs, there were three notable studies. Amornrattanasirichok and Jaroongkhongdach (2017) explored the use of *Engagement* in LRs from Thai and international applied linguistics journals. The findings showed that *Heterogloss* was used mostly in the corpora of this study. In particular, the dialogic *Expand* prevailed, in which the use of *Acknowledge* was of highest frequency. Naghizadeh and Afzali (2018) also compared *Engagement* in LRs of research articles published in Iranian local and international journals and found that international journals favored *Expand*, while local journals relied more heavily on *Monogloss*. Xie (2016) analyzed evaluation in 25 Chinese English-major MA thesis LRs. The study found a high proportion of *Monogloss*. Within *Heterogloss*, the *Expand* slightly outweighed *Contract*, with *Acknowledge* being the most used. Besides, *Entertain* was used as a hedging strategy in students' works, but their lexical variety was limited compared to expert writers, which called for improved instruction in academic evaluation.

Engagement was also examined across sections in research articles (Fryer, 2013; Nguyen & Nguyen, 2024; Loi et al., 2016). Fryer (2013) studied medical research articles and identified that *Expand* was used at twice the frequency of *Contract*, especially in the introduction and discussion sections, with *Entertain* being used most frequently. The most common *Engagement* strategies observed in this study were *Entertain* and *Disclaim.* It also noted a trend towards increased *Expand* in the discussion, while introduction sections frequently employed the *Counter* like *'however'* and *'although'*. In contrast, the *Contract* was dominant in the methods and results sections. The study by Nguyen and Nguyen (2024) revealed that Vietnamese authors prioritized

Contract over *Expand*, with *Disclaim* frequently used in introductions to highlight research gaps and *Proclaim* used in discussions and conclusions to affirm findings. The rarity of *Concur* and *Acknowledge* in the findings reflects a preference for asserting knowledge and minimizing external views. In a similar vein, Loi et al. (2016) analyzed the evaluative and dialogic stances in English and Malay research article conclusions using the *Appraisal* framework and genre analysis and they found that English conclusions employed more *Engagement* than Malay conclusions, and that Malaysian writers often employed *Contract*, which is arguably attributed to cultural and local academic influences.

As far as data sources are concerned, a number of studies looked into dissertations at the doctoral level. For example, Loghmani et al. (2020) examined how native English-speaking Ph.D. students in TEFL used the *Engagement* in their dissertation discussion sections. The findings showed these students relied more on *Contract* to assert authority and limit dialogic alternatives, in which *Counter* and *Deny* were the most frequently used. Less common resources, such as *Endorse* and *Acknowledge*, reflected the formal and dialogic nature of academic writing. Geng and Wharton (2016) investigated the *Engagement* in twelve discussion sections of doctoral theses written by first-language (L1) Chinese and L1 English writers. The findings revealed no statistically significant differences between the two groups in their use of *Engagement*. In addition, the use of *Heterogloss* was realized by the dominance of *Contract* type over *Expand* one. Within *Contract*, *Disclaim* was more common than *Proclaim*, whereas *Entertain* was the most frequently used within *Expand*.

This review discloses that how MA students make use of *Engagement* in their theses has captured little attention in research. The present study, therefore, addresses this gap by examining the use of *Engagement* within the LR sections written by MA students to enhance the understanding of dialogic strategies in academic writing and provide pedagogical recommendation for the teaching of academic writing within the EFL context.

Methodology

Context and Design

In Vietnam, postgraduate programs in English Language Studies (ELS) and English Language Teaching (ELT) are offered in English to broaden students' knowledge in the English language and provide them with knowledge and skills in ELT. As students must write a thesis on a relevant topic as part of the program requirements, they are provided with a course in Academic English with a focus on thesis writing. Nonetheless, due to the short duration of the course, there is little practice and their theses are arguably their first product of professional writing. To gain an in-depth understanding of the nuances of *Engagement* usage in their theses, the present study adopts a descriptive qualitative research design along with quantitative support from UAM Corpus Tool. This design is essential to achieve the goal.

The Corpus

The corpus comprises 20 MA theses by students majoring in ELT and ELS (submitted during 2022–2024) with approvals from the thesis writers. The data source ensures temporal relevance that reflects current academic writing practices in a particular context in Vietnam. The LR sections were then extracted, anonymized by assigning each LR to an identifier (LR1 to LR20), and converted to plain texts, which totaled a corpus of 115,214 words (M = 5,760.7; SD =

2,216.94). The corpus was built with meticulousness to ensure that the integrity of the data remained intact.

Data Processing and Analysis

Once the data source was shaped, the annotation and quantification of the *Engagement* system posited by Martin and White (2005) was conducted. This process was supported by the UAM Corpus Tool, which has been utilized in previous studies in the *Appraisal* framework (Fryer, 2013; Gheng & Wharton, 2016; Ngo, 2013), with the most updated version 6.2j as of 2023. Some linguistic realizations outlined by Naghizadeh and Afzali (2018) were also used to complement the *Engagement* realizations in Martin and White's (2005) work, which they acknowledged as not exhaustive. To ensure credibility, the annotation process strictly adhered to the *Appraisal* framework, and authors as coders regularly discussed and resolved discrepancies, if any, in coding to achieve a consensus on classifications.

As Martin and White (2005) pointed out, a challenge of Appraisal analysis is that the potential change of a word under the influence of co-textual factors. In this study, however, this challenge was overcome by meticulous examination of the co-text to ensure that the *Engagement* is categorized and analyzed accurately. Tabulation was also provided for a clear overview of the findings to facilitate further analysis.

Results

This section presents the findings, with the quantitative data being aggregated metrics for the entire corpus, then local subtypes of the *Engagement*. Reporting data this way can minimize misleading interpretations and obtain a more balanced overview of the data because some categories with limited occurrences might appear disproportionately significant when viewed in isolation as local percentages.

The distribution of Monogloss and Heterogloss

Statistically, the use of *Heterogloss* in LRs of Vietnamese EFL students' MA theses (78.8%) is strongly dominant over *Monogloss* (21.2%). Such an outstripping dominance demonstrates the preference for the *Engagement* in academic discourse. The findings resonate with those from studies by Amornrattanasirichok and Jaroongkhongdach (2017), who observed a relatively similar proportion of *Heterogloss* (79.58%) and *Monogloss* (20.42%) in LRs of local and international journals, and Naghizadeh and Afzali (2018), who also reported that both Thai and international research articles in applied linguistics favoured *Heterogloss* (67% and 83% respectively) in LRs.

Most *Monogloss* in this study's corpus was used when writers reported studies or research methods as facts, uncontested information. In (9) below, the proposition is presented as a factual report of the study's focus, with no acknowledgment of competing interpretations or additional perspectives. Similarly, (10) asserts methodological details without inviting dialogic alternatives, emphasizing the factual nature of the information.

> (9) The study by Rahmat and Dianita (2024) focuses on [monogloss] analyzing […].
>
> (10) In this study, a descriptive design and a mixed-methods approach were used [monogloss].

ISSN: 2229-9327

These examples highlight how *Monogloss* serves to establish the credibility of the review by presenting background studies and methods as uncontested knowledge. The usage ensures clarity and avoids overcomplicating the discourse, especially in contexts where presenting foundational information is essential for building the LR.

The use of Heterogloss

Among 1,874 instances of *Heterogloss* used in the corpus, *Expand type* (60.0%) was favored over *Contract type* (18.8%). Vietnamese EFL students, in this regard, primarily create space for alternative perspectives and engage more flexibly with prior academic works.

Table 1. *The distribution of Heterogloss*

Contract 446 (18.8%)		Expand 1428 (60%)	
Disclaim	Proclaims	Entertainment	Attribute
133 (5.6%)	313 (13.2%)	39 (1.6%)	1389 (58.4%)

These findings were consonant with those reported in previous studies (Amornrattanasirichok & Jaroongkhongdach, 2017; Fryer, 2013). Nevertheless, it is observed in other studies (Loghmani et al., 2020; Nguyen & Nguyen, 2024), that *Contract* type was preferred. This pattern was also evident in Geng and Wharton's (2016) study, where *Contract* realizations outnumbered *Expand* ones in the discussion sections of doctoral theses written by L1 Chinese and English writers. Such contrast can be attributed, in part, to writers' academic levels. For instance, Loghmani et al. (2020) and Geng and Wharton (2016) analyzed discussion sections written by doctoral candidates. At this advanced level, the writers might have aimed to assert their authority, emphasizing their original contributions, and thus probably relying more on *Contract* strategies.

The use of Contract types

As shown in Table 1, *Disclaim* accounts for 5.6% of the total *Engagement*, while *Proclaim* makes up a larger proportion of 13.2%. In the corpus, *Disclaim,* comprising 1.8% of *Deny* and 3.8% of *Counter,* was primarily used to establish research gaps. These figures resonate with the patterns observed in Nguyen and Nguyen's (2024) study, where a high frequency of *Disclaim* resources was used in the introduction sections of Vietnamese research articles to pinpoint research gaps. However, the findings contrast with results from Geng and Wharton's (2016) study, which found that *Disclaim* resources were slightly more common than *Proclaim* resources. The findings suggest that Vietnamese EFL MA students rely more on affirming and aligning with propositions rather than challenging alternative views in their LRs. This usage may reflect the students' attempt to build credibility by aligning with existing studies rather than overtly contesting them.

Table 2. *The breakdown of Disclaims and Proclaims*

Contract Subtypes		Occurrences	Percentage
Disclaim	Deny	43	1.8%
	Counter	90	3.8%
Proclaims	Concur (Affirm/Concede)	11	0.5%
	Pronounce	6	0.3%
	Endorse	296	12.4%

Table 2 provides a close look at the distribution of these two subtypes of *Contract* in the corpus, which reveals that these two types of *Disclaim* are often used in combination to establish

research gaps or point out weaknesses in previous studies, suggestive of a key rhetorical strategy in academic writing among the MA students. They primarily used *Deny* to place an explicit emphasis on the absence of specific studies or limitations in existing research and *Counter* to introduce their critique or alternative perspectives.

A combined use of *Counter ('nonetheless')* and *Deny ('little')* to highlight a research gap is evident in (11). The writer counters the implicit expectation that there has been significant research in this area while simultaneously denying the existence of sufficient studies. This rhetorical strategy effectively positions the writer's work as addressing a critical gap in the literature. In a similar fashion, in (12), *Deny ('without')* is used to critique a methodological oversight, while *Counter ('although')* accentuates this critique with an acknowledgment of other considerations. This combination demonstrates an elaborate evaluation when contributions and specific weaknesses in the cited research are recognized at the same time.

(11) […] **Nonetheless** [counter], **little** [deny] research delved into […].

(12) TK items in Lux et al.'s (2011) research were written **without** [deny] examining technological affordances of technologies, **although** [counter] TK items were considered in various dimensions.

The breakdown of *Proclaim* in Table 2 shows that *Endorse* stood out as the most frequently used (12.4%). *Concur* and *Pronounce* were much rarer, with only 0.5% and 0.3% occurrences, respectively. Both *Affirm* and *Concede* within the *Concur*-type make up only a humble share of 0.5% of the corpus. Examples (13) and (14) are extracted from the corpus, illustrating their uses.

(13) **It is not surprising** [affirm] that Pajares (1992) considers beliefs as a messy construct.

(14) **Undoubtedly,** the researches *[sic]* on TPACK, as aforementioned, have been increased in recent years, but [concede] very few studies that have been done in the world […].

In (13), the phrase *"it is not surprising"* subtly affirms the cited author's perspective, suggesting that the writer aligns with this established understanding. In contrast, the expression *"undoubtedly… but"* in (14) acknowledges the progress in TPACK research while shifting the focus to an evident research gap. The use of the adversative conjunction *but,* combined with the explicit limitation noted in the expression *"very few studies,"* positions the writer's research as addressing a critical and underexplored area. The use of *Pronounce* is even rarer, with only six instances identified:

(15) **There is no doubt** [pronounce] that teachers' beliefs play a pivotal role in […].

(16) **The results offer valuable proof** [endorse] that dialogic exchanges during group projects are essential […].

In (15), the writer uses *there is no doubt* to present an authoritative stance on the importance of teacher beliefs. This explicit emphasis eliminates room for alternative interpretations, a signal that the writer is confident in the assertion. Meanwhile, *Endorse* claims the highest proportion within *Proclaim* resources. To illustrate in (16), the phrase *the results offer valuable proof* explicitly aligns the writer with the cited findings. Such alignment with external evidence lends credibility to the writer's claims, which highlights the importance of dialogic exchanges in L2 acquisition.

The use of Expand types

Within *Expand*, the occurrences of *Attribute* overwhelmingly dominated with 58.4%, while *Entertain* makes up a much smaller proportion (1.6%), as displayed in Table 3. The high proportion of *Expand* types reflects the writers' strong preference for opening dialogic space and engaging with multiple perspectives in their academic writing.

Table 3. *The breakdown of Entertainment and Attributes*

Expand type		Occurrences	Percentage
Entertain		39	1.6%
Attribute	Acknowledge	1632	57.3%
	Distance	27	1.1%

Albeit very infrequent, *Entertain* was used by Vietnamese EFL students to primarily introduce propositions as subjective, tentative, or open to alternative interpretations.

(17) However, the findings from this paper **may be** [entertain] stereotypical.

In (17), the writer used *"may be"* to acknowledge potential weaknesses in the findings and ultimately to leave room for alternative interpretations or critiques. It also demonstrates how *Entertain* enables writers to maintain a cautious stance in the discussion or interpretation of the findings that may be subject to variability or critique. The low frequency of *Entertain* in this corpus may indicate a more cautious approach in engaging with tentative or speculative statements in comparison to the doctoral theses analyzed by Geng and Wharton (2016).

Regarding the use of *Attribute*, there was an asymmetry between *Acknowledge* (57.3%) and *Distance* (1.1%). This result implies students' strong preference for neutrally reporting external perspectives rather than expressing skepticism or doubt. *Acknowledge* as the leading *Attribute* type, including 57.3% occupied more than half of the total *Engagement* instances. It allows the writers to neutrally report the claims, arguments, or findings of other scholars without explicitly evaluating or questioning their validity.

(18) Newmark (1988) **further refined** [acknowledge] the concept of CST by differentiating […]. He **argued** [acknowledge] that […].

The use of *"further refined"* and *"argued"* reflects the writer's reliance on authoritative voices to ground their discussion without making evaluative judgments. This acknowledgment is instrumental for defining key concepts and establishing the context for the writer's research, implying agreements without critique. Obviously, the *Acknowledge* strategy is significant in neutrally incorporating external voices into students' academic conversation. Avoiding overt alignment or critique can be a strategy to ensure that their tone remains objective and scholarly, which is crucial in the early stages of argument development.

Distance, on the contrary, was rarely used, reflecting students' cautious approach to overtly critique or question established views. For example, the use of the verb "claim" in example (19) to introduce the author's conclusion signaled that the writer is reserved about its generalizability or applicability.

(19) The author **claimed** [distance] that professional development has a limited impact on […].

Discussion

How Vietnamese EFL students use Engagement resources in the evaluation of prior academic works

These consistent findings suggest that academic writers, whether novice or professional, commonly place a priority on *Heterogloss* engagement in their LR sections. In the examination of introduction sections of research articles, Alotaibi (2019) also noted a predominance in usage of *Heterogloss* in both SSCI-ranked and non-SSCI-ranked journals. This consistency may further show that the writers have their preference on dialogic *Engagement* across different sections of scholarly texts and publication contexts. However, Chinese EFL students were in favor of *Monogloss* (56.1%) over *Heterogloss* (43.9%) in their MA theses (Xie, 2016), which may reflect variations in academic writing practices where Chinese writers assert propositions with greater confidence, potentially at the expense of dialogic interaction. By comparison, Vietnamese EFL students in this study demonstrate a stronger inclination towards engaging with multiple perspectives. The difference might be due to the influence of pedagogical practices that emphasize dialogic *Engagement* and critical evaluation. However, more empirical evidence is required to reach a substantive conclusion as to the different EFL academic writing practices between Vietnamese and Chinese MA students.

Another notable difference is between our corpus and Geng and Wharton's (2016). The lower frequency of the *Engagement* in our data could be accounted for the different academic genres analyzed. While we focused on LR sections, Geng and Wharton (2016) investigated discussion sections of doctoral theses, which are evaluative and argumentative, and thus more intensive in the use of *Engagement*. The LR section is arguably more inclined towards presenting prior studies and establishing research contexts rather than engaging in detailed argumentation or critique. Another plausible explanation is that novice academic writers' texts often exhibit a deficiency in evaluative elements (Chang & Schleppegrell, 2011; Xie, 2016). In addition, the writer's proficiency also hinders the frequency of *Engagement* usage. The writers in Geng and Wharton's (2016) study are doctoral-level, whereas those in our study are MA students.

Regarding the use of *Heterogloss*, however, the Vietnamese MA students in the present study might have a less advanced academic writing level than those of doctoral candidates or professional researchers. Therefore, they may rely more on the *Expand* strategies to engage with prior works and create a less assertive stance, possibly due to limited confidence in claiming authority or less exposure to advanced writing practices. This is in line with Liu's (2013) argument that novice writers' texts are characterized by a limited use of intra-vocalizing (*contract*) resources.

Notably, there were contradictory patterns in studies with different groups of Asian writers. While Thai writers use a wide range of *Expand* type, (Amornrattanasirichok and Jaroongkhongdach, 2017), Chinese and Vietnamese writers preferred using *Contract* type in applied linguistics research articles and theses (Nguyen & Nguyen, 2024; Geng & Wharton, 2016). These disparities possibly stem from differences in the specific subfields within applied linguistics, as well as the sections of academic texts under investigation. Arguably, the complexity and variability of *Engagement* usage is more evident, even within the same discipline. Further intradisciplinary studies are therefore needed to explore these differences in greater detail.

Having a closer look at the *Contract* type, the combined use of *Deny* and *Counter* strategies in the corpus underlines their critical role in constructing academic arguments since highlighting research gaps or weaknesses in previous studies using these resources helps writers justify the relevance and necessity of their research. The relatively higher occurrence of *Counter* compared

to *Deny* indicates that Vietnamese EFL MA students tend to critique prior research by opposing or replacing ideas rather than outright rejecting or negating them. In this regard, their critiques possess a sense of elaboration with a balance between the identification of weaknesses and recognition of contributions or strengths in the cited studies. It also reflects their effort to establish their research focus without being overly dismissive, thus maintaining scholarly credibility and fostering dialogic *Engagement*. However, the overall occurrence of *Disclaim* in this study is significantly scarce, only 5.6% of all the *Engagement* in the entire corpus. Therefore, although the previous research findings showed interesting patterns in how *Disclaim* strategies were employed, the limited frequency of their use means that these conclusions are unnecessarily overarching, but rather indicative of tendencies within this specific dataset. Further research with larger corpora should be necessary to draw more substantive conclusions. Meanwhile, the usage patterns of *Endorse*, *Concur* and *Pronounce* indicate an ability to manage rhetorical *Engagement* in academic writing. The higher use of *Endorse* suggests that Vietnamese EFL MA students prioritize external validation to support their arguments, a rather prudent approach for establishing credibility in scholarly discourse. However, their highly occasional use of *Pronounce* and *Concur* reveals a rare willingness to explore more assertive and critical engagement strategies, albeit cautiously. This usage pattern points out the status quo of their academic writing, where they balance deference to established knowledge with an inactive authorial presence. Over time, as these writers gain confidence and experience, their use of rhetorical resources might evolve further. To gain a deeper understanding of how these practices may develop over time or respond to targeted pedagogical interventions, future research could benefit from longitudinal studies tracking the evolution of students' rhetorical strategies across their academic careers. Alternatively, experimental studies could explore the impact of specific instructional approaches on their use of rhetorical resources to shed light on how these strategies can be fostered and enhanced.

Considering the use of *Expand* resources, Vietnamese MA students employ them more extensively to engage with external perspectives and create a dialogic space for scholarly discourse. While *Attribute* was the primary tool for them to reference prior research, the rare use of *Entertain* was still a signal of their awareness of the need to present claims cautiously and leave room for alternative viewpoints. Within the *Attribute*, the high frequency of *Acknowledge* in this study resonates with Xie's (2016) results, suggesting that the MA students, possibly due to their novice status, rely on the strategy that situates their research within existing literature and demonstrate familiarity with authoritative sources. However, the sparing use of *Distance* hinted at students' hesitancy to engage critically or challenge established views. In general, the use of *Attribute* uncovers students' effort to express neutrality, signaling their caution in asserting an evaluative stance, likely due to their novice status and a lack of confidence in navigating complex academic arguments. Jalilifar et al. (2013) explained that their MA experience might be lacking in academic engagement training, which could result in limited familiarity with strategies for critical evaluation and dialogic interaction in academic writing.

Implications for Teaching Academic English

This study's results have underlined significant pedagogical implications. First, academic English instruction should place greater emphasis on cultivating critical thinking skills regarding how to question, critique, and propose alternatives to existing literature. The limited application of *Pronounce* and *Distance* indicates that students in this study did not critically engage in scholarly discourse. Inadequate teaching instruction has caused a lack of stance in students' writing (Chang & Schleppegrell, 2011). Managing the degree of personal voice is crucial for the formation of a

persuasive argument, which can be achieved through fostering interactive relationships with their readers, critically assessing their topics, and situating their perspectives within the context of alternative viewpoints (Hyland, 2005). In line with this perspective, activities such as debate-style assignments, peer reviews, and critical literature evaluations can be of great use.

Research on academic writing has also shown consistency in its emphasis on the *Engagement* in effective academic communication (Pham, 2024). The incorporation of explicit instruction on *Engagement* is paramount. Given the predominance of one category over the other (e.g., *Pronounce* versus *Endorse*, *Acknowledge* versus *Distance*) in this study, academic writing courses should explicitly familiarize students with the function and application of the *Engagement* to enable them to diversify their evaluation strategies in academic writing. Importantly, given that Vietnamese L1 writing is also rich in evaluative language, as suggested by Chung et al. (2022), instructors are encouraged to leverage positive L1 transfer by recognizing and integrating appropriate L1 rhetorical features into L2 academic texts.

Furthermore, the results of this study have proved the utility of corpus tools for analyzing linguistic patterns in academic writing. Therefore, educators should incorporate corpus-based instruction so that students can analyze examples of the *Engagement* in authentic academic texts, including their writing. Such activities may help students recognize the rhetorical impact of different *Engagement* strategies and fine-tune their use of these resources in other sections of their theses. To this end, tailored word banks, including phrases for *Appraisal*, as suggested by McKinley (2018), or concordance in Kwan's (2006) words, should be used to tap into linguistic representations of each category of the *Appraisal* framework, all of which is to strengthen students' ability to construct appropriate writer identities. In parallel with corpus-assisted instruction, Kara (2024) demonstrated that visual tools like infographics, when embedded into EFL writing classrooms, not only enhance students' organization and coherence in essay writing but also foster broader rhetorical and critical thinking skills. Our study suggests this teaching method since these visual tools can complement text-based instruction by helping students conceptualize complex argument structures, identify contrasting viewpoints, and represent authorial stance more effectively. Such corpus-assisted and multimodal approaches can thus offer a solid foundation for students' writing ability.

In addition, L2 motivation research also shows the significance of course relevance in shaping learner engagement. For example, Altalib (2019) found that students in English for Specific Purposes courses reported significantly stronger ideal L2 selves and more positive learning experiences than those in English for General Purposes courses. This study points to the potential of aligning academic writing instruction with students' disciplinary interests and long-term goals, which may foster not only motivational development but also deeper investment in critical engagement practices.

Last but not least, since LRs are among the most challenging sections for novice researchers, particularly those whose mother tongue is not English (Uzuner, 2008), we suggest scaffolding approaches in writing instruction. Evaluation should be first instructed, not during the writing of LRs, but rather during the reading process. Educators and facilitators of academic writing must recognize that training in evaluation cannot be an afterthought but a foundational step in a student's educational journey. Embedding evaluation skills into the reading process early on is crucial (Kwan et al., 2012) because it equips students to approach scholarly texts with a critical lens. It also enables them to construct well-informed and evaluative arguments in their academic

writing. This proactive approach is arguably essential for fostering not only effective academic writing but also deeper intellectual involvement in the literature.

Conclusion

This study investigated the use of *Engagement* resources in the Literature Review sections of Vietnamese EFL students' MA theses. The findings indicate that Vietnamese EFL students tend to engage dialogically in their LRs, gravitating toward *Heterogloss* resources rather than *Monogloss* resources. Among the former, *Expand* reigned dominant over *Contract*, suggesting a greater focus on opening dialogic space. Among the *Expand* resources, *Attribute* strategy was employed most, with *Acknowledge* being of the highest frequency. This result elucidates the students' reliance on neutrality when reporting external voices to situate their research within established literature. Besides, the scant use of *Deny*, *Counter*, *Concur*, *Pronounce*, *Entertain*, and *Distance* is indicative of either a hesitancy to exercise critical dialogic *Engagement*, an amount of caution, and a lack of confidence in employing these rhetorical strategies, or a potential unfamiliarity with their usage due to insufficient academic writing training.

Naturally, limitations are inevitable. First, the study's small sample size of 20 theses may not capture the whole spectrum of Vietnamese EFL students' academic writing practices. Additionally, focusing solely on the LR sections excludes other thesis components that may exhibit distinct *Engagement* patterns. The reliance on corpus-based analysis limits insights into the cognitive and instructional factors influencing students' rhetorical choices. Accordingly, future studies could expand the sample size and include theses from various disciplines and linguistic backgrounds to enhance generalizability and provide a more comprehensive view of students' writing practices. Nevertheless, this study is still significant as it elucidates how Vietnamese EFL students employ *Engagement* in LR sections of their MA theses and offers actionable pedagogical implications regarding the use of *Engagement* in particular and *Appraisal* in general in the EFL context that empower students to engage more actively with scholarly discourse.

Funding: This research is not funded.
Conflicts of Interest: There is no conflict of interest.
Authenticity: All the authors named in the article have contributed to the manuscript preparation. The authors used AI tool for polishing the language where the connection of different sections written by the authors needs a smooth flow.

About the authors

Dr. Lien-Huong Vo is a senior lecturer of Linguistics and Translation Studies at Hue University, Vietnam. Her research expertise pertains to linguistics and applied language studies. She is also interested in integrating linguistic theories into teaching the English language and literature in the EFL context. ORCID: https://orcid.org/0000-0001-7922-7416

Long Viet Le (MA) is a lecturer at Faculty of English, Hue University, Vietnam. His research interests include Linguistics and TESOL, with a recent focus on corpus linguistics and the emerging trend of Discursive News Values Analysis. ORCID: https://orcid.org/0009-0002-5527-4810

Mai-Duyen Pham (MA) is a lecturer at Can Tho University, Vietnam. She is currently pursuing the study leading to the doctoral degree in Linguistics at Hue University. Her research project is related to a linguistic comparative study of attitudinal language in cinematic discourse of Vietnamese-language and English-language movies. ORCID: https://orcid.org/0009-0005-4387-8424

References

Alotaibi, H. (2019). An exploration of authorial stance in SSCI-ranked journals versus non-SSCI-ranked Journals. *3L: The Southeast Asian Journal of English Language Studies*, *25*(3), 65–78. http://doi.org/10.17576/3L-2019-2503-05

Altalib, A. (2019). L2 motivation in ESP and EGP courses: An investigation of L2 motivational selves among learners of English in Saudi Arabia. *Australian Journal of Applied Linguistics*, *2*(1), 1–16. https://doi.org/10.29140/ajal.v2n1.113

Amornrattanasirichok, S., & Jaroongkhongdach, W. (2017). Engagement in literature reviews of Thai and international research articles in applied linguistics. *Proceedings of the International Conference: DRAL 3/19th ESEA 2017*, 312–327.

Chang, P., & Schleppegrell, M. (2011). Taking an effective authorial stance in academic writing: Making the linguistic resources explicit for L2 writers in the social sciences. *Journal of English for Academic Purposes*, *10*, 140–151. https://doi.org/10.1016/j.jeap.2011.05.005

Fryer, D. L. (2013). Exploring the dialogism of academic discourse: Heteroglossic Engagement in medical research articles. *English corpus linguistics: Variation in time, space and genre*, *77*, 183–207. https://doi.org/10.1163/9789401209403_011

Geng, Y., & Wharton, S. (2016). Evaluative language in discussion sections of doctoral theses: Similarities and differences between L1 Chinese and L1 English writers. *Journal of English for Academic Purposes*, *22*, 80–91. http://dx.doi.org/10.1016/j.jeap.2016.01.001

Gil-Salom, L., & Soler-Monreal, C. (2014). Writers' positioning in literature reviews in English and Spanish computing doctoral theses. *Journal of English for Academic Purposes*, *16*, 23–39. https://doi.org/10.1016/j.jeap.2014.08.002

Goudong, G. U. & Afzaal, M. (2023). Understanding political debates from the perspective of Appraisal theory of China's Belt and Road initiative: A corpus-based critical discourse analysis. *BRU ELT Journal*, *1*(1), 83–98. https://doi.org/10.14456/bej.2023.7

Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies*, *7*(2), 173–192. https://doi.org/10.1177/1461445605050365

Jalilifar, A. R., Hayati, A. M., & Mashhadi, A. (2013). Evaluative strategies in Iranian and international research article introductions: Assessment of academic writing. *Journal of Research in Applied Linguistics*, *3*(1), 81–109.

Kara, S. (2024). Enhancing students' writing competence and 21st-century skills simultaneously via infographics. *Australian Journal of Applied Linguistics*, *7*(3), 2044. https://doi.org/10.29140/ajal.v7n3.2044

Kwan, B. S. (2006). The schematic structure of literature reviews in doctoral theses of applied linguistics. *English for Specific Purposes*, *25*(1), 30–55. https://doi.org/10.1016/j.esp.2005.06.001

Kwan, B. S., Chan, H., & Lam, C. (2012). Evaluating prior scholarship in literature reviews of research articles: A comparative study of practices in two research paradigms. *English for Specific Purposes*, *31*(3), 188–201. https://doi.org/10.1016/j.esp.2012.02.003

Lee, J. J., & Deakin, L. (2016). Interactions in L1 and L2 undergraduate student writing: Interactional metadiscourse in successful and less-successful argumentative essays. *Journal of second language writing*, *33*, 21–34. https://doi.org/10.1016/j.jslw.2016.06.004

Lian, Y. (2018). Analysis of Xi's diplomatic speeches from the perspective of Appraisal theory. *Journal of Language Teaching and Research*, *9*(4), 759–764. http://dx.doi.org/10.17507/jltr.0904.12

Liu, X. (2013). Evaluation in Chinese university EFL students' English argumentative writing: An appraisal study. *Electronic Journal of Foreign Language Teaching*. *10*(1), 40–53.

Loghmani, Z., Ghonsooly, B., & Ghazanfari, M. (2020). Engagement in doctoral dissertation discussion sections written by English native speakers. *Journal of English for Academic Purposes*, *45*, 100851. https://doi.org/10.1016/j.jeap.2020.100851

Loi, C.-K., Lim, J. M.-H., & Wharton, S. (2016). Expressing an evaluative stance in English and Malay research article conclusions: International publications versus local publications. *Journal of English for Academic Purposes*, *21*, 1–16. http://dx.doi.org/10.1016/j.jeap.2015.08.004

Luo, D., Lin, Y. & Zhang, L. (2022). Positive discourse analysis of journalistic discourses about Sino-foreign relations from the Attitude perspective of Appraisal theory. *International Journal of Frontiers in Sociology*, *4*(2), 46–50. https://dx.doi.org/10.25236/IJFS.2022.040209

Martin, J. & White, P. (2005). *The language of evaluation*. Palgrave Macmillan London. https://doi.org/10.1057/9780230511910

Matthiessen, C. M., & Teruya, K. (2023). *Systemic functional linguistics: A complete guide*. Taylor & Francis. https://doi.org/10.4324/9781315675718

McKinley, J. (2018). Integrating appraisal theory with possible selves in understanding university EFL writing. *System, 78*, 27-37. https://doi.org/10.1016/j.system.2018.07.002

Mori, M. (2017). Using the Appraisal framework to analyze source use in essays: a case study of engagement and dialogism in two undergraduate students' writing. *Functional Linguistics*, *4*(1), 11. https://doi.org/10.1186/s40554-017-0046-4

Naghizadeh, M., & Afzali, K. (2018). Author engagement in the Literature Review of research articles published in international and Iranian local journals: Some insights from Appraisal theory. *Applied Research on English Language*, *7*(1), 111–142. https://doi.org/10.22108/are.2019.113133.1369

Ngo, T. T. B. (2013). *The deployment of the language of evaluation in English and Vietnamese spoken discourse* [Unpublished doctoral dissertation]. University of Sydney.

Nguyen, H. B., & Nguyen, P. T. (2024). An analysis of the Engagement system of Vietnamese research articles in the light of the Appraisal framework. *Theory and Practice in Language Studies*, *14*(5), 1595–1603. https://doi.org/10.17507/tpls.1405.33

Pham, T. P. (2024). Hooking the readers via the use of attitudinal language in economic research article abstracts. *3L: Language, Linguistics, Literature - The Southeast Asian Journal of English Language Studies*, *30*(4), 325–343. http://doi.org/10.17576/3L-2024-3004-22

Ross, A. S. & Caldwell, D. (2020). 'Going negative': An appraisal analysis of the rhetoric of Donald Trump on Twitter. *Language & Communication, 70*, 13–27. https://doi.org/10.1016/j.langcom.2019.09.003

Uzuner, S. (2008). Multilingual scholars' participation in core/global academic communities: A literature review. *Journal of English for Academic Purposes, 7*, 250–263. https://doi.org/10.1016/j.jeap.2008.10.007

White, P. R. R. (2003). Beyond modality and hedging: A dialogic view of the language of intersubjective stance. *Text*, *23*(2), 259–284. https://doi.org/10.1515/text.2003.011

Xie, J. (2016). Direct or indirect? Critical or uncritical? Evaluation in Chinese English-major MA thesis literature reviews. *Journal of English for Academic Purposes, 23*, 1–15. https://doi.org/10.1016/j.jeap.2016.05.001

Yuliyanti, A. Y. (2023). Appraisal system on Twitter: An attitudinal analysis toward alleged Islamic Blasphemy case of M Kece. *Ranah: Jurnal Kajian Bahasa, 12*(1), 221–235. http://dx.doi.org/10.26499/rnh.v12i1.5657

Zhang, W., & Cheung, Y. L. (2018). The construction of authorial voice in writing research articles: A corpus-based study from an APPRAISAL theory perspective. *International Journal of English Studies*, *18*(2), 53–75. https://doi.org/10.6018/ijes/2018/2/320261

AWEJ for Translation & Literary Studies
International Peer Reviewed Journal ISSN: 2550-1542

AWEJ for Translation & Literary Studies, Volume 9. Number 4. October 2025 Pp. 107 -122
DOI: http://dx.doi.org/10.24093/awejtls/vol9no4.9

Translation and L2 Arabic Teaching during Hajj: Linguistic Challenges in Multilingual Religious Contexts

Areej A. Naeem
Department of Arabic Language
Faculty of Languages and Translation
University of Jeddah
Jeddah, Kingdom of Saudi Arabia
Email: naareeg@uj.edu.sa

Received: 07/21/2025 Accepted: 09/21/2025 Published: 10/24/2025

Abstract

The present paper explores the pedagogical dimension of linguistic facilitation during the Hajj pilgrimage. It focuses on the potential integration of Arabic language instruction as a second language alongside formal translation services provided to non-Arabic-speaking pilgrims. Set against the backdrop of a highly multilingual and spiritually charged environment, the study investigates whether communicative Arabic instruction can effectively reduce linguistic barriers. It also examines how such instruction may support ritual comprehension and improve procedural fluency. The study is guided by a primary research question: to what extent does this integrated approach enhance linguistic accessibility? Its main aim is to examine the feasibility of embedding Arabic instruction within translation services in the Hajj context. Methodologically, the study adopts a descriptive-analytical method to examine the linguistic landscape of the Hajj and a developmental method to construct an instructional model tailored to the pilgrimage. The significance of the study lies in showing how the integration of Arabic instruction with translation services can enrich pilgrims' spiritual engagement, enhance their safety, and strengthen Saudi Arabia's broader vision for linguistic accessibility. The main findings indicate that integrating Arabic instruction with translation services enhances pilgrims' comprehension, ritual performance, and spiritual connection. The study concludes by recommending the institutionalization of a seasonal Arabic language immersion model, which could be extended to religious tourism, diplomacy, and Arabic for specific purposes.

Keywords: Arabic for specific purposes, communicative competence, Hajj, Krashen, religious tourism, second language acquisition, translation services

How to cite: Naeem, A. A. (2025). Translation and L2 Arabic Teaching during Hajj: Linguistic Challenges in Multilingual Religious Contexts. *Arab World English Journal for Translation & Literary Studies 9* (4): 107-122. http://dx.doi.org/10.24093/awejtls/vol9no4.9

Introduction

The Hajj is not merely an act of worship or a human gathering; it is a recurring historical and human phenomenon with global political, economic, and ethical repercussions (Al-Rashidi, 2016; Youssef, 2023). Beyond its spiritual core, Hajj represents a fusion of divine revelation with daily practices, encompassing financial preparation and social interaction. It invites scholarly engagement from multiple disciplines, employing theories from the social and human sciences to analyze crowd dynamics, and addressing emotional, cognitive, and economic dimensions (Siddiqi & Al-Ruwaithi, 2023). Language, as a vital bridge of communication among nations, has long been recognized by Saudi Arabia as no less critical than physical infrastructure, a perspective reinforced in the era of Vision 2030 (Al-Mahmoud, 2020; Saudi Arabia, 2025a).

Academic inquiry into Hajj has expanded significantly, with 2024 witnessing a surge in multidisciplinary studies spanning health (Mashat, 2024), communication intersecting with cultural anthropology (Al-Zahrani, 2024), social-cognitive constructs (Al-Harithi, 2023), spatial orientation (Mustafa, 2024), environmental factors (Shaaban, 2024), and artificial intelligence applications (Al-Sayigh, 2024; Qazzaz, 2024). While these studies enrich the understanding of Hajj services, there remains a critical gap regarding the linguistic dimension. Only one prior work (Qadhi et al., 2021) has addressed this aspect by compiling a multilingual dictionary for non-Arabic-speaking pilgrims. However, this approach is limited to lexical support. The unexplored area, and the one that this study specifically addresses, is the pedagogical potential of teaching Arabic as a second language to pilgrims alongside translation services.

The main aim of this study is to examine whether integrating Arabic instruction as a second language with formal translation services can enhance pilgrims' comprehension, ritual performance, and overall spiritual engagement.

The significance of the study lies in presenting an innovative approach: moving beyond translation-only solutions to incorporate pedagogical interventions. This integration is expected to enrich the Hajj experience, enhance safety, and align with Saudi Arabia's broader strategy for linguistic empowerment in religious tourism. This study seeks to:

1. Explore the feasibility of embedding Arabic language instruction within existing translation services.
2. Identify the linguistic barriers faced by non-Arabic-speaking pilgrims.
3. Evaluate Saudi Arabia's efforts in translation and multilingual accessibility.
4. Propose a sustainable, context-sensitive instructional model for Arabic language integration during Hajj.

To achieve these objectives, the study addresses the following questions:

1. To what extent can communicative Arabic instruction reduce linguistic barriers during Hajj?
2. How can integrating Arabic instruction alongside translation services enrich pilgrims' ritual and spiritual experiences?
3. What are the main linguistic challenges faced by non-Arabic-speaking pilgrims?
4. How effective are Saudi Arabia's current translation and language accessibility initiatives in addressing these challenges?

Literature Review

The Importance of Enhancing the Linguistic Dimension in the Pilgrimage Journey

The Hajj season is characterized by extraordinary linguistic and cultural plurality, bringing together pilgrims from over 185 nations, speaking fifty major languages and more than two hundred dialects (Saudi Arabia: General Authority for Statistics, 2023). This unprecedented diversity makes communication a central axis of the pilgrimage experience, not only for logistical efficiency but also for safety, well-being, and spiritual immersion. Without effective linguistic facilitation, even routine activitiessuch as navigating entry points, performing rituals in crowded sites, or accessing medical servicesmay become sources of confusion or danger. Saudi authorities have increasingly recognized this reality, positioning language infrastructure as a strategic pillar equal to transportation and security (Al-Mahmoud, 2020). Evidence supports this approach: Hadi et al. (2025) demonstrated that deploying an AI-based multilingual assistant in a Hajj hospital context significantly reduced CT examination times and minimized errors, highlighting the measurable impact of communication technologies on service quality. Such findings confirm that language mediation is not an auxiliary service but a functional necessity with direct implications for safety and efficiency. At the same time, the linguistic dimension of Hajj extends into the spiritual and affective domain. Pilgrims' ability to receive invocations, jurisprudential guidance, and religious sermons in their native languages fosters tranquility and devotion. Survey data from 2023 showed that over 90% of pilgrims who accessed real-time translation at guidance centers reported enhanced satisfaction and emotional comfort (Saudi Arabia, 2023c). The language medium, therefore, acts not merely as a conduit for information but as a bridge to deeper engagement with the sacred, transforming ritual acts (mansak) into symbols of meaning (shaʿīrah). Recent translation scholarship adds further depth to this understanding. Ghazala (2025) argues that cultural translatability can be achieved through sustainable strategies neutralization, adoption, reconciliation, or managed "clash" offering a framework for negotiating religiously sensitive content across languages. Similarly, Abdul Magid (2025) emphasizes that teaching translation theory equips practitioners to select appropriate strategies under real-world conditions, enhancing their cultural and communicative competence. These theoretical insights resonate strongly with the Hajj context, where linguistic mediation must balance fidelity to sacred texts with accessibility for diverse audiences. Moreover, the development of domain-specific linguistic resources, such as Hajj-FQA (Aleid & Azmi, 2025), indicates a growing recognition of the need for tailored tools that go beyond generic translation. By compiling a fatwa-based Q&A dataset, this work underscores how contextualized guidance can reduce confusion and enhance pilgrims' confidence in performing rituals. Together, these studies confirm that language is not a supplementary service in Hajj but a central mechanism through which safety, spirituality, and inclusivity are realized.

Gaps and Challenges in Current Linguistic Facilitation

Despite significant progress, formidable challenges remain in delivering comprehensive linguistic services during Hajj. One major issue is the shortage of qualified interpreters relative to the growing number of non-Arabic-speaking pilgrims (Abu-Ghararah, 2005). This shortage is compounded by high seasonal costs (Salem, 2025) and a lack of specialists in rare dialects such as Hausa, Amharic, and Kurmanji. Institutions often resort to booklets or digital aids (Amini et

al., 2021), yet these cannot substitute for real-time, context-sensitive interpretation. Technology, while promising, has its limitations. For instance, even with widespread 5G coverage, the density of over 1.8 million simultaneous users can overload networks, compromising real-time translation services (Bapty, 2024). Moreover, interpreters and language staff face exhausting working conditions, balancing long shifts with the need for absolute accuracy under pressure (Saudi Arabia, 2024a). These operational constraints undermine the effectiveness of existing solutions. Recent research confirms these challenges but also highlights potential pathways forward. Maddah et al. (2025), through qualitative interviews on wearable AI devices for disease detection at Hajj, revealed that stakeholders see strong value in such technologies but identified barriers including infrastructure gaps, privacy concerns, and training needs. Alquayt et al. (2025), in a large-scale task force report, similarly emphasized that AI innovations in healthcare spanning triage, imaging, and multilingual communication cannot succeed without clear metrics, ethical frameworks, and workforce preparation. Together, these findings underscore that technological fixes alone are insufficient without systemic capacity-building. Environmental and contextual pressures also exacerbate these gaps. High temperatures in Mecca and the risk of sudden emergencies demand rapid communication of new mobility or safety protocols (Madbuli et al., 2008). Yet producing and disseminating such messages in multiple languages both digital and printed remains an ongoing logistical challenge. Thus, the literature consistently points to a multidimensional problem: human-resource shortages, technological fragility, financial constraints, and environmental stressors collectively impede the provision of seamless linguistic facilitation.

Identified Research Gap

What remains largely absent in this expanding literature is the pedagogical perspective. While translation and AI-enabled mediation have received extensive scholarly attention, the structured teaching of Arabic as a second language (L2) to pilgrims has not. This gap is striking, given the potential benefits of even limited Arabic instruction: supporting ritual comprehension, strengthening safety communication, and fostering deeper spiritual connection. The current study addresses this gap by proposing a structured, context-sensitive Arabic teaching model that complements translation services. By doing so, it responds to the call for innovative, sustainable approaches that move beyond translation alone and position language learning as a strategic axis of Hajj service design.

Method

The study adopted two complementary research methodologies. The first was the descriptive-analytical method, employed to examine the linguistic-communicative phenomena associated with the experience of non-Arabic-speaking pilgrims during the Hajj season. In this phase, no direct field data were collected; instead, the analysis relied on institutional reports, governmental statistics, and documented surveys from 2023–2024, as well as academic literature on multilingual communication and Second Language Acquisition (SLA) in immersive contexts. This method enabled the researcher to systematically identify communicative barriers and challenges documented in prior studies and official sources. The second methodology was Design-Based Research (DBR), a flexible and iterative paradigm widely applied in educational

research and instructional model design. As articulated by Wang and Hannafin (2005), DBR enhances educational practice through successive cycles of analysis, design, development, and refinement. In this study, DBR guided the formulation of a pedagogical model for teaching Arabic as a second language during Hajj, ensuring that the model was theoretically grounded and context-sensitive.

Participants

The study did not engage direct human participants. Instead, the term "participants" is used conceptually to denote the target population non-Arabic-speaking adult pilgrims performing Hajj. Data for the descriptive-analytical phase were drawn from existing sources, including institutional reports, official Hajj statistics, published field surveys, and prior scholarly analyses of pilgrim experiences. This conceptualization reflects the study's pedagogical orientation: the proposed model is tailored to the needs of this population, even though empirical data were derived from documented sources rather than from direct interaction with pilgrims.

Research Instruments

The study belongs to the domain of qualitative research, and the instruments were selected accordingly. The tools employed included: (1) analysis of government and institutional reports (e.g., Ministry of Hajj and Umrah statistics, General Authority for Statistics); (2) review of published field surveys and observational studies conducted during the 2023 and 2024 Hajj seasons; (3) a systematic review of academic literature on SLA in short-term immersive and high-stakes contexts; and (4) thematic analysis of documented pilgrim testimonials and case studies available in published reports. These instruments provided a triangulated basis for identifying recurring linguistic barriers, gaps in current communicative practices, and opportunities for pedagogical intervention.

Procedures

The research unfolded in three structured phases. The first phase employed the descriptive-analytical method to assess the significance of linguistic facilitation during Hajj and to map existing challenges, based on the above-mentioned instruments. The second phase applied the DBR approach to design a pedagogical model aligned with the temporal and communicative demands of the pilgrimage. The final phase refined the model into two integrated components: a pre-travel digital Arabic module and an on-site communicative track. All analyses and design activities were conducted between February and May 2025.

Results

Saudi Arabia's Translation Efforts for Non-Arabic-Speaking Pilgrims

In response to the complex linguistic challenges of the Hajj season, Saudi Arabia has articulated a comprehensive linguistic strategy founded on three pillars: developing qualified human capital (Saudi Arabia, 2025a), integrating AI-powered technologies for real-time interpretation (Saudi Data and AI Authority, 2025), and continuously refining multilingual informational content (Saudi Arabia: General Commission for Audiovisual Media, 2024). This strategy reflects the Kingdom's view that linguistic infrastructure is as vital as physical

infrastructure. Central to this vision is the preparation of skilled personnel (Al-Qusaibi & Abed, 2018), the deployment of smart translation tools, and the systematic evaluation of all communicative materials to ensure clarity and accessibility.

Translating this strategy into practice, Saudi Arabia has advanced initiatives across these pillars. Human capital development has focused on training translators and language guides under national programs (Saudi Arabia, 2021a; Guests of Allah Service Program, 2023b). Technological integration includes mobile apps and smart devices with AI-based translation functions, notably the Nusuk platforms and the Nusuk smart wristband (Nusuk, 2025a; SDAIA, 2025). Simultaneously, multilingual media undergoes regular review, as seen in initiatives like Mediathon and the production of multilingual resources evaluated through empirical tools (Saudi Arabia, 2024b; Aljawhara Magazine, 2024).

Institutionally, Saudi Arabia has embedded language services within the Hajj system through the establishment of a specialized linguistic academy (King Salman Global Academy for Arabic Language & Saudi Arabia, 2023d), which produces multilingual glossaries, including *The Arab Tongue for Hajj* lexicon. This is complemented by platforms such as the Annual Scientific Forum for Hajj and Umrah Research and the "Researchers in the Service of the Guests of Allah" Award, which promote localized knowledge production and translation innovations (Saudi Arabia, 2025b; Umm Al-Qura University, 2020).

Furthermore, Saudi Arabia has institutionalized workforce training via the Holy Mosques Training Center and other collaborative programs aimed at enhancing linguistic and cultural competencies (Saudi Arabia, 2021b; Takamol Holding, 2021; Umm Al-Qura University, 2024). The deployment of multilingual female interpreters, AI translation tools like Tarjuman, and the expansion of real-time interpretation to over 50 languages reflect a layered approach to linguistic facilitation (Al Arabiya, 2023; Saudi Arabia, 2023c; SPA, 2019). These efforts are reinforced by digital innovations, such as interactive Qur'an teaching projects and AI-enhanced sermon translations (SPA, 2025; Manarat Al-Haramain, 2022).

Collectively, these initiatives underscore Saudi Arabia's sustained commitment to embedding linguistic accessibility within the Hajj experience, aligning technology, human capacity, and institutional frameworks to meet the evolving needs of pilgrims in multilingual religious settings.

However, the analysis also reveals a gap: while Saudi Arabia's initiatives are extensive, they remain primarily translation-oriented and technologically driven. There is limited focus on structured Arabic language acquisition for pilgrims, particularly through short-term, context-specific pedagogical programs. This unaddressed need serves as the foundation for the DBR model proposed in this study.

Proposed Linguistic Solutions to Enrich the Pilgrim Experience in a Multilingual Religious Context (Key Result of DBR Process)

Following the identification of this gap, this study proposes a complementary approach that draws from the unique spiritual and linguistic environment of the Holy Sites. Non-Arabic-speaking pilgrims, many of whom are devout Muslims, arrive in Saudi Arabia yearning to connect spiritually and linguistically with the language of revelation. They aspire to absorb fundamental Arabic terms and expressions that aid in performing rituals meaningfully. Many return home

spiritually transformed, carrying linguistic and cultural experiences that foster long-lasting relationships with the Two Holy Mosques and disseminate knowledge about Islam and its language.

The initiative *"Labbayka bi-jamīʿ al-lughāt, wa-al-ʿArabīyah tajmaʿunā"* ("Labbayk in All Languages, and Arabic Unites Us") presents a dual framework. The first component highlights the Kingdom's multilingual translation services and proposes integrating personalized linguistic assistance within the Nusuk platform, positioning interpreters as cultural mediators (Ruwad al-Aamal Magazine, 2024). The second component envisions a six-day Arabic program delivered by qualified instructors in pilgrims' camps, interweaving linguistic competencies with ritual engagement to deepen cognitive understanding and spiritual immersion.

The program's curriculum is designed for non-Arabic-speaking pilgrims, integrating essential linguistic competencies with the spiritual and procedural aspects of ritual observance. It offers a faith-anchored educational journey rooted in the Kingdom's unique sociolinguistic environment, where immersion in Arabic combines formal instruction with natural interaction with Saudi society whether through security personnel, service providers, or local residents. This exposure strengthens religious identity and deepens the connection to the land of revelation. The initiative instills fundamental Arabic expressions relevant to sacred sites, health, and safety, fostering cognitive understanding and spiritual immersion. A six-day structured course, delivered by trained instructors in low-pressure, ritual-adjacent settings, emphasizes implicit learning and contextual relevance.

Tables originally placed in this section have been relocated to the appendices for clarity. The full instructional framework is presented in Appendix A, Table 1 ("Proposed implementation framework"), while Appendix B, Table 2 provides an "Introductory guide to Arabic noun types."

Digital delivery via the Nusuk platform prior to arrival further facilitates linguistic transition and preparedness, leaving a lasting linguistic imprint as a spiritual keepsake from the Holy Land. The initiative is operationalized through curriculum development, dissemination via Nusuk, instructor training, institutional integration, deployment in the camps, and continuous evaluation.

The instructional design is grounded in Krashen's Second Language Acquisition Theory (1982), applying the Natural Order, Acquisition-Learning, Monitor, Comprehensible Input, and Affective Filter hypotheses. The program provides exposure through interactive, meaningful contexts, introducing grammar after patterns are internalized, supported by visual aids, gestures, and Total Physical Response activities in a psychologically safe environment.

Pedagogically, the program integrates direct, communicative, grammar-translation, and community-based methods, aligned with Richards and Rodgers (2014) and contemporary models such as the Silent Way, Suggestopedia, and Community Language Learning. It incorporates AI-supported personalization (Mukhlisin, 2017; Nasaruddin, 2024), and addresses the cultural and generational diversity of pilgrims (Al-Khouli, 2000). In this context, differentiated instruction may also be considered, as it is perceived as the most effective and preferred form of teaching (AlAmir, 2025). The conceptual framework rests on four pillars: the target learner group, limited temporal scope, communicative-functional objectives, and the unique religious context, guiding both macro-level design and micro-level lesson planning.

In alignment with linguistic empowerment objectives, instructors introduce participants to Arabic language institutes in Saudi universities and accredited digital platforms for continued learning upon their return. The initiative also creates opportunities to employ Arabic language graduates in seasonal Hajj roles, supporting long-term professional integration. This positions Arabic as a strategic bridge connecting pilgrims to the spiritual and intellectual essence of Islamic rites, echoing ʿUmar ibn al-Khaṭṭāb's advice: *"Learn Arabic, for it is part of your religion"* (Ibn Taymiyyah, 1995).

Ultimately, the initiative aspires to deliver not merely a functional service but a spiritually enriching, pedagogically transformative experience that enhances communication, deepens ritual engagement, and fosters intercultural dialogue. It aligns with Saudi Arabia's Vision to elevate the Hajj experience through holistic, human-centered services (Saudi Arabia, 2019a).

Discussion

This study set out to address four guiding research questions, and the discussion proceeds by responding to each in turn before synthesizing the broader theoretical and practical implications. First, regarding whether Arabic instruction could complement translation services during the Hajj (RQ1): the findings affirm that while Saudi initiatives have invested heavily in translation and AI-powered interpretation, a persistent gap remains in equipping pilgrims with even the most basic Arabic competencies. The proposed instructional model responds directly to this gap, emphasizing that translation and instruction function best as complementary rather than competing solutions. Second, concerning the role of language in enriching the pilgrim's experience (RQ2): analysis highlights that language is not merely a logistical tool but a spiritual and cognitive bridge. The model situates Arabic as both a means of ritual comprehension and a vector for deeper devotional engagement. Here, the fusion of communicative function and sacred meaning underscores the dual role of language in procedural accuracy and spiritual immersion. Third, on identifying the key linguistic barriers (RQ3): the study confirms that shortages in specialized human resources, the cost of professional interpreters, and the technical limitations of real-time translation remain major obstacles. These findings validate the need for a preventive pedagogical intervention that reduces reliance on external services by empowering pilgrims with essential Arabic expressions before and during Hajj. Fourth, regarding the effectiveness of current Saudi efforts and the potential for pedagogical intervention (RQ4): Saudi initiatives demonstrate impressive breadth training programs, AI-enabled platforms, and multilingual content but they remain predominantly service-oriented. The absence of systematic language instruction leaves a clear gap that the present study addresses by proposing an instructional model rooted in Design-Based Research (DBR). Beyond answering these questions, the discussion synthesizes theory with practice. The model's design draws explicitly on Krashen's Second Language Acquisition hypotheses. The Hajj environment provides an unparalleled opportunity for comprehensible input (pilgrims are surrounded by authentic, meaningful Arabic) and reduces the affective filter because spiritual motivation lowers anxiety, enhances receptivity, and fosters confidence in communication conditions difficult to replicate in a classroom. This validates Krashen's framework in a high-stakes, short-term learning context and shows how theory can be adapted to religious tourism.

At the same time, a strong discussion must recognize the study's limitations. This is currently a theoretical model yet to be empirically tested. Future implementation may face challenges such as:

- fatigue or limited uptake by pilgrims after long travels,
- the need for intensive training of seasonal instructors,
- coordination with Hajj authorities to integrate language modules into official schedules, and
- balancing the model's pedagogical ambitions with logistical constraints of mass gatherings.

Despite these challenges, the study demonstrates that linguistic empowerment through Arabic instruction is both feasible and strategically necessary, complementing Saudi Arabia's translation initiatives and enhancing the pilgrim's safety, comprehension, and spiritual experience.

Conclusion

The present study affirms that the Hajj pilgrimage represents one of the most powerful and natural forms of language immersion available to non-native speakers of Arabic. Surrounded by authentic Arabic expressions in religious, logistical, and interpersonal contexts, pilgrims find themselves in an environment that fosters meaningful and sustained language exposure. This immersion facilitates not only real-time language acquisition but also deepens pilgrims' engagement with the religious and cultural dimensions of the experience. Despite these advantages, non-Arabic-speaking pilgrims often face significant communicative challenges that may impede their ability to access essential services, comprehend instructions, and fulfill their ritual obligations effectively.

It is within this complex multilingual setting that the present study seeks to address the linguistic barriers that, if left unaddressed, may exacerbate other challenges whether medical, technical, or logistical. The study proposes an applied instructional model anchored in an integrative educational vision: a six-day Arabic language program specifically designed for non-Arabic speakers and aligned with the temporal and functional demands of Hajj. This program is preceded by a digital preparatory phase delivered via the Nusuk platform as part of the registration or orientation process, with the primary aim of priming and motivating pilgrims for active participation in the subsequent field-based, interactive instruction. Conceived as a complementary pillar to the comprehensive translation services already provided by Saudi Arabia, this model aspires to enhance pilgrims' linguistic competence, thereby fostering deeper spiritual engagement, procedural fluency, and overall well-being throughout their sacred journey.

Future research should move beyond conceptualization and empirically test the proposed instructional model under real-world Hajj conditions. A mixed-methods design is recommended: quantitative surveys to measure linguistic gains, comprehension accuracy, and satisfaction rates among a pilot group of non-Arabic-speaking pilgrims; and qualitative interviews or focus groups to capture experiential insights from both learners and instructors. Additionally, quasi-experimental comparisons between participants who receive the program and those who rely solely on translation services could provide robust evidence of the model's effectiveness. Such empirical testing would not only validate the theoretical underpinnings of the model but also

guide its refinement and scalability for broader contexts. From this standpoint, the study advocates capitalizing on this seasonal linguistic immersion as a strategic entry point for promoting subsequent Arabic language programs that are rapid, effective, and communicative in nature programs that may serve wider purposes including commerce, tourism, and other functional domains within the broader socio-cultural and economic landscape of the Arab region.

Funding: This research is not funded.
Acknowledgments: Not applicable.
Conflicts of Interest: I declare that there is no conflict of interest.
Authenticity: This manuscript is an original work
Declaration of Generative AI and AI-assisted Technologies : During the preparation of this work, the author used AI-assisted technologies (e.g., ChatGPT) for language editing and to improve the readability of the manuscript. After using these tools, the author carefully reviewed and revised the content as needed and takes full responsibility for the content of the publication.

About the author:

Dr. Areej Naeem is an Associate Professor of Translation Studies and Linguistics at the University of Jeddah, KSA. She has published several research papers, and her research interests include Arabic-English translation, translation pedagogy, translation technology, discourse analysis, language acquisition, and second language teaching.
ORCID ID: https://orcid.org/0000-0002-6732-4144

References

Abu-Ghararah, B. A. (2017). The gap between translator training and the translation industry in Saudi Arabia. *Arab World English Journal for Translation & Literary Studies, 1*(4), 107–118. https://doi.org/10.24093/awejtls/vol1no4.8

AlAmir, B. A. (2025). "Teach it, and then set me free": Saudi EFL learners' voice and choice in grammar pedagogies. *Arab World English Journal, 16*(2), 96–115. https://doi.org/10.24093/awej/vol16no2.6

Al-Harithi, G. T. S. (2023). Smart management for knowledge of pilgrims. In *The 22nd Scientific Forum for Hajj, Umrah, and Visit Research* (p. 35). The Custodian of the Two Holy Mosques Institute for Hajj & Umrah Research, Umm Al-Qura University.

Aljawhara Magazine. (2024). Kingdom's efforts during Hajj 2024: Applications, initiatives, and distribution of guidebooks to pilgrims. *Aljawhara Magazine*, *1*(74), 10–13.

Al-Khouli, M. (2000). *Methods of teaching the Arabic language*. Dar Al-Falah for Publishing and Distribution.

Al-Mahmoud, M. (2020). Language planning and language policy: A theoretical foundation. *Journal of King Abdulaziz University: Arts and Humanities*, *28*(13), 193–230.

Al-Owdh, A. (2019). Glorifying the rituals of Hajj: A cognitive-communicative analysis. *Arab Journal of Sciences and Research Publishing*, *5*(4), 155.

Al-Qahtani, R. (2023). A sustainable scenario for enhancing operational efficiency and preparedness for mass gatherings in the Holy Sites during Hajj and Umrah seasons in Mecca. *Educational Journal, Sohag University*, (105), 449–495.

Al-Qusaibi, M., & Abed, S. (2018). A strategic plan to establish a training center for Hajj and Umrah system workers. *Journal of King Abdulaziz University: Engineering Sciences, 29*(1), 61–81.

Al-Rashidi, H. (2016). The role of Hajj in enriching political and economic relations between Mecca and the Levant during the Ottoman era through the rule of Al-Azm. *Journal of Arabic and Human Sciences, Qassim University, 9*(3), 1417–1469. https://hajj.uqu.edu.sa/bank/6/serch.php?i

Al-Sayigh, A. A. A. R. (2024/1445 AH). Analyzing pilgrim satisfaction from social media platforms using artificial intelligence. *The Custodian of the Two Holy Mosques Institute for Hajj and Umrah Research, Umm Al-Qura University.* https://hajj.uqu.edu.sa/bank/6/serch.php?i

Al-Zahrani, A. D. (2024). The effectiveness of new media in supporting pilgrims' health promotion (Research Report No. 44521001). *The Custodian of the Two Holy Mosques Institute for Hajj and Umrah Research, Umm Al-Qura University.* https://hajj.uqu.edu.sa/bank/6/serch.php?i

Amini, Z., Mohammadamini, M., Hosseini, H., Mansouri, M., & Jaff, D. (2021). Central Kurdish machine translation: First large-scale parallel corpus and experiment. *arXiv preprint*. https://arxiv.org/abs/2107.14131

Custodian of the Two Holy Mosques Institute for Hajj and Umrah Research, Umm Al-Qura University. (2024). *Scientific record of the 24th Scientific Forum for Hajj, Umrah, and Ziyara Research (Excellence in transportation and crowd management).*

Custodian of the Two Holy Mosques Institute for Hajj and Umrah Research, Umm Al-Qura University. (2023). *Scientific record of the 22nd Forum for Hajj, Umrah, and Ziyara Research* (p. 35).

Custodian of the Two Holy Mosques Institute for Hajj and Umrah Research, Umm Al-Qura University. (2020). *Proceedings of the 20th Scientific Forum for Hajj, Umrah, and Visit Research.*

General Presidency for the Affairs of the Grand Mosque and the Prophet's Mosque. (2025). *Homepage*. https://alharamain.gov.sa/public/?page=page_795020

General Presidency for the Affairs of the Two Holy Mosques. (2022). *Simultaneous interpretation of sermons at the Grand Mosque* [Video]. YouTube. https://www.youtube.com/watch?v=mK86XILlflo

Ibn Hisham Al-Ansari. (2007). *Qatr al-Nada wa Ball al-Sada* (M. M. Abdul Hamid, Ed.). Cairo: Dar Al-Hadith.

Ibn Taymiyyah, A. (1995). *Iqtida' al-Sirat al-Mustaqim li-Mukhalafat Ashab al-Jahim* (Vol. 2, N. A. Al-Aql, Ed.). Riyadh: Dar Alam Al-Fawa'id.

King Salman Global Academy for Arabic Language, & Ministry of Hajj and Umrah. (2023). *Pilgrims' and Umrah performers' terminology dictionary*. https://siwar.ksaa.gov.sa/public-dict-information/73ce2496-7b2a-4bb0-a8c0-84ea5b1bd677

King Salman Global Academy for Arabic Language, Center for Arabization and Translation. (2025). *The linguistic reality of Arabic in Saudi Arabia and the GCC: Language planning* (pp. 33–123). KSAA.

King Salman Global Academy for Arabic Language. (2025). *Language planning policy*. KSAA. https://ksaa.gov.sa/ksaa.gov.sa/areas-of-work-and-initiatives/language-planning-and-policy

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.

Madbuli, N., Noweir, M. H., Bafail, A. O., Ibrahim, M. G., & Jomoah, I. M. (2008). Study of heat exposure during Hajj (pilgrimage). *Environmental Monitoring and Assessment, 147*(1–3), 279–295.

Manarat Al-Haramain. (2022). Live streaming – English [Webpage]. https://manaratalharamain.gov.sa/speech/235274240/البث-المباشر-(English)

Mansour, R. (2021). The role of language immersion in teaching Arabic to non-native speakers. *Journal of the Faculty of Arabic Language in Girga, Al-Azhar University, 25*(7).

Mashat, B. H. H. (2024). *Enhancing the internal health environment visited by pilgrims to improve the pilgrimage experience of the Guests of Allah at the sacred sites* (Research Report No. 44533001). The Custodian of the Two Holy Mosques Institute for Hajj and Umrah Research, Umm Al-Qura University. https://hajj.uqu.edu.sa/bank/6/serch.php?i

Ibn Hisham Al-Ansari. (2007). *Qatr al-nada wa-ball al-sada* (M. M. Abdul Hamid, Ed.). Cairo: Dar Al-Hadith.

Ibn Taymiyyah, A. (1995). *Iqtida' al-sirat al-mustaqim li-mukhalafat ashab al-jahim* (Vol. 2, N. A. Al-Aql, Ed.). Riyadh: Dar Alam Al-Fawa'id.

King Salman Global Academy for Arabic Language, & Ministry of Hajj and Umrah. (2023). *Pilgrims' and Umrah performers' terminology dictionary*. https://siwar.ksaa.gov.sa/public-dict-information/73ce2496-7b2a-4bb0-a8c0-84ea5b1bd677

King Salman Global Academy for Arabic Language, Center for Arabization and Translation. (2025). *The linguistic reality of Arabic in Saudi Arabia and the GCC: Language planning* (pp. 33–123). KSAA.

King Salman Global Academy for Arabic Language. (2025). *Language planning policy*. KSAA. https://ksaa.gov.sa/ksaa.gov.sa/areas-of-work-and-initiatives/language-planning-and-policy

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.

Madbuli, N., Noweir, M. H., Bafail, A. O., Ibrahim, M. G., & Jomoah, I. M. (2008). Study of heat exposure during Hajj (pilgrimage). *Environmental Monitoring and Assessment, 147*(1–3), 279–295. https://doi.org/10.1007/s10661-007-0100-x

Manarat Al-Haramain. (2022). *Live streaming – English* [Webpage]. https://manaratalharamain.gov.sa/speech/235274240/البث-المباشر-(English)

Mansour, R. (2021). The role of language immersion in teaching Arabic to non-native speakers. *Journal of the Faculty of Arabic Language in Girga, Al-Azhar University, 25*(7).

Mashat, B. H. H. (2024). *Enhancing the internal health environment visited by pilgrims to improve the pilgrimage experience of the Guests of Allah at the sacred sites* (Research Report No. 44533001). The Custodian of the Two Holy Mosques Institute for Hajj and Umrah Research, Umm Al-Qura University. https://hajj.uqu.edu.sa/bank/6/serch.php?i

Muhlisin, M. (2015). Methods of teaching Arabic to non-native speakers. *Alsinatuna, 1*(1), 1–15. https://doi.org/10.28918/alsinatuna.v1i1.637

Mustafa, B. H. H. (2024). *Enhancing the internal health environment visited by pilgrims to improve the pilgrimage experience of the Guests of Allah at the sacred sites* (Research Report No. 44533001). The Custodian of the Two Holy Mosques Institute for Hajj and Umrah Research, Umm Al-Qura University. https://hajj.uqu.edu.sa/bank/6/serch.php?i

Nasaruddin, N. (2024). Using ChatGPT in teaching Arabic as a foreign language. *Arabiyatuna: Journal Bahasa Arab, 8*(1), 1–24. https://doi.org/10.29240/jba.v8i1.9413

Nusuk. (2025). *Nusuk for Hajj*. https://hajj.nusuk.sa

Nusuk. (2025). *Nusuk*. https://www.nusuk.sa/ar

Qadi, A., Al-Tayeb, A., & Diop, A. (2021). *Preparing a multilingual abridged dictionary to assist non-Arabic-speaking pilgrims and visitors of the Two Holy Mosques in performing their religious duties.* The Custodian of the Two Holy Mosques Institute for Hajj and Umrah Research, Umm Al-Qura University. https://hajj.uqu.edu.sa/bank/6/serch.php?i

Qazzaz, O. B. O. (2024). *Applications of artificial intelligence in media and their role in enhancing awareness patterns among visitors to the Prophet's Mosque.* The Custodian of the Two Holy Mosques Institute for Hajj and Umrah Research, Umm Al-Qura University. https://hajj.uqu.edu.sa/bank/6/serch.php?i

Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching* (3rd ed.). Cambridge University Press.

Ruwad al-Aamal Magazine. (2023). Business ideas suitable for Hajj season: An opportunity to boost profits. *Entrepreneur Magazine, 1*(169), 58–61.

Ruwad al-Aamal Magazine. (2024). Hajj 2024: A technological revolution shaping the pilgrimage experience and opening investment opportunities for entrepreneurs. *Entrepreneur Magazine,* (177), 24–29.

Salem, R. (2025). Translation prices in Saudi Arabia. *Fast Trans.* https://fast4trans.com/أسعار-الترجمة-في-السعودية/

Saudi Arabia. (2019a). *Guests of Allah Service Program: About the program.* https://pep.gov.sa/ar/about

Saudi Arabia. (2019b). *Ministry of Health: Innovation (Health) provides instant translation for pilgrims as part of Smart Hajj.* https://www.moh.gov.sa/Ministry/MediaCenter/News/Pages/news-2019-09-08-024.aspx

Saudi Arabia. (2020). *Ministry of Communications and Information Technology: Annual report 2020* (p. 90). Riyadh.

Saudi Arabia. (2021a). *Ministry of Human Resources and Social Development: Cooperation memorandum for training Hajj season workers.* https://www.hrsd.gov.sa/media-center/news/843260

Saudi Arabia. (2021b). *Ministry of Human Resources and Social Development: News center.* https://www.hrsd.gov.sa/media-center/news/843260

Saudi Arabia. (2022a). *Human capability development program: Media document* (pp. 220–225).

Saudi Arabia. (2022b). *Ministry of Islamic Affairs, Dawah and Guidance: A mobile robot provides fatwas to pilgrims at Namira Mosque in Arafat.* https://www.moia.gov.sa/MediaCenter/News/Pages/09121443_7.aspx

Saudi Arabia. (2023a). *Guests of Allah Service Program: Annual report.* Vision 2030. https://www.vision2030.gov.sa/media/0hwbcwev/pep-annual-report-2023-english.pdf

Saudi Arabia. (2023b). *Guests of Allah Service Program: The Languages and Translation Agency at the Two Holy Mosques provides services in over 50 global languages.* https://pep.gov.sa/ar/news/487

Saudi Arabia. (2023c). *Ministry of Islamic Affairs, Dawah and Guidance: Awareness programs for pilgrims 1445H* (p. 15).

Saudi Arabia. (2023d). *Ministry of Hajj and Umrah: Annual report on the Hajj season 1445H* (p. 47).

Saudi Arabia. (2024a). *Ministry of Islamic Affairs, Dawah and Guidance: Translators in 19 languages at the Custodian of the Two Holy Mosques Guests Camps to facilitate Hajj rituals.* https://www.moia.gov.sa/MediaCenter/News/Pages/12121445_8.aspx

Saudi Arabia. (2024b). *Ministry of Media: Hajj and Umrah Mediathon.* https://mediathon.media.gov.sa/Home/Index

Saudi Arabia. (2025a). *Human capability development program media document: Language policy program.* King Salman Global Academy for Arabic Language (KSAA).

Saudi Arabia. (2025b). *Ministry of Hajj and Umrah: Research excellence award in the Hajj system.* Ministry of Hajj and Umrah. https://www.haj.gov.sa/Researchers

Saudi Arabia: General Authority for Statistics. (2024). *Annual Hajj statistics 1445H* (p. 1). General Authority for Statistics.

Saudi Arabia: General Commission for Audiovisual Media. (2024). *News.* https://gmedia.gov.sa/en/news

Saudi Data and Artificial Intelligence Authority (SDAIA). (2021). *Signing a memorandum of understanding between SDAIA, STC, and the Guests of Allah Service Program.* https://sdaia.gov.sa/ar/MediaCenter/Events/Pages/EventsDetails.aspx?EventID=7

Saudi Press Agency (SPA). (2019). *Hajj: Tarjuman application helped pilgrims translate signs in Mecca and the holy sites.* https://www.spa.gov.sa/1958311

Saudi Press Agency (SPA). (2024). *Global languages to broadcast the Arafat sermon and scholarly lessons from the Grand Mosque to the world.* https://spa.gov.sa/N2122643

Saudi Press Agency (SPA). (2025). *Hajj/Translation and guidance in multiple languages in Medina to ensure pilgrims' comfort.* https://www.spa.gov.sa/N2325905

Saudi Press Agency (SPA). (2025). *The General Presidency for the Affairs of the Two Holy Mosques launches real-time translation project for Arafat sermon in 34 languages.* https://www.spa.gov.sa/N2331689

Saudi Press Agency (SPA). (2025). *The Religious Affairs Presidency launched an educational e-project for Surah Al-Fatihah.* https://www.spa.gov.sa/N2286556

Shaaban, F. F. A. (2024). *Mitigating geo-environmental risks at Mount Uhud and the Archers' Mountain in Medina by assessing slope stability factors and using geographic information systems* (Research Report No. 44533007). The Custodian of the Two Holy Mosques Institute for Hajj and Umrah Research, Umm Al-Qura University. https://hajj.uqu.edu.sa/bank/6/serch.php?i

Siddiqi, Z., & Al-Ruwaithi, A. (2023). *Developing a guest experience culture in Hajj, Umrah, and Ziyara activities through enhancing internal motivation among service providers. Scientific Record of the 22nd Scientific Forum for Hajj*, 63. https://drive.uqu.edu.sa/_/hajj/files/multaqa/%D8%A7%D9%84%D8%B3%D8%AC%D

9%84%20%D8%A7%D9%84%D8%B9%D9%84%D9%85%D9%8A%2021%20%D8%A7%D9%86%D9%82%D9%84%D8%B4.pdf

Swain, M., & Johnson, R. K. (1997). *Immersion education: International perspectives.* Cambridge University Press.

Takamol Holding. (2021). *Specialty section.* Takamol. https://takamolholding.com/ar#speciality-section

Technical and Vocational Training Corporation. (2023). *News center.* TVTC. https://tvtc.gov.sa/ar/MediaCenter/News/Pages/hmc-s.aspx

Umm Al-Qura University. (2024). *Rafid Al-Haramain training program for Ministry of Hajj and Umrah staff concluded.* https://uqu.edu.sa/App/News/166069

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development, 53*(4), 5–23. https://doi.org/10.1007/BF02504682

World Tourism Organization. (2024). *International tourism recovers to pre-pandemic levels in 2024.* https://www.unwto.org/news/international-tourism-recovers-pre-pandemic-levels-in-2024

Youssef, M. (2023). Religious economy and concept investment: The Malaysian Hajj Fund (Tabung Haji). *Civilization Center for Studies and Research, Issues and Perspectives, 32.*

Abdul Magid, I. H. (2025). The significance of instructing translation theory to students majoring in translation. *Arab World English Journal for Translation & Literary Studies, 9*(2), 97–105. https://doi.org/10.24093/awejtls/vol9no2.6

Aleid, H. A., & Azmi, A. M. (2025). Hajj-FQA: A benchmark Arabic dataset for developing question-answering systems on Hajj fatwas. *Journal of King Saud University – Computer and Information Sciences, 37,* 135. https://doi.org/10.1007/s44443-025-00128-w

Alquayt, A., Aljuhani, O., Alharthi, A. F., … & England, A. (2025). AI-driven healthcare innovations for enhancing clinical services during mass gatherings (Hajj): Task force insights and future directions. *BMC Health Services Research, 25,* 876. https://doi.org/10.1186/s12913-025-13045-5

Ghazala, H. S. (2025). Rethinking the translatability of culture: A sustainable approach. *Arab World English Journal for Translation & Literary Studies, 9*(1), 2–16. https://doi.org/10.24093/awejtls/vol9no1.1

Hadi, Y. H., Altalhi, F. K., Ali, H. M., Shabli, M. A., Abu Aqil, A. I., & England, A. (2025). Enhancing CT examination efficiency with ChatGPT-4o for multilingual Hajj pilgrims: A short communication. *Journal of Medical Imaging and Radiation Sciences, 56*(1), 101781. https://doi.org/10.1016/j.jmir.2024.101781

Maddah, N., Verma, A., & Ainsworth, J. (2025). Opportunities and challenges surrounding the use of wearable sensor bracelets for infectious disease detection during Hajj: Qualitative interview study. *JMIR Formative Research, 9,* e60484. https://doi.org/10.2196/60484

Al Arabiya. (2023, April 8). Female interpreters serve women pilgrims 24/7 in 14 languages at the Grand Mosque. https://www.alarabiya.net/saudi-today/2023/04/08/-مترجمات-الحرم-المكي-يخدمن-المعتمرات-على-مدار-الساعة-ب14-لغة

Bapty, R. (2024, June 15). Tracking mobile network performance daily during the Hajj pilgrimage. *OpenSignal*. https://www.opensignal.com/2024/06/15/tracking-mobile-network-performance-daily-during-the-hajj-pilgrimage

Massoud, A. (2025, June 10). Robots and AI fatwas during Hajj 2025. *CNN Business Arabic.* https://cnnbusinessarabic.com/economic-stories/1119243

Asharq Al-Awsat. (2025, June 15). Pilgrims' prayers embrace Arafat's sky… Mina awaits their return to complete the rituals. https://aawsat.com/العالم-العربي/الخليج/5151376-دعوات-الحجيج-تعانق-سماء-عرفات-ومِنى-تترقب-عودتهم-لاستكمال-المناسك

Appendices

Appendices (A)

Table 1. *Proposed implementation framework*

Instructional Unit Topic	Contextual Setting	Gregorian Date	Suggested Activities	Targeted Skills
Greetings and Self-Introduction	Airport	2026-05-05	Flashcards, repetition, simple dialogue enactments	Listening – Speaking
Dialogue in Reception and Transportation	Hotel – Bus	2026-05-06	Role-playing, gestures	Listening – Speaking – Reading
Sacred Sites and Ritual Spaces	Holy Mosque and its Facilities	2026-05-07	Simulated video, map of the Haram, recitation of supplications	Listening – Pronunciation – Vocabulary
Interactions in the Pilgrim Camps	Mina and Arafat	2026-05-08	Functional phrases, flashcards, role-based scenarios	Conversation – Vocabulary
Emergency Situations	Hospital – Telephone	2026-05-09	Picture-based dialogue, emotional expression activities	Conversation – Reading
Recap and Introduction to Arabic Learning Centers in Saudi Arabia	Pilgrim Camps	2026-05-10	Interactive review, assessment stations, reflective exercises	Conversation – Reading

Appendices (B)

Table 2. Introductory guide to Arabic noun types

Noun Type	Definition	Examples
Proper Noun	A word indicating a specific person, place, or entity	Fatimah, Saudi Arabia, Badr
Demonstrative	A word used to point to a specific object or idea	this, that (hādhā, hādhihi)
Pronouns	Words indicating speaker, addressee, or third person	I, you, he, she (ana, anta, huwa)

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp.185-195
DOI: https://dx.doi.org/10.24093/awej/vol16no4.10

Unlocking the Potential of AI-Enhanced Mixed Reality: Enhancing Technical English Proficiency for Engineering Students in the Gulf Region

Shafiq ur Rehman
Foundation Program Unit, University of Doha for Science and Technology, Qatar.
Corresponding Author: shafiqur.rehman@udst.edu.qa

Rabia Khatoon
English Language Institute, King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia
Email: rkkhan@kau.edu.sa

Intakhab Alam Khan
King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia
Email: ikhan1@kau.edu.sa

Received: 08/29/2025 Accepted: 11/5/2025 Published: 12/10/2025

Abstract
This study examines the influence of artificial intelligence (AI) augmented mixed reality (MR) technologies on the enhancement of technical English proficiency among engineering students in the Gulf countries. The main aim of the research is to explore how the gap between foreign language learning and domain-specific technical knowledge can be reduced by integrating the immersive and adaptive MR environments with AI. The mixed method design was used to evaluate the effectiveness of AI-enhanced MR technologies in a 12-week course. The sample size consists of 300 students from the Gulf region. The treatment group used AI-enhanced MR technologies, which simulated real-life engineering scenarios to facilitate content-based English learning. The quantitative data from the pre-and post-tests revealed higher proficiency with the mean score of 72.5 to 80.3 that shows a medium-to-large statistically significant improvement. The qualitative data indicated enhanced engagement, confidence, and understanding of technical vocabulary. The study's results suggest that AI-driven MR applications provide an practical pedagogical approach for advancing English for Specific Purposes (ESP), particularly within a Science, Technology, Engineering, and Mathematics (STEM) setting. Technical issues, usability problems, and misalignment with culture were highlighted as challenges. Improving the technical support and cultural appropriateness of the content would address these challenges. The study's significance is shown through its contribution to practical pedagogical strategies by combining AI and MR for ESP instruction in the engineering context. It also reflects that these technologies can be effective in other fields.

***Keywords*:** Artificial intelligence, engineering, English for specific purposes, English proficiency, Gulf countries, human-computer interaction, immersive learning environments, mixed reality
How to cite: Rehman, S.U., Khatoon, R., & Khan, I.A. (2025). Unlocking the Potential of AI-Enhanced Mixed Reality: Enhancing Technical English Proficiency for Engineering Students in the Gulf Region. *Arab World English Journal, 16*(4): 185-195.
https://dx.doi.org/10.24093/awej/vol16no4.10

Introduction

Teaching English, primarily in an English for Specific Purposes (ESP) context, has become easier than ever because of the integration of artificial intelligence (AI)-enhanced and mixed reality (MR) technologies, which are known for their innovative and interactive nature. It helps students in the Gulf countries to deal with their two major issues of improving their English proficiency and comprehending the complex disciplinary content by providing integrated solutions that combine immersive, adaptive, and interactive learning experiences. AI-driven tools personalize and tailor content delivery, adapting to individual learners' needs, enhancing comprehension, retention, and engagement. For instance, AI can evaluate dynamic ways of understanding, and it changes instructional materials to meet the needs, applicable mainly in technical areas in which students need to understand more of the terminology and concepts (Ai, 2024).

MR, based on augmented reality (AR) and virtual reality (VR), provides an immersive environment that helps enhance learning results. MR allows for the simulation of engineering scenarios, preparing students to practice and apply technical English within an authentic context. This dual functionality supports both language acquisition and subject-specific understanding, aligning with the cognitive demands of engineering education (Ai, 2024). This has been noticed from multiple studies that AI-enhanced MR technologies support vocabulary acquisition, particularly when used in ESP settings (Divekar et al., 2021). However, the use of AI and MR is under-researched in the Gulf region for teaching English to engineering students. The available studies mainly discuss the general English proficiency and technological adoption, but present little findings on the contextual need of technical communication. Therefore, the current study aims to answer the following research questions:

- What is the effectiveness of AI-enhanced MR technologies in enhancing technical English proficiency in engineering students in the Gulf region?
- What are the specific challenges and benefits of using AI-enhanced MR for teaching technical English in engineering contexts?
- To what extent do AI-enhanced MR technologies influence students' human-technology interaction skills to understand and apply technical English in practical scenarios?

Literature Review

The phenomenon of AI and MR has received increasing attention in educational research, particularly in language teaching acquisition and technical education. This literature review provides an overview of current research on the effect of AI-enabled MR technology on the technical English proficiency of engineering students in the Gulf Region (Chen et al., 2021; Strielkowski et al., 2023; Yannier et al., 2024).

Several studies have demonstrated how AI-MR integration supports second- or foreign-language learners and enhances retention, motivation, and engagement. For example, AlSaleem (2023) and Khodabandeh et al. (2025) show that augmented reality (AR) and AI-enhanced MR can significantly enhance vocabulary teaching and learning in the ESP contexts. Further research shows that AI-enhanced MR applications are excellent tools to merge AI with MR, fostering both language development and content mastery (Celik & Kara, 2024). Such applications, which simulate authenticity and domain-specific experiences and tasks where language use is naturally embedded, align with Content-Based Instruction (CBI) and ESP principles.

AI-enabled MR settings can create the effect of interactive classrooms where personalized learning is delivered that adapts to the pace and needs of the student. During engineering and

medical education, the integration of MR with AI can assist in developing models for complex procedures, which may be helpful for students to practice in real-life scenarios with instant feedback (Syahrin & Akmal, 2024). The studies from engineering, medical surgeries, fashion, and retail industries (Aljuaid, 2024; Gil-López et al., 2023; Hemphill et al., 2021; Kim, 2023; Kumar et al., 2022; Luan et al., 2020; Taşdemir et al., 2025) show how these technologies can be used to improve efficiency and effectiveness.

This study is guided by theoretical frameworks, including Cognitive Load Theory (CLT) and Social Constructivism. It shows how an AI-enhanced MR model can be integrated in pedagogy to support students in understanding of the complex educational challenges they face. A study by Divekar et al. (2021) has used a similar kind of model focusing on the importance of personalized, immersive, and culturally appropriate design to support language learning in authentic contexts. Yannier et al. (2024) further support that these tools can be practical for providing students with adaptive feedback and scaffolding complex tasks. The Cognitive Load Theory (CLT) states that the use of instructional design load should be minimized on the learners so that the learns can focus their working memory on core learning tasks. As Van Nooijen et al. (2024) explain, "Scaffolding regulates the flow of information within the learner's working memory, thereby reducing cognitive load" (p. 1). AI-enhanced MR environments reduce this load (Yannier et al., 2024). Moreover, Social Constructivism theory supports the pedagogical rationale of the study, as it focuses on learning as a socially mediated activity, where interaction and collaboration drive knowledge construction. The research also addresses broader educational goals, including improved engagement, adaptation to regional educational needs, and fostering human-technology interaction. Aguayo and Eames (2023) justify the collaborative learning principles of the MR environment by arguing that learning among learners is because of socially mediated interactions: "students, educators, and parent helpers were involved in spontaneous learning interactions during the immersive snorkeling and visitor center experiences" (p. 69).

Methodology

Research Design

The convergent mixed-methods design of research was used by combining both the quantitative and the qualitative methodologies as suggested by Creswell & Plano Clark (2017). The learner's performances were gauged by using pre-and post-tests in the quantitative part. To explore students' learning experiences, their engagement, and contextual factors, the study used interviews, surveys, and classroom observations.

Population

The target population for this study is the undergraduate and postgraduate engineering students from universities in the Gulf region. A purposive stratified sampling strategy was used to ensure proportional representation from each country and academic level. The final sample of 300 participants was distributed approximately equally across the countries and disciplines. The following calculation shows the validity of the selected sample from the Gulf countries. It was calculated using the population estimate of approximately 102,000 engineering students across the Gulf region, at a 95% confidence level and a 5% margin of error. The recommended sample size was 383 students. The following sample size formula is used for the population:

$$n = \frac{N \times Z^2 \times p \times (1-p)}{E^2 \times (N-1) \times Z^2 \times p \times (1-p)}$$

Substituting the values:

$$n = \frac{102000 \times (1.96)^2 \times 0.5 \times (1-0.5)}{(0.05)^2 \times (102000-1) + (1.96)^2 \times 0.5 \times (1-0.5)} \approx 383$$

Instruments

Table 1 provides an overview of each instrument and data collection.

Table 1. *Instrumental details*

Instrument	Description	Purpose	Data Type
Standardized Tests	Pre- and post-tests to measure ESP proficiency.	Assess pre-and post-improvement in language proficiency	Quantitative
Surveys	Questionnaires include Likert-scale items and open-ended questions.	Capture students' perceptions, challenges, benefits, and satisfaction.	Quantitative and Qualitative
Semi-Structured Interviews	One-on-one interviews	Investigate human-technology interaction	Qualitative
Classroom Observations	Real-time observations of AI-enhanced MR integration in classrooms.	Evaluate student engagement, participation, and learning dynamics.	Qualitative

Data Collection Methods

The researchers used research assistants from each of the six countries to collect data. The researchers mandated these assistants to organize, follow up with various universities, arrange, and facilitate video-conferenced focus group discussions and semi-structured interviews. The rubrics used by them to train faculty focused on critical aspects of classroom dynamics, including student engagement, use of MR technology, and interaction quality. These criteria included the frequency and quality of the student-teacher interactions in the MR settings, integrating technical English, and the effectiveness of the technology in supporting learning outcomes.

Reliability & Validity Measures

Unified rubrics were used for the standardized tests developed. The surveys were designed using standard questions and were piloted for validity and clarity by the experts. For the semi-structured interview, the open-ended questions were matched with key aspects of human-technology interaction. Lastly, the calibration sessions were conducted to use the classroom observation rubrics.

AI-Enhanced MR Curriculum Module for Technical English in Engineering

Each teaching session integrated AI-powered features to personalize the MR experience and lasted for 90 minutes a week. The AI algorithms analyzed students' interactions, provided real-time corrections to their language use, and adjusted task difficulty. First week, reviewed the course objectives and technology tutorial. Secondly, basic engineering-related terminologies were introduced during this module. The second module, during weeks two and four, focused on terminology related to core engineering disciplines. Additionally, MR was introduced to visualize the engineering processes. The third module, during weeks five and seven, focused on materials science and engineering methodologies. Moreover, MR activities were performed interactively on material properties and testing procedures. The fourth module during weeks 8 and 10 covered

research on engineering systems, i.e., HVAC or electrical grids. It also integrated MR exercises for building or repairing systems. The fifth module during week 11 focused on health, safety, and environmental standards. In the last module, the students prepared a small project to describe one engineering solution using technical English. Lastly, peer and instructor feedback were provided through MR-enabled sessions.

Results

Pre-and Post-test Quantitative Data

To evaluate the effectiveness of AI-enhanced MR technologies in improving technical English proficiency, standardized technical English proficiency tests were administered before and after using AI-enhanced MR technologies. Pre- and post-test scores were analyzed by a paired t-test to evaluate significant differences and help decide on effectiveness. Table 2 below elaborates on the descriptive statistics for the pre- and post-test scores.

Table 2. *Pre-and post-test results*

Metric	Pre-Test Mean	Post-Test Mean	Change in Mean
Overall	72.5	80.3	+7.8
Saudi Arabia	70.0	78.2	+8.2
UAE	75.0	82.5	+7.5
Qatar	73.0	81.0	+8.0
Oman	71.5	79.0	+7.5
Kuwait	74.0	82.0	+8.0
Bahrain	72.0	79.5	+7.5

The t-test results showed a significant improvement in scores, with a p-value of 0.001, well below the conventional level of 0.05. This outcome indicates that the intervention effectively enhanced the students' technical English skills and had a positive effect on student outcomes. The effect size for this study was calculated using Cohen's d formula. For instance, with a mean post-test score of 80.3, a pre-test score of 72.5, and a standard deviation of 10, the calculated effect size is $d = \frac{80.3 - 72.5}{10} = 0.78$. This value indicates a medium to large effect size, suggesting AI-enhanced MR intervention significantly improved the technical English proficiency of students.

ANOVA result

Since the data were collected from various Gulf countries, a one-way Analysis of Variance (ANOVA) was conducted to evaluate the degree of improvement. The independent variable was the country of origin. The dependent variable was the test score improvements, calculated as the change in performance from the pre-test to the post-test. The analysis produced an F-value of 2.34 together with a p-value of 0.045. The F-value represents the ratio of variance between the country groups to the variance within those groups. A higher F-value indicates a general likelihood of differences existing between groups. The p-value of 0.045 is lower than the standard alpha threshold of 0.05, which shows that the observed differences in improvement scores reach statistical significance. Figure 1 below presents a boxplot visualization showing the distribution of post-test score variation and central tendency across countries. Each box represents the distribution of scores within a country, showing the median, interquartile range, and outliers.

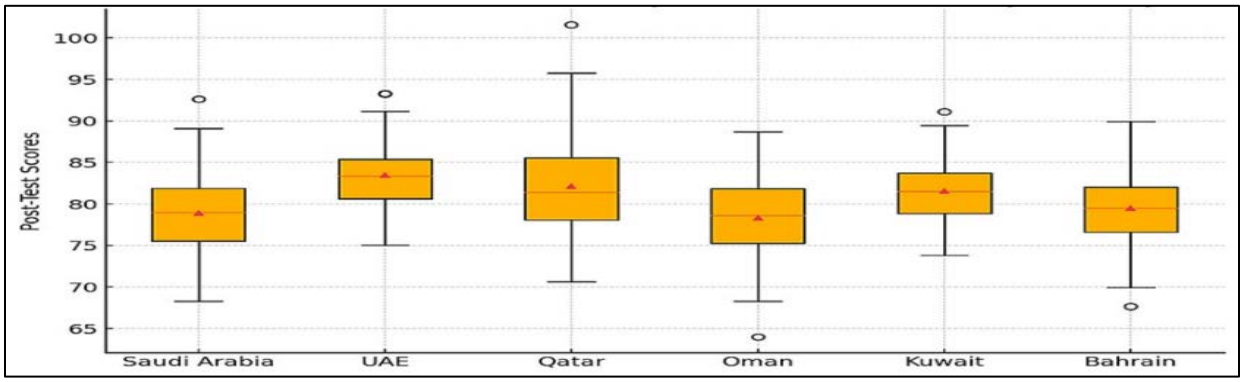


Figure 1. Variation and central tendency in post-test scores by country

The ANOVA test showed no significant difference among countries regarding the effectiveness of the AI-enhanced MR technology: F = 1.58 (p = 0.17) for five groups and 377 degrees of freedom. The technology achieved successful results in all educational environments in which it was used.

Survey Results

A quantitative survey was conducted to evaluate participants' perceptions of AI-enhanced MR technology. The survey contained four essential constructs, which used a 5-point Likert scale to measure participants' ratings for ease of use, technical English confidence, engagement level, and general satisfaction.

Table 3. *Metrics discussion*

Metric	Mean Score	Standard Deviation
Ease of Use	4.2	0.8
Confidence in Technical English	3.9	0.9
Engagement Level	4.1	0.85
Overall Satisfaction	4.0	1.0

Descriptive statistics on questionnaire data showed that, on average, students mainly experienced the AI-enhanced MR technology as easy to use, at 4.2. Also, the student's confidence level in technical English post-intervention was, on average, applicable at 3.9. Besides, the technology considerably increased engagement in learning activities, as reflected by an average score of 4.1 out of 5, indicating that students were more involved and active in the learning process with the use of technology. Figure 2 below shows the average survey scores for three essential areas: Ease of Use, Confidence in Technical English, and Engagement in Learning for a 12-week course.

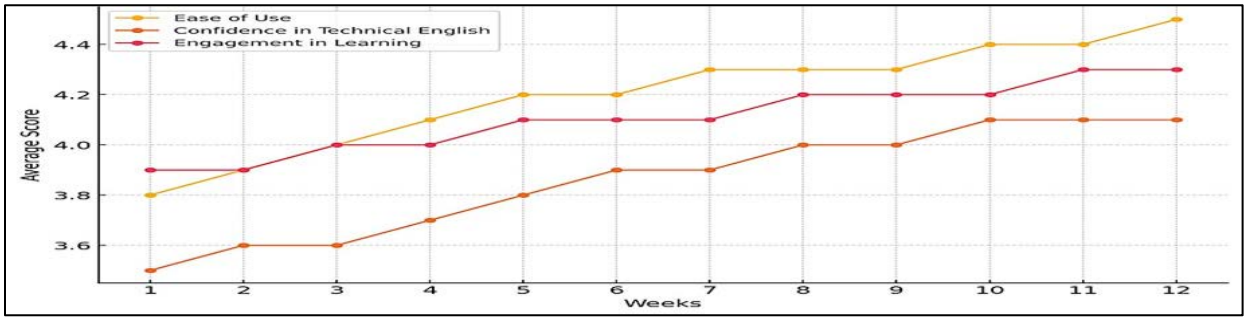


Figure 2. Survey scores over time

Qualitative Results

Thematic Analysis

The open-ended items in the survey revealed several recurring challenges and perceived benefits of using AI-enhanced MR for technical English instruction. Thematic analysis was conducted following the steps of familiarisation, coding, theme identification, and refinement, as proposed by Braun and Clarke (2006).

Semi-Structured Interview

The semi-structured interviews were analyzed to draw inferences from the qualitative data and capture a rich understanding of participants' experiences with AI-enhanced MR technology for learning technical English. Verbatim transcriptions were first created for all interviews, followed by repeated reading to familiarise the researchers with the content and identify emerging patterns.

Theme Development

Codes were further aggregated under possible themes reflecting a distinct feature of how the technology appears to aid learning in the engineering context. For "better understanding", some students reported "the visual and interactive nature of the MR environment made technical terms more understandable." For "Application in Real Life", many instructors highlighted "students are not just learning the terms; they are using them in context, which shows real learning progress." For "User Experience and Interaction", some students commented, "The interface is user-friendly most of the time, but because of occasional bugs, the sessions become frustrating." For the "Challenges and Barriers", students commented, "Sometimes the content does not align well with the local context".

Classroom Observations

For the classroom observation, multiple factors were considered to see the application of these technologies. It focused on the classroom setting and environment to determine the implementation of AI-enhanced MR technologies. Moreover, it focused on students' participation during their learning process. It was noticed that the classroom setup was appropriate for MR technology, and there were ample supplies for MR when needed in the session. In addition, MR technology was frequently embedded into observed classroom activities, which are constantly or frequently repeated within each lesson. The primary type of application used was subject-related simulations and interactive lectures intended to clarify abstract ideas. Furthermore, the use of MR technologies has resulted in a high level of engagement among the students during individual, group-based, and class-wide activities. The most significant ones were peer-to-peer teaching moments in which students taught one another concepts by using the MR applications' interactive features. Finally, the use of MR technology fulfilled the learning outcomes established for the sessions very successfully, showing significant gains in students' understanding and competency.

Discussion

Discussion on Quantitative Insights

The findings of this study strongly support the main research questions, which suggest that AI-enhanced MR technologies significantly improve technical English proficiency and that technology-integrated ESP and CBI models are effective in engineering education in the Gulf

region. Additionally, the significant gains in pre- and post-tests, with medium-to-large effect size validate that immersive AI-integrated learning environments can enhance students' language competencies in STEM settings. The results of the study also align with Cognitive Load Theory (CLT) by reducing cognitive overload and supporting dual-channel processing, allowing students to visualize complex engineering concepts (Hong & Guo, 2025). As noted in various other studies (Khodabandeh et al., 2025; Parmaxi & Demetriou, 2020), these models appear to be efficient primarily when there are high scores for usability and satisfaction.

Additionally, the study reports a 0.045 p-value in the ANOVA, which shows that these scores were not the same across the Gulf countries. These findings show the importance of more efficient technological integration and appropriate cultural contexts. The research should determine particular country pairs through post-hoc tests, which include Turkey's Honestly Significant Difference (HSD), to confirm their statistical differences.

The research results show that AI-based MR learning platforms achieve successful outcomes, which work well across different cultural settings for teaching ESP. Additionally, the absence of significant differences in the perceived effectiveness of geographical indicates the universality of the technology-enhanced language learning when properly localized and supported. The research validates CLT principles because student satisfaction depends on usability, which CLT seeks to reduce unnecessary mental work for improved task achievement. The growing learner confidence and participation levels follow CLT principles because students learn language and technology skills by using digital tools in authentic ways.

Discussion on Qualitative Insights

The semi-structured interviews, open-ended surveys, and classroom observations produced qualitative data that supported the quantitative results by documenting students' social interactions and their opinions on system usability and the educational value of these technologies. The survey results, interview data, and observational findings show that teachers found AI-enhanced MR environments to be educationally beneficial for teaching technical and STEM subjects. The data sources indicate that MR is a practical tool for teaching content-based language instruction. The students repeatedly mentioned the benefits of the visual and immersive elements of these technologies in making the complex ideas more comprehensible.

The qualitative results showed that learners obtained substantial learning advantages through their work with authentic tasks and peer collaboration, and interactive feedback during their activities. These results align with Social Constructivist theory, which supports the idea that students learn better through social interaction and active engagement with real-world problems. Students applied peer work during MR sessions to build their language skills and concept knowledge through activities, which combined ESP vocabulary for effective learning. The research evidence supports Shaddad and Jember (2024), who demonstrate that students learn language skills by studying educational materials that teach academic content through meaningful interactions and scenarios.

The qualitative data showed that students and instructors faced various technical problems, which involved software failures, hardware problems, and unreliable internet connections. The students encountered cultural differences, which created understanding challenges because their backgrounds did not match. The system needs particular educational modifications at every deployment location before implementation, and system-related problems require staff members who specialize in technical support to overcome these challenges.

Conclusion

This study investigates the effectiveness of AI-enhanced MR technologies in improving technical English proficiency among engineering students in the Gulf region. A 12-week MR-enhanced curriculum was implemented, where students engaged in context-rich, interactive simulations that combined engineering content with targeted language use. Analysis of pre-and post-test standardized data on the number of significant gains by students in their technical English skills was then validated with paired t-tests and ANOVAs for all groups of participants, native language, and previous experience with MR. Based on Cohen's d, the effect size was medium to large, signifying that the MR technologies practically amplified students' understanding and use of technical English within engineering contexts. The class observations showed a relatively high level of engagement and practical application of concepts learned in MR sessions. Qualitative data revealed that students' and instructors' reception of MR technology was favorable, as the immersive and interactive nature of the MR environment facilitated deeper conceptual understanding and improved communicative competence. Nevertheless, the study also identified several difficulties, such as technical problems, a steep learning curve, and limited cultural contextualization in the MR scenarios. The study suggested more culturally relevant content and better technical support. It concluded that AI-enhanced MR technologies represent a promising instructional tool for English for Specific Purposes (ESP) contexts, especially in STEM education. The study has implications for other disciplines where effective communication is also essential. Future research and development need to be in place to update such educational tools and address identified challenges with wider adoption in global education. Finally, ongoing research and development are key to optimizing AI-MR platforms, ensuring they remain pedagogically sound, technologically reliable, and inclusive across diverse educational settings.

Funding: This research is not funded.

Acknowledgment: Not applicable.

Conflict of Interest: The authors declare no conflict of interest.

Authenticity: This manuscript is an original work.

Artificial Intelligence Statement: During the preparation of this work, the authors used ChatGPT, Grammarly, and Keenious in order to proofread, improve readability, and find related studies with complete human supervision. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

About the Authors:

Dr. Shafiq ur Rehman holds a Ph.D. in Applied Linguistics and TESOL with 17 years of experience in EFL teaching and academic leadership. He currently works at the University of Doha for Science & Technology, Qatar. His research interests include EdTech, language policy, EMI, and ESP. http://orcid.org/0000-0002-0412-4129

Dr. Rabia Khatoon holds a Ph.D. in Applied Linguistics and multiple teaching certifications. She works at King Abdulaziz University. Her research interests include writing instruction, language pedagogy, and the use of technologies in language teaching. https://orcid.org/0009-0001-7471-233X

Dr. Intakhab Alam Khan is a professor at King Abdulaziz University and holds over 27 years of experience in education, ESP, and curriculum development. He has authored 18 books and over 75 papers and led several funded research projects at various universities. https://orcid.org/0000-0001-7058-2200

References

Aguayo, C., & Eames, C. (2023). Using mixed reality (XR) immersive learning to enhance environmental education. *The Journal of Environmental Education*, *54*(1), 58–71. https://doi.org/10.1080/00958964.2022.2152410

Ai, Z. (2024). Challenges faced by English education in ethical minority areas under the background of Bilingual Education Policy for ethical minority students: Taking Yunnan Province as an example. *Transactions on Social Science, Education and Humanities Research*, *5*, 319–324. https://doi.org/10.62051/m83s2d29

Aljuaid, H. (2024). The Impact of Artificial Intelligence Tools on Academic Writing Instruction in Higher Education: A Systematic Review. *Arab World English Journal*, *1*(1), 26–55. https://doi.org/10.24093/awej/ChatGPT.2

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Celik, B., & Kara, S. (2024). Reaping the fruits of technology-integrated grammar instruction in EFL classes at the tertiary level through web 2.0 tools. *Novitas-ROYAL (Research on Youth and Language), 18*(1), 18–36. https://doi.org/10.5281/zenodo.10973103

Chen, X., Zou, D., Xie, H., & Cheng, G. (2021). Twenty years of personalized language learning. *Educational Technology & Society, 24*(1), 205-222. https://www.jstor.org/stable/26977868

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. SAGE Publications, Incorporated.

Divekar, R. R., Drozdal, J., Chabot, S., Zhou, Y., Su, H., Chen, Y., Zhu, H., Hendler, J. A., & Braasch, J. (2021b). Foreign language acquisition via artificial intelligence and extended reality: design and evaluation. *Computer Assisted Language Learning*, *35*(9), 2332–2360. https://doi.org/10.1080/09588221.2021.1879162

Gil-López, C., Guixeres, J., Marín-Morales, J., Torrecilla, C., Williams, E., & Alcañiz, M. (2023c). Is mixed reality technology an effective tool for retail? A vividness and interaction perspective. *Frontiers in Virtual Reality, 4.10.3389* https://doi.org/10.3389/frvir.2023.1067932

Hemphill, S., Rodriguez, S., Wang, E., Koeppen, K., Aitken-Young, B., Jackson, C., Simons, L., & Caruso, T. J. (2021b). Virtual reality augments movement during physical therapy. *American Journal of Physical Medicine & Rehabilitation*, *101*(3), 229–236. https://doi.org/10.1097/phm.0000000000001779

Hong, X., & Guo, L. (2025). Effects of AI-enhanced multi-display language teaching systems on learning motivation, cognitive load management, and learner autonomy. *Education and Information Technologies 30*(12), 17155–17189 (2025). https://doi.org/10.1007/s10639-025-13472-1

Issa Ahmad AlSaleem, B. (2023). The efficiency of metaverse platforms in language learning based on Jordanian young learners' perceptions. *Arab World English Journal*, *14*(1), 334–348. https://doi.org/10.24093/awej/vol14no1.21

Khodabandeh, F., Bagheri, A., & Mobarakeh, M. I. (2025b). Augmented reality's role in EFL learning: enhancing language skills in clothing and style through flipped and blended methods. *Virtual Reality*, *29*(2),90. https://doi.org/10.1007/s10055-025-01151-y

Kim, S. J. (2023). Virtual fashion experiences in virtual reality fashion show spaces. *Frontiers in Psychology*, *14*, 1276856. https://doi.org/10.3389/fpsyg.2023.1276856

Kumar, H., Gupta, P., & Chauhan, S. (2022). Meta-analysis of augmented reality marketing. *Marketing Intelligence & Planning*, *41*(1), 110–123. https://doi.org/10.1108/mip-06-2022-0221

Luan, F., Qu, B., & Pan, X. (2020). Human-machine integration interactive system based on mixed reality technology. *Journal of Physics Conference Series*, *1549*(2), 022128. https://doi.org/10.1088/1742-6596/1549/2/022128

Parmaxi, A., & Demetriou, A. A. (2020). Augmented reality in language learning: A state-of-the-art review of 2014–2019. *Journal of Computer Assisted Learning*, *36*(6), 861–875. https://doi.org/10.1111/jcal.12486

Ramadan Elbaioumi Shaddad, A., & Jember, B. (2024). A step toward effective language learning: an insight into the impacts of feedback-supported tasks and peer-work activities on learners' engagement, self-esteem, and language growth. *Asian-Pacific Journal of Second and Foreign Language Education*, *9*(1),39. https://doi.org/10.1186/s40862-024-00261-5

Strielkowski, W., Grebennikova, V., Lisovskiy, A., Rakhimova, G., & Vasileva, T. (2025). AI-driven adaptive learning for sustainable educational transformation. *Sustainable Development*, *33*(2), 1921-1947. https://doi.org/10.1002/sd.3221

Syahrin, S., & Akmal, N. (2024). Navigating the Artificial Intelligence frontier: Perceptions of instructors, students, and administrative staff on the role of artificial intelligence in education in the Sultanate of Oman. *Arab World English Journal*, *1*(1), 73–89. https://doi.org/10.24093/awej/chatgpt.4

Taşdemir, F., Atalay, E., & Çelikpazu, E. E. (2025). Development of the Grammar Learning Awareness Scale. *Frontiers in Education*, *9*, 1480823. https://doi.org/10.3389/feduc.2024.1480823

van Nooijen, C. C. A., De Koning, B. B., Bramer, W. M., Isahakyan, A., Asoodar, M., Kok, E., Van Merrienboer, J. J. G., & Paas, F. (2024). A Cognitive Load Theory Approach to Understanding Expert Scaffolding of Visual Problem-Solving Tasks: A Scoping review. *Educational Psychology Review*, *36*(1), 12. https://doi.org/10.1007/s10648-024-09848-3

Yannier, N., Hudson, S. E., Chang, H., & Koedinger, K. R. (2024). AI adaptivity in a Mixed-Reality system improves learning. *International Journal of Artificial Intelligence in Education*, *34*(4), 1541–1558. https://doi.org/10.1007/s40593-023-00388-5

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp. 196-209
DOI: https://dx.doi.org/10.24093/awej/vol16no4.11

Evaluative Adjectives with Positive Connotations in Modern English Advertising Language

Eva Leláková
Institute of Mediamatics and Cultural Heritage
University of Žilina, Slovakia
Corresponding author: eva.lelakova@fhv.uniza.sk

Nora Kupčíková
University of Žilina, Slovakia
Email: nora.kupcikova.uni@gmail.com

Received:08/19/2025 Accepted:11/07/2025 Published: 12/10/2025

Abstract
The present study explores the use, frequency, and communicative impact of positively connoted evaluative adjectives in the contemporary English advertising language. The theoretical framework is grounded in the works of Goddard et al. (2016), Hall (2002), Ke & Wang (2013), and Rigby & Lee (2024). Empirically, the study draws on a custom-built corpus titled *Language of Advertising* and developed using the Corpus Architect function of the Sketch Engine software and a seed word methodology. Eight seed words – brand, campaign, offer, discount, newsletter, launch, product, and marketing – have been identified and selected for their frequency, positive connotation, and prominence in existing advertising texts and manuals. The corpus serves as the primary research source, providing linguistic data that are extracted and analysed both quantitatively and qualitatively, enabling comparison and supporting inferences and conclusions. The findings offer insights into the strategic use of evaluative adjectives in the advertising sector and have potential applications in English for Media Purposes (EMP) as well as in the theory and practice of applied linguistics and the broader field of marketing communication.

Keywords: Advertising discourse, consumer psychology, custom-built corpus, emotional appeal, evaluative adjectives, persuasive language, positive connotations

How to cite: Leláková, E., & Kupčíková, N. (2025).Evaluative Adjectives with Positive Connotations in Modern English Advertising Language. *Arab World English Journal*, *16*(4):196-209. https://dx.doi.org/10.24093/awej/vol16no4.11

Introduction

In today's world, advertising plays an irreplaceable role in marketing and has become a pervasive element across various domains of everyday life (Bhatia, 2016; Kotler & Keller, 2016). It is not merely a tool for informing potential buyers but also a powerful means of persuasion aimed at shaping consumer preferences and encouraging brand loyalty (Janoschka, 2004). Effective advertising relies heavily on language – particularly on the strategic use of evaluative expressions that create desirable associations with a product or service (Darics & Koller, 2018). Among the most prominent linguistic tools in advertising are evaluative adjectives with positive connotations, which enhance the appeal of products by evoking favourable emotions and imagery (Bednarek, 2018). The subtle yet impactful role of such adjectives highlights the importance of understanding their function in modern English advertising discourse (Hommerberg & Don, 2015).

This study brings findings from real-life advertising texts and sheds light on how evaluative adjectives contribute to brand identity, emotional appeal, and persuasive strategies (Handzilevska et al., 2020; Pollaroli & Rocci, 2015). The outcomes of the research are relevant not only to linguists and discourse analysts but also to marketing professionals seeking to optimize the effectiveness of their campaigns through targeted language use (Bhatia, 2021; Kotler & Keller, 2016).

Literature review

The most important component of advertising is knowing the customer's wants and needs (Danciu, 2014). With the growth of social media, marketing teams can engage a wide audience, as well as target and interact with their followers (Zhu & Chen, 2015). Thanks to social media collaboration, benevolence, and goodwill, they build their business reputation and gain a better understanding of customers (Isidro, 2014). Consequently, they can inform, persuade and effectively remind the target group of the marketing message (Karlíček et al., 2016). Horňák (2003) expresses the interconnection of advertisements to the location of the action, campaign character, target group, the newness of the information, transmission media, and the method of contact with the receiver. All of them affect the success and attractiveness of promoted goods, as well as the overall building of an outstanding brand name.

Language is an integral part of advertising and is "jointly covered by linguists and business scholars, thus integrating theoretical foundations from both disciplines" (Hornikx, et al., 2023, p. 3). It pre-determines the success of the advertisement and affects the emotions of the customers when reading the names or qualities of the products introduced. "Advertisers use language quite distinctively – there are advantages in making bizarre and controversial statements in unusual ways as well as communicating with people using simple, straightforward language" (Vasiloaia 2009, p. 294). Brahim (2021) agrees and underlines the significance of a slogan that should be written in simple language to grab the reader's attention and be memorable. In this way, advertisers ensure that customers recall the advertised product effortlessly. El-Dali (2019) prioritizes the use of language to highlight the benefits of promoted products, rather than focusing solely on their provision; advertising language should primarily emphasize the customer's gain and its positive contribution to his or her life. This bonus makes the product seem like a necessity for having a fulfilling life which promotes consumerism.

Simple adjectives used in advertisements may appear insignificant; however, they often carry substantial persuasive power. Belonging to an open-word class, they alter, clarify, and adjust the meaning contributions of nouns (Huddleston & Pullum, 2003) as well as provide additional

information. Ward (2024) categorizes adjectives into proper, distributive, compound, and indefinite adjectives. Building on this classification, Ittianath (2025) further includes descriptive, quantitative, possessive, demonstrative, and interrogative types. Mykhalevych (2024) asserts that, with the help of descriptive adjectives, the person can arouse customer interest and communicate more precisely by creating vivid images and offering richer depictions. Grace (2024) argues that descriptive adjectives provide context in the sentence generally regarding colours, shapes, sizes, or any other fine details – they make the sentences influential and informative. Their importance lies in painting a precise picture, voicing inner perceptions, and adding fullness to the portrayal. Simkins (2023) also refers to colour, age, material, opinion, purpose, origin, and size descriptive adjectives. They not only describe nouns or pronouns but also affect the tone and style of a text. The correct adjective selection provokes desired feelings towards a noun in the reader's or listener's perception. This is particularly noticeable in the language of advertising – the author of an advertisement can utilize specific descriptive adjectives to ignite the customer's interest (Millar, 2024).

"Evaluative language includes the range of linguistic resources that may be used to express language users' attitude or stance to entities in the real world or to propositions" (Pounds, 2015, p. 1) and it helps the writer and speaker portray his or her idea clearly and exactly. We recognize evaluative adjectives as a part of the evaluative language (Benamara et al., 2017). Their essence is not based on an objective perspective, since by their usage, the author gives his/her stance (Martin & White, 2005). Thus, they can be positive or negative, following the intended purpose. Goddard et al. (2019) view evaluative adjectives as highly biased since the meaning may be interpreted differently by different readers. Consequently, the semantics of evaluative adjectives are subjective.

Evaluative adjectives are identifiable based on their usage and the degree of personal implication in a text or speech (Stojanovic, 2015). They serve as a valuable tool in the field of advertising, enabling promoters to express evaluations of products or services and influence potential buyers. In his study, Kheovichai (2014, p. 5) separates evaluative adjectives based on their "emotive impact, aestheticism, quality, convenience, and cost-effectiveness". Goddard et al. (2016) distinguish among thirty-nine positively and negatively connotative evaluative adjectives and categorize them into five arrangements: first-person thought-plus-feeling words (A+, A-), experiential evaluators (B1+, B1-), experiential evaluators (B2+, B2-), subsequent ongoing effect on the viewer (C+, C-); and cognitive evaluations (D+, D-). The symbol '+' represents positively connotative evaluative adjectives, while '−' indicates those with negative connotations. As Rigby & Lee (2024) explain, advertisements mostly employ positively impactful adjectives to direct consumer's attention toward their emotional appeal. By incorporating storytelling and personalization, advertisements can evoke positive feelings and thoughts, making them more successful and profitable. Hall (2002) supports this idea but focuses more on emotions than cognition, explaining that consumer neuroscience explores the deep and complex areas of the human psyche that are vulnerable to emotive advertising strategies instead of cognitive rationalism.

Based on Goddard et al. (2016) classification, twenty-seven evaluative adjectives with positive connotations are chosen for the research purposes:

- A+ (great, wonderful, terrific, fabulous, awesome),
- B1+ (entertaining, delightful, fascinating, compelling, interesting, touching),
- B2+ (gripping, exciting, tense, suspenseful, stunning),

- C+ (powerful, memorable, haunting, inspiring),
- D+ (complex, excellent, outstanding, impressive, brilliant, original, clever).

Evaluative adjectives with positive connotations emphasize desirable qualities or prime aspects of people and things, reflecting favourable recognition and judgement. They are considered a vital element of advertising language, particularly in the creation process. Favourable adjectives confirm the desirable qualities the products have. They help readers develop a positive attitude toward the products (Linghong, 2006), enable potential buyers to envision the advertised goods and conceptualize their qualities, improve communication between companies and clients, and make advertisements sound more concrete, unreserved, and vivid (Khedri et al., 2022; Liu & Zhang, 2021).

Methodology

In the research process, corpus linguistics was employed as one of the primary methodologies, due to its capacity to systematically analyse authentic language data (Lacková, 2020; Mizin et al., 2025) and uncover patterns of usage that may not be immediately visible through intuition alone (Fabian, 2024). The Sketch Engine software, in combination with its Corpus Architect function, was used to construct and customize the corpus *Language of Advertising*. Sketch Engine allows automatic crawling and filtering of web data via seed terms, enhancing the efficiency and thematic relevance of the resulting corpus (Suchomel & Pomikálek, 2012). A personalized, purpose-built corpus provides a modular and highly customizable framework, allowing researchers to narrow the output to data aligned with specific research questions and linguistic parameters (McEnery & Brookes, 2022).

Building the corpus was guided by a seed word methodology, a well-established practice in corpus linguistics. Seed words, or lexical triggers, are predefined words or phrases selected to represent the core semantic field or thematic domain under investigation (McEnery & Hardie, 2012). This method is particularly useful in ad hoc corpus creation where researchers aim to explore specific lexical or discourse phenomena (Koester, 2022). Seed word filtering ensures that the retrieved data remain topically relevant and linguistically representative of the communicative domain – in this study, advertising discourse. It also allows for corpus representativeness within the thematic scope without needing to include an overwhelming volume of general texts (Hunston, 2002; Tognini-Bonelli, 2001). Additionally, seed-word-driven corpus construction aligns well with the principles of collocational priming (Hoey, 2005), where language users are sensitive to the repeated co-occurrence of certain words in context. By tracking seed items and their surrounding collocates or grammatical patterns, the research can reveal how evaluative language is primed and recycled in commercial messages. In the research, eight seed words – brand, campaign, offer, discount, newsletter, launch, product, and marketing – were chosen for their strong association with advertising and their capacity to guide corpus building and extract relevant texts rich in affective and persuasive language. Their selection was based on frequency, positive connotation, and prominence in established advertising corpora and manuals. Given the focus on the communicative effect of positively evaluative adjectives in advertising, the scope remained narrow and thematically defined. We referred to a specific functional register, in line with Koester (2022), who notes that for small, specialized corpora, even 5-10 well-chosen seeds can yield a meaningful and analysable dataset. Also, Baker (2022) emphasizes thematic representativeness over quantity, especially when the research aim is not linguistic generalization but in-depth functional analysis.

Language of Advertising thus enabled a focused investigation of how positively connoted evaluative adjectives function in authentic advertising discourse, both in terms of frequency and their communicative or persuasive impact. The integration of both quantitative and qualitative analyses allowed for a comprehensive understanding of the specialised language within the genre.

Research Objective and Hypotheses

The primary research aim was to analyse twenty-seven evaluative adjectives with positive connotations, selected from Goddard et al. (2016), with respect to the corpus *Language of Advertising*, created for this exact purpose. The research analysis depended on identifying the frequency and distribution, semantic prosody and collocational behaviour of these adjectives.

The research hypotheses were formulated based on insights from three key theoretical sources, which also guided the development of the corresponding research questions. The first research question (Q1), and consequently the first hypothesis (H1), were inspired by the list of evaluative adjectives provided by Goddard et al. (2016), as well as by Ke & Wang's (2013) assertion that, among one hundred English-language advertising slogans, eight adjectives – beautiful, perfect, pure, real, better, good, great, and best are the most frequently used.

Q1: Which of the analysed twenty-seven evaluative adjectives with positive connotations is present the most frequently in the *Language of Advertising* corpus?

H1: The evaluative adjective *great* is the most frequently used adjective in the *Language of Advertising* corpus.

Question 2 (Q2) and Hypothesis (H2) were derived both from Goddard et al. (2016) and Rigby & Lee (2024) who claim that advertisements mostly select positively impactful adjectives to shift the consumer's focus toward their emotional appeal.

Q2: Which category of evaluative adjectives appears the most frequently in the *Language of Advertising* corpus?

H2: The most frequent category of evaluative adjectives in the *Language of Advertising* corpus is the *First Person Thought-plus-Feeling Words* category.

And finally, Question 3 (Q3) and Hypothesis 3 (H3) were motivated by Hall (2002) who expounds the shift from the cognitive processes used in the language of advertising towards the use of feelings and emotions at the present time.

Q3: Which category of evaluative adjectives appears the least frequently in the *Language of Advertising* corpus?

H3: The least frequent category of evaluative adjectives in the *Language of Advertising* corpus is the category of *Cognitive Evaluations.*

Results

In the first stage of the research procedure, we examined 27 positively connoted evaluative adjectives, classified into four semantic categories. Each adjective (i.e., the node word) was subjected to both quantitative and qualitative corpus analysis, primarily using the concordance function. The concordance lines enabled us to identify and extract both the raw frequency (number of hits) and normalized frequency (per million tokens), as well as observe semantic references and patterns of usage.

Table 1. *Statistical overview of the semantic categories of evaluative adjectives*

Category	*No*	*Hits*	*Mean (%)*	*Median (%)*	*Range (%)*	*Most Frequent (hits)*	***Least Frequent (hits)***
First Person Thought …	5	1,506	190.54	12.02	885.00	great (1,406)	wonderful (19)
Experiential Evaluators	11	1,225	55.11	8.86	306.81	compelling (485)	gripping, tense (0)
Ongoing Effect on Viewer	4	796	125.39	103.43	296.69	powerful (469)	haunting (0)
Cognitive Evaluations	7	601	74.06	72.75	115.13	excellent (209)	outstanding (27)

Table 1 presents a statistical overview of four semantic categories of evaluative adjectives. Each category is assessed through key metrics – number of adjectives, total frequency (hits), mean and median percentage occurrence, range, and extremes of usage. The results offer insight into how various adjective types contribute differently to persuasive communication in advertising discourse.

First Person Thought-plus-Feeling Words

First Person Thought-plus-Feeling Words category contains only five adjectives, yet it produces the highest total number of hits (1,506) and the highest mean occurrence percentage (190.54%). The adjective *great* alone accounts for 1.406 hits – 93.4% of the category's total. The adjectives *wonderful*, *terrific*, *fabulous* and *awesome* occur much less frequently, as reflected in the low median (12.02%) and an extensive range (885.00%). The adjective *great* cooccurs in the corpus texts with nouns like customers, sales, brand, social media, newsletter, market, discount, offers, marketing, and launch, enriching the language of advertising substantially.

<s> *"It's great to involve your audience and make them part of the brand-building process."* <s>

<s> *"Building a newsletter list is a great opportunity to send news about products or offers."* <s>

The adjective *wonderful* collocates with advertising words such as product, purchase, marketing tactics, value, product launch emails, discount marketing, brand building, and money. It is used frequently while expressing gratitude and praise.

<s> *"It is a* ***wonderful*** *moment to create a convincing value offer."* <s>

<s> *"Today, we're going to delve into the* ***wonderful*** *world of discount marketing."* <s>

The adjective *terrific* functions as a keyword in the context with words like price point, promotions, discount coupons, gift cards, customers, and sales. Collocations semantically refer to giving advice or information.

<s> *"Ocean to Orbit offers top-notch form and function at a* ***terrific*** *price point."* <s>

<s> *"Seasonal promotions are a* ***terrific*** *way to make consumers feel valued and appreciated."* <s>

Fabulous is mainly used in contexts such as an online marketplace, discount websites, email newsletters, and campaigns, and specifically conveys a sense of similarity and persuasive intent.

<s> *"Well, creating a* ***fabulous*** *newsletter is almost like creating a lip-smacking dish."* <s>

ISSN: 2229-9327

<s> *"We initiated a guide for you, from scratch to the creation of a* ***fabulous*** *email newsletter."* <s>

Awesome can be frequently found in collocations with web design, feedback, profit, free trial, leadership, product launch emails, target audience, or email coupons. It highlights a perspective of positive warning or appreciation.

<s> *"Customers, employees, analysts, and even competitors will take notice of this* ***awesome*** *alignment, but even more importantly, this alignment will drive rapid success and market leadership!"* <s>

<s> *"If you're a web designer, send out* ***awesome*** *designs you found on Dribbble, and if you're a photographer, share your favourite Instagram photos of the week."* <s>

Experiential Evaluators

With eleven adjectives and a total of 1,225 hits, experiential evaluators have a modest mean occurrence (55.11%) and the lowest median (8.86%) across all categories. Notably, *compelling* (485 hits) and *exciting* (178 hits) contribute significantly, while *gripping* and *tense* had zero occurrences. C*ompelling* and *exciting* are strongly associated with storytelling or entertainment-based advertising, while others are less functional or appealing in advertising contexts. These findings suggest that contextual relevance plays a critical role in adjective choice within this semantic category. *Entertaining* is observed with expressions like commerce, business, brand recognition, tech marketing, and headlines. Such usage demonstrates its vitality when informing, recommending, or even judging certain factors.

<s> *"They present news in an* ***entertaining*** *way using catchy headlines, and fun graphics."* <s>

<s> "*When your audience finds your content valuable or* ***entertaining****, they're more likely to share it with their networks, increasing brand visibility and attracting new potential customers*." <s>

Delightful is widely used in advertising and marketing collocations, appearing with terms such as newsletters, subscribers' inbox, discount, brand ambassadors, and business potential.

<s> "*Fun newsletters are the sprinkle of joy in your inbox, offering a* ***delightful*** *break from the usual*." <s>

<s> "*When customers get* ***delightful*** *deals, they're more likely to become loyal patrons*." <s>

Fascinating appears in collocations with product development, price, strategy, branding, and the pre-launch period. It is used to emphasize a person's or product's abilities and characteristics.

<s> "*First off, John's* ***fascinating*** *insights on product development can really help you if you're trying to get better at managing projects that involve a bunch of different teams*." <s>

<s> "*What makes this strategy more* ***fascinating*** *is that the brand gains more traction and this makes the business to benefit much more than the discount's monetary value*." <s>

Compelling occurs with launch, database, subscribers, packaging, company, teasers, brand message, logos and taglines, the market, and flash sales.

<s> "*The insights will help you craft a* ***compelling*** *and relevant brand story*." <s>

<s> "*Creating a Sense of Urgency Popups create urgency,* ***compelling*** *customers to act fast*." <s>

Interesting can be detected in 164 concordance lines, signifying its wild use in the advertising language context. The most frequent collocating words/phrases are prizes, stocks, campaigns, subscribers, newsletters, emails, surveys, money, team, or early access. The adjective signals subjective opinions of marketers.

ISSN: 2229-9327

<s> "*It brings the focus on real-life product use cases with an* ***interesting*** *example*." <s>
<s> "*One particularly* ***interesting*** *campaign from Super Bowl LV was a contest from Mountain Dew to promote their new flavour MTN DEW MAJOR MELON*." <s>
Touching pops up in the context of newsletters and emails, exhibiting a rather descriptive character with a focus on emotions.
<s> "*Asana covers a few topics,* ***touching*** *upon productivity and career pathing.*" <s>
<s> "*This* ***touching*** *message comes to a movie about the power of running.*" <s>
The linguistic co-text of *exciting* predominantly includes e-newsletters, discounts, influencers, product launches, deals, audiences, participants, rebranding campaigns, and community-building events. Its main purpose is to bring newness and necessity to the spotlight.
<s> "*If you spend months working on a new product and you don't send an* ***exciting*** *product launch announcement email, no one will notice your product*." <s>
<s> "*By offering* ***exciting*** *prizes and asking participants to engage with your brand, you increase brand awareness and encourage users to become more involved with your business*." <s>
Suspenseful has been detected in collocates with the words campaign, viewer's attention, and new product release. Its importance lies in the edgy emotions.
<s> "With **suspenseful** music, the campaign ad initially sets a chilling tone." <s>
<s> "Apple's keynote presentations are a great example of how the company turns new product releases into **suspenseful** events." <s>
We can find *stunning* in the texts referring to brand identity, design, branding strategies, email campaigns, target audience, editing, commercials, and giveaways. It is used to prioritise recommendations or to reinforce the characterisation of the external appearance of people or objects.
<s> "*You can also create* ***stunning*** *designs using beginner-friendly free design tools and repurpose existing content from your blogs, articles, and social media or use user-generated content*." <s>
<s> "*To change this perception, the brand launched the "The Man Your Man Could Smell Like" campaign, featuring memorable commercials with a* ***stunning*** *model*." <s>

Subsequent Ongoing Effect on the Viewer

Subsequent Ongoing effect on the Viewer category includes only four adjectives but achieves a high total of 796 hits and a mean occurrence of 125.39%. *Powerful* (469 hits) and *memorable* (250 hits) dominate while *haunting* is not used. Adjectives in this category contribute to the lasting emotional or psychological impact of an advertisement. Their strength lies in creating enduring associations with the product. The strong performance of *powerful* and *memorable* indicates that advertisers strategically leverage long-term affective responses as persuasive tools. *Powerful* appears in collocations with social media platforms, marketing toolbox, slogan, eye-catching logo, consumer behaviour, discounts, entrepreneurs, cost-effectiveness, and bargain.
<s> "*It's no secret that the psychological concept known as FOMO, or "fear of missing out", can be a* ***powerful*** *way to influence consumer behaviour*." <s>
<s> "*Partnering with influencers is one of the most* ***powerful*** *product or brand launch ideas in today's social media-driven world*." <s>
Memorable presents itself through the usage in context with ads, product launch parties, checkout, sales, and promotion techniques. The semantics of the collocations refer to the

ISSN: 2229-9327

catchiness and long-lasting positive impact of adverts.
<s> "*Along with helping your customers to buy the right products the first time, you'll remove another barrier to checkout when you give them a* ***memorable*** *discount code*." <s>
<s> "*Develop an engaging and* ***memorable*** *brand name that resonates with the target audience.*" <s>
Inspiring appears in connection with words and word combinations such as email campaigns, banners, flyers, business resources, newsletters, product teams, product orders, planning, and management. It primarily conveys an enthusiastic approach.
<s> "*The email opens with the* ***inspiring*** *tagline: "Set new hair goals with the Dyson Airwrap*." <s>
<s> "*As the former VP of Product Design at Facebook, Julie uses this newsletter to share her feelings, musings, and thoughts about product management, all compelling and deeply* ***inspiring***." <s>

Cognitive Evaluations

Cognitive evaluators – 7 adjectives appearing in 601 hits – represent the most statistically stable category. It has a moderate mean (74.06%), a closely aligned median (72.75%), and the narrowest range (115.13%). No adjective in this category had zero hits and frequently used lexical items like *excellent* (209) and *original* (115) demonstrate its relevance. Cognitive evaluators show consistent usage, suggesting that appeals to intellectual judgment and product quality are central to persuasive advertising. The balanced distribution points to a broad applicability across genres, making this category reliable and versatile.
Complex can be observed in the context of business operations, online stores, audiences, marketing strategies, promotional offers, engagement, product catalogues, and warehouses.
<s> "*Planning a marketing launch is a* ***complex*** *process that requires careful consideration*." <s>
<s> "*Support is an important element in a customer offering if you market products that are* ***complex****, difficult to install or require maintenance*." <s>
Excellent coexists with the lexical items such as sales promotion, coupons, users, free courses, purchases, BOGO (Buy One, Get One) deals, competitive prices, and pre-orders. It refers to the correctness or high quality of products.
<s> "*His newsletter is an* ***excellent*** *example of how an all-text message can be both elegant and engaging, without relying on visual elements*." <s>
<s> "*Behind every big business in the world is an* ***excellent*** *brand marketing strategy*." <s>
Outstanding can be found along with the collocates customer support team, marketers, campaign, sales, newsletters, and product launch ideas.
<s> "*Its customized interface and* ***outstanding*** *support make it preferable to users*." <s>
<s> "*The companies share a common goal: helping marketers deliver* ***outstanding*** *digital experiences*." <s>
Impressive appears together with the words such as clients, followers, returns, campaigns, investments, banks, companies, and others.
<s> "*Brand promotion campaigns are launched to engage audiences and produce* ***impressive*** *results*." <s>
<s> "*They look* ***impressive*** *but don't show the actual impact on your brand image*." <s>
Brilliant coexists with the words and expressions such as brand performance, promoter scores,

email campaigns, money, products, marketing, designs, companies, or digital platforms. This adjective refers to the greatness and overall level of satisfaction.
<s> "*People given a chance to contribute and get involved in a plan early are more likely to be motivated to execute it efficiently and come up with* ***brilliant*** *creative assets*." <s>
<s> "*The Apple iPod is a great example of a* ***brilliant*** *product branding strategy*." <s>
Original has an unambiguous connection with lexical items such as product clearance sales, desires, advertising activities, prizes, brand stories, social media, discounted products, purchases, and feedback. It makes individuality and creativity prominent.
<s> "*Use our* ***original*** *research to learn how Gen Z likes to engage with brands through….*" <s>
<s> "*The company's product marketing was colourful and youthful, complementing the* ***original*** *"hippie" vibe the brand was once known for*." <s>
Clever is frequently used in corpus texts with terms such as viral marketing, design, limited edition products, logos, coupons, finance, websites, and customer tracking. It stresses the importance of being smart.
<s> "*Its* ***clever*** *and minimalistic approach redefined how cars were marketed*." <s>
<s> "***Clever*** *product launches engage existing customers and entice new ones*" <s>

Discussion

The research set out to explore the frequency, contextual usage, and persuasive functions of 27 positively connoted evaluative adjectives grouped into four semantic categories within advertising discourse. Through a combination of quantitative frequency counts and qualitative concordance analysis, the study provided a nuanced understanding of how evaluative adjectives function rhetorically in advertising texts.

The *First Person Thought-plus-Feeling* category, although comprising only five adjectives, showed overwhelming dominance both in total hits (1,506) and in the mean occurrence percentage (190.54%). The adjective *great* alone accounted for over 93% of this category's usage, confirming its ubiquitous role in promoting user-centric and emotionally charged brand messaging. The extensive range in this category (885.00%) indicates a heavy skew toward a few adjectives, which could imply a lexical overreliance in persuasive writing.

In contrast, the *Experiential Evaluators* category, though it contained the highest number of adjectives (11), had the lowest overall performance, both in terms of mean and median percentage and frequency spread. This result reflects the selective and context-dependent nature of adjectives in this category; words such as *compelling* and *exciting* proved effective and frequent while others like *gripping* and *tense* were unused. The data suggest that certain adjectives may not resonate within promotional language due to their emotional complexity or inappropriateness for commercial tone.

The Ongoing Effect on Viewer category, despite comprising only four adjectives, demonstrated strong effectiveness in generating lasting emotional resonance. The words *powerful* and *memorable* dominated this category, confirming that advertisers aim to evoke enduring psychological engagement rather than transient impact. The absence of *haunting* aligns with advertising's generally positive and uplifting communicative goals.

The Cognitive Evaluations category was the most statistically balanced and stable. With a narrow range and consistent frequency, this category reflected a rational, credibility-based appeal strategy in advertising. Adjectives like *excellent*, *original*, and *clever* were well represented,

ISSN: 2229-9327

indicating that intellectual evaluation continues to be a persuasive element across advertising genres, though not as dominant as emotional appeals. Hypotheses testing further supported the reliability of the findings.

Hypothesis 1 and Hypothesis 2 were confirmed by the corpus data, demonstrating both the prevalence of *great* and the frequency dominance of the *First Person Thought-plus-Feeling* category. However, Hypothesis 3 was refuted, as the *Experiential Evaluators*, not the *Cognitive Evaluations*, formed the least frequent category on a per-word basis. This finding reflects the need to differentiate between emotionality and functionality in advertising discourse and suggests that the presence of certain adjectives may depend more on a rhetorical fit than on their semantic category alone.

Conclusion

The research highlighted the relevance of using positively connotative evaluative adjectives in advertising language to enhance credibility and perceived impact among consumers. It revealed that modern advertising heavily prioritizes emotional resonance, relatability, and positive reinforcement over intellectual evaluation or descriptive subtlety. The *First Person Thought-plus-Feeling Words* as the most recurring and dominant category of adjectives indicated the influence of social media on customer decision-making. These adjectives, with marketing potential, can be used when sellers want to promote the authentic attributes of their products/services.

The findings of the research also confirmed that evaluative adjectives in advertising are not used uniformly across semantic categories. Instead, their selection is highly strategic, reflecting both the desired emotional impact and the rhetorical purpose of the advertisement. Adjectives such as *great, powerful, memorable,* and *excellent* dominated due to their flexibility and high persuasive value, while more context-sensitive or nuanced terms appeared less frequently.

The study has offered valuable insights into lexical choices within promotional discourse and presented a solid foundation for further exploration in the fields of evaluative language, consumer psychology, and advertising rhetoric. Its findings have practical applications for advertising professionals, educators, and discourse analysts, as they highlight the critical role of carefully selected evaluative adjectives in creating persuasive and engaging content.

Additionally, the study can serve as a valuable resource for students across a variety of study programmes and academic levels. By drawing on concrete examples from authentic advertising contexts, it enhances their understanding of the strategic use of language in promotional texts. Furthermore, the collected linguistic material can be effectively utilized by teachers and linguists to support their analyses and discussions of English, enriching classroom instruction and scholarly debate alike.

Funding: Cultural and Educational Grant Agency of the Slovak Ministry of Education

Acknowledgments: This paper has been published with the support of the Cultural and Educational Grant Agency of the Slovak Ministry of Education, project: 005ŽU-4/2023

Conflicts of Interest: The authors declare no conflict of interest.
Authenticity: This manuscript is an original work.

Artificial Intelligence Statement: During the research process, the authors employed Sketch Engine, a corpus query system that uses computational linguistics and natural language processing techniques. They also used a conversational model GPT-5 solely to support editing and improving language clarity, readability and grammaticality during the writing process. All outputs were subsequently reviewed and revised by the authors, who take full responsibility for the final content of the publication.

About the Authors:

Dr. Eva Leláková works as a senior lecturer at the Institute of Mediamatics and Cultural Heritage, University of Žilina, Slovakia. In her research she focuses on the comparative analyses of specific lexico-semantic fields, syntactic structures of literary texts, corpus linguistics and application of the results of the matrix theory in the teaching theory and practice. ORCID ID: https://orcid.org/0000-0002-5294-2505

Bc. Nora Kupčíková is a bachelor's graduate of the study programme Teaching English Language and Literature and Citizenship Education at the University of Žilina, Slovakia. Her areas of research include general and corpus linguistics. ORCID ID https://orcid.org/0009-0002-7524-4205

References

Baker, P. (2022). *Using corpora in discourse analysis.* Bloomsbury Publishing.

Bednarek, M. (2018). *Language and television series: A linguistic approach to TV dialogue.* Cambridge University Press. https://doi.org/10.1017/9781108559553

Benamara, F., Taboada, M., & Mathieu, Y. (2017). Evaluative language beyond bags of words: Linguistic insights and computational applications. *Computational Linguistics, 43* (1), 201-264. http://dx.doi.org/10.1162/COLI_a_00278

Bhatia, V. K. (2016). *Critical genre analysis: Investigating interdiscursive performance in professional practice.* Routledge. http://dx.doi.org/10.4324/9781315690315

Brahim, M. (2021). The linguistic features of advertising English. *Algerian Journal Platform, Society and Sports Journal, 5*(1), 545-553.

Danciu, V. (2014). Manipulative marketing: persuasion and manipulation of the consumer through advertising. *Theoretical and Applied Economics, 2*(591), 19-34.

Darics, E., & Koller, V. (2018). *Language in business, language at work.* Red Globe Press. http://dx.doi.org/10.1057/978-1-349-93808-7

El-Dali, H. (2019). The language of consumer advertising: linguistic and psychological perspectives. *Studies in Linguistics and Literature, 3*(2), 95-126. https://doi.org/10.22158/sll.v3n2p95

Fabian, M. (2024). Corpus-based methods of *power* notion analysis in English. C*ontemporary Studies in Foreign Philology, 1* (25), 196-208. https://doi.org/10.32782/2617-3921.2024.196-208

Goddard, C., Taboada, M., & Trnavac, R. (2016). *Semantic descriptions of 24 evaluational adjectives for application in sentiment analysis.* Canada: Simon Fraser University. http://dx.doi.org/10.48550/arXiv.1608.06697

Goddard, C., Taboada, M., Trnavac, R., & Collett, P. (2019). The semantics of evaluational adjectives: Perspectives from natural semantic metalanguage and appraisal. *Functions of Language*, *26*(3), 308-342. https://www.jbe-platform.com/content/journals/10.1075/fol.00029.god

Grace, E. (2024). 515 Descriptive adjectives to describe everything in English! htpp://www.7esl.com/descriptive-adjectives/

Hall, B. F. (2002). A new model for measuring advertising effectiveness. *Journal of Advertising Research 42*(2), 23-31. http://dx.doi.org/10.2501/JAR-42-2-23-31

Handzilevska H., Balashov, E., Shturkhetskyy, S., Kostiuchenko, O., & Shulyk, R. (2020). Barriers and resources for journalists' success in life. *Journal of Education Culture and Society, 11*(1), 174-185. https://doi.org/10.15503/jecs2020.1.174.185

Hoey, M. (2005). *Lexical priming: A new theory of words and language.* Routledge.

Hommerberg, C., & Don, A. (2015). Appraisal and the language of wine appreciation: A critical discussion of the potential of the Appraisal framework as a tool to analyze specialized genres. *Functions of Language*, *22*(2), 161-191. http://dx.doi.org/10.1075/fol.22.2. 01hom

Hornikx, J., Meurs, F., & Tenzer, H. (2023). Foreign languages in advertising: Theoretical implications for language-related IB research. *Journal of International Business Studies*, *55*(2), 1-10. http://dx.doi.org/10.1057/s41267-023-00639-6

Horňák, P. (2003). *Nová abeceda reklamy.* Bratislava: Central European Advertising, s. r. o.

Huddleston, R. & Pullum, G. K. (2003). The Cambridge grammar of the English language. *The Modern Language Review, 98*(3), 803-805. http://dx.doi.org/10.2307/3738393

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge University Press. https://doi.org/10.1017/CBO9781139524773

Isidro, I. (2014). How to use yahoo answers to promote your business. http://www.powerhomebiz.com/%20internet-marketing/%20social-media/yahoo-answers.htm

Ittianath, A. (2025). Kinds of adjectives: Understanding different types of adjectives. http://www.vedantu.com/english/kinds-of-adjectives

Janoschka, A. (2004). *Web advertising: New forms of communication on the internet.* John Benjamins Publishing Company. https://doi.org/10.1075/pbns.131?locatt=mode:legacy

Karlíček, M., Hatoňová, T., Charvát, M., Duchek, K., Smrt, M., Poucha, T., Jesenský, D., Hejl, M., Kozáková, L., Tesař, M., Král, P., Stříteský, V., Zamazalová, M., Hodboď, J., & Kúta, P. (2016). *Marketingová komunikace.* Praha: Grada Publishing, a. s.

Ke, Q. & Wang, W. (2013). The adjective frequency in advertising English slogans. *Theory and Practice in Language Studies, 3* (2), 275-284. http://dx.doi.org/10.4304/tpls.3.2.275-284

Kheovichai, B. (2014). Evaluative language in online product advertising. *Veridian E-Journal*, *7*(5), 1-13.

Khedri, S., Hasan, E., & Kritsis, K. (2022). Rhetorical structure and persuasive features of advertising: an intercultural analysis of English and Arabic online advertisements. *Russian Journal of Linguistics, 26* (3), 596-624. http://dx.doi.org/10.22363/2687-0088-24591

Kotler, P., & Keller, K. L. (2016). *Marketing management*. Pearson.

Koester, A. (2022). Building small, specialized corpora. *The Routledge Handbook of Corpus Linguistics*, 66-79. Routledge. https://doi.org/10.4324/9780367076399

Lacková, M. (2020). Corpus-based analysis of lexicosemantic behaviour of nervous system diseases names. *Arab World English Journal*, *11* (4), 50-65. https://dx.doi.org/10.24093/awej/vol11no4.4

Linghong, Z. (2006). The linguistic features of English advertising. *CELEA Journal*, *9*(1), 71-78.

Liu, S., & Zhang, J. (2021). Using metadiscourse to enhance persuasiveness in corporate press releases: A corpus-based study. *SAGE Open*, *11*(3), 1-11. https://doi.org/10.1177/21582440211032165

Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. London: Palgrave Macmillan. http://dx.doi.org/10.1057/9780230511910

McEnery, T., & Brookes, G. (2022). Building a written corpus: What are the basics? *The Routledge Handbook of Corpus Linguistics.* London: Routledge. 35-47.

McEnery, T., & Hardie, A. (2012). *Corpus linguistics: method, theory and practice.* Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511981395

Millar, S. (2024). *100 descriptive adjectives (+ Examples).* http://www. fictionary.co/journal/descriptive-adjectives/

Mizin, K., Petrov, O., Petrova. A., & Letiucha, L. (2025). Exploring cross-cultural equivalence of emotion concepts ENVY and NEID: A corpus-based study. *Arab World English Journal*, *16*(1), 336-348. https://dx.doi.org/10.24093/awej/vol16no1.21.

Mykhalevych, N. (2024). A beginner's guide to descriptive adjectives in English. http://www. preply.com preply.com/en/blog/descriptive-adjectives-in- english/

Pollaroli, C., & Rocci, A. (2015). The argumentative relevance of pictorial and multimodal metaphor in advertising. *Journal of Argumentation in Context*, *4*(2), 158-199. http://dx.doi.org/10.1075/jaic.4.2.02pol

Pounds, G. (2015). Evaluative language. *International Encyclopedia of Language and Social Interaction 3.* https://doi.org/10.1002/9781118611463.wbielsi063

Rigby, B. D., & Lee, R. (2024). Creating emotional resonance in advertising language strategies and consumer identity. *Journal of Linguistics and Communication Studies*, *3*(3), 1-13. https://doi:10.56397/JLCS.2024.09.01

Simkins, L. B. (2023). Adjectives. htpp://www. englishwithlucy.com/lesson/adjectives/

Stojanovic, I. (2015). Evaluative adjectives and evaluative uses of ordinary adjectives. *Proceedings of LENLS12: Language Engineering and Natural Language Semantics*, 1-13.

Suchomel, V., & Pomikálek, J. (2012). Efficient web crawling for large text corpora. *Proceedings of the Seventh Web as Corpus Workshop*, 39-43.

Tognini-Bonelli, E. (2001). Corpus linguistics at work. *Studies in Corpus Linguistics, Vol. 6*. Amsterdam: John Benjamins Publishing Company.

Vasiloaia, M. (2009). Linguistic features of the language of advertising. *Economy Trandisciplinarity Cognition*, *No. 1*, 294-298.

Ward, K. (2024). Nine types of adjectives all writers should know. http://www. elevenwriting.com/blog/types-of- adjectives-all-writers-should-know

Zhu, Y- Q., & Chen, H-G. (2015). Social media and human need satisfaction: Implications for social media marketing. *Business Horizons*, *58*(3), 335-345. https://doi.org/10.1016/j.bushor.2015.01.006

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp. 210-223
DOI: https://dx.doi.org/10.24093/awej/vol16no4.12

A Retrospective Narrative Inquiry into Vietnamese Learners' Learning Experiences in General English Classrooms

Bui Thi Thuc Quyen
Faculty of Foreign Languages, Ho Chi Minh City Open University, Ho Chi Minh City, Vietnam
Corresponding author: quyen.btt@ou.edu.vn

Phan Nhat Hao
School of Foreign Languages, Can Tho University, Can Tho City, Vietnam
Email: phannhathao85@gmail.com

Nguyen Thi Phuong Hong
School of Foreign Languages, Can Tho University, Can Tho City, Vietnam
Email: ntphong@ctu.edu.vn

Pham Xuan Binh
School of Foreign Languages, Can Tho University, Can Tho City, Vietnam
Email: pxbinh@ctu.edu.vn

Huynh Anh Huy
School of Education, Can Tho University, Can Tho City, Vietnam
Email: hahuy@ctu.edu.vn

Received:08/11/2025 Accepted:10/09/2025 Published:12/10/2025

Abstract

Despite growing interest in General English (GE) curricula within Vietnamese higher education, limited attention has been given to learners' lived experiences, particularly regarding how they recall and make meaning of their learning in these curricula. To address this gap, the current study employed a retrospective narrative inquiry to explore the self-reflections of undergraduates from diverse disciplines who completed three levels of a GE curriculum at a public university in the Mekong Delta region. The study was informed by the question: *How do Vietnamese EFL tertiary students experience their learning in GE classrooms?* Qualitative data were collected from five purposefully selected participants and analyzed through convergent narrative-framed thematic analysis. The findings revealed a dynamic learning process characterized by a shift from pre-course expectations to initial engagement with GE courses, mismatches between expectations and classroom realities, and the development of adaptive learning strategies. Participants initially perceived GE courses as obligatory requirements but gradually acknowledged their long-term value. They identified unmet needs within these courses. These insights highlighted the necessity for more communicative, student-centered approaches that better align with learners' academic and linguistic aspirations. Pedagogical implications include integrating small-group interaction, cultivating learner autonomy and conveying affective factors to construct safe environments for oral participation. The study contributes theoretically illustrating the interplay between learners' learning motivation, their adaptive strategies and the design of English-as-a-Foreign-Language (EFL) curricula. The study concluded by acknowledging its research limitations and proposing directions for future research

Keywords: Common European Framework of Reference for Languages for Vietnam, Communicative approach, English language learning, General English, higher education, learning experiences, Mekong Delta region, narrative inquiry
How to cite: Bui, T. T. Q., Phan, N. H., Nguyen, T. P. H., Pham, X. B., & Huynh, A. H. (2025). A Retrospective Narrative Inquiry into Vietnamese Learners' Learning Experiences in General English Classrooms. *Arab World English Journal, 16*(4):210-223. https://dx.doi.org/10.24093/awej/vol16no4.12

Introduction

Vietnam's foreign language education has been dramatically affected by its long-lasting history of outsider domination with prolonged periods under Chinese, French, American, and Soviet influences (Nguyen & Hamid, 2015; Phan, 2009). As Bui et al. (2019) asserted, such a historical background, accompanying the nation's integration into the global economy and relations, has prioritized critical roles of foreign language proficiency among its human workforce (Phan et al., 2025). Recognizing this imperative, the Ministry of Education and Training (MoET) launched the project titled "Teaching and Learning Foreign Languages in the National Education System during the Period 2008-2020," known as the National Foreign Language Project 2020 (NFLP 2020) (MoET, 2008). This initiative was subsequently extended until 2025 under the Decision No. 2080/QĐ-TTg (MoET, 2017), referred to as the Refreshed NFLP 2020. These policy enactments collectively endeavor to ameliorate the number of learners capable of communicating fluently and confidently in foreign languages (Le & Pham, 2019; Trinh et al., 2024). Among such endeavors, one of the most crucial objectives lies in enhancing "the quality of foreign language education, including the General English [GE] programs" (Lam, 2018, p. 87). Correspondingly, GE courses have systematically been established as a core component of governmental and non-governmental tertiary education institutions (Pham, 2022), becoming a compulsory requirement for all tertiary students (Bui et al., 2019; Phan & Tran, 2025; Trinh et al., 2025).

In the field of English Language Teaching (ELT), GE courses offered by higher institutions in Vietnam are considered foundational programs for undergraduates (Pham, 2022). Indeed, such courses serve as essential prerequisites for non-English-majored students to participate effectively in international exchanges and to compete successfully in the ever-changing globalized job market (Quach & Nguyen, 2024). According to Doan and Hamid (2019), whereas these courses reflect international trends, concerns have emerged regarding how they affect academic performance and learning habits of language learners, given unique localization dynamics in Vietnamese language education (Selçuk, 2025). Even though Vietnamese research into GE courses has been ubiquitously implemented, a notable lack of studies concentrating on learners' lived experiences within these classrooms is reported. Building upon these foundations, this study endeavors to explore lived experiences of Vietnamese learners in GE classrooms within the local context of the Mekong Delta region. The insights from this study are anticipated to offer critical recommendations for relevant stakeholders operating in comparable settings wherein English is formally instructed as a foreign language.

Literature Review

General English in Tertiary Education

GE in higher education refers to English-related courses endeavoring to cultivate students' proficiency in English (Far, 2008), ultimately for versatile communicative purposes. According to Campion (2016), in contrast to English for Specific Purposes (ESP), which targets the language demands of particular disciplines, GE in turn provides a broader linguistic foundation suitably applicable to different and cross-cultural contexts (Shirvan & Taherian, 2021). More pressingly, Waluyo (2020) mentioned that it functions as a linguistic cornerstone for academic advancement, continuous professional development, and continuing participation in the global arena. Expanding this perspective, Johnson and Tweedie (2021) emphasized that across universities worldwide, GE fosters communicative competence, academic literacy, and higher-order skills while nurturing transferable attributes. Although differences in curriculum orientation, class size, and pedagogical emphasis are more likely to shape learners' learning experiences, GE retains its position as a pivotal

component of undergraduate programs that prepare learners for further educational pursuits and diverse workplace contexts (Tran et al., 2023). At its core, the learning outcomes for GE articulate the expected achievements in terms of knowledge, skills, attitudes, and competencies attained by the end of each course. Drawing on Guilbert's (1987) tripartite model, these learning outcomes span cognitive, psychomotor, and affective domains, encompassing intellectual amelioration, practical skill mastery, and the cultivation of constructive dispositions. The formulation of such outcomes benefits from deliberate alignment with curriculum design and assessment practices to promote coherence and enrich learning experiences (Pham, 2022; Uygun & Cesur, 2025; Warnby, 2025).

General English in Tertiary Education in Vietnam

In response to the increasing imperative of ameliorating graduates' employability and their competencies to partake effectively in the global workforce, the Ministry of Education and Training (MoET) promulgated Circular No. 01/2014/TT-BGDDT. This mandates that tertiary students participate in GE courses and attain Level 3 proficiency in the six levels of the adapted Common European Framework of Reference for Languages for Vietnam (CEFR-VN) (Council of Europe, 2001). This adaptation embodies dynamic interconnections between globalization and localization in English language education (Lam & Albright, 2019; Nguyen & Hamid, 2020). More significantly, it functions as a pivotal catalyst for Vietnamese higher education institutions to refine their English program design, articulate explicit learning outcomes, and implement assessment mechanisms conducive to sustainable language development (Albright, 2019; Le & Chen, 2019). Furthermore, as Lam (2018) described, the CEFR-VN framework's conceptual alignment with Guilbert's (1987) tripartite model of learning outcomes provides a comprehensive foundation for language education.

The institutionalization of CEFR-VN instilled in GE courses as a compulsory subject for all non-English majors reflects strategic prioritization of English language proficiency in Vietnam's higher education. Expected proficiency thresholds are delineated in different student groups; specifically, "university students need to achieve level three (B1), and college students need to achieve level two (A2) to graduate; students who are considered ethnic minorities (including highland and tribal students) need only to achieve a lesser proficiency level of 2/6 (A2)" (Lam, 2018, p. 88). To accommodate diverse learner profiles and proficiency levels, different universities implement varied class level structures. These range from granular subdivisions of proficiency levels (e.g., A1.1, A1.2, A2.1) to broader categorizations (e.g., A1, A2.1, B1.1), exemplifying institutional selections in adapting the CEFR-VN framework to align with their educational contexts and anticipated graduate employment sectors (Pham, 2022). Accordingly, based on their assessed proficiency level, students are assigned to different course levels and curricula (Nguyen & Hamid, 2015) designed to best meet their linguistic needs and support their progress toward the mandated competency targets (Le & Pham, 2019). This tiered approach ensures that instruction is tailored, progressive, and relevant, enabling learners at each stage to build the necessary language skills required for academic success and professional readiness (Bui et al., 2019; Lam, 2018).

Related Studies

Research on GE courses in higher education has grown substantially over the past decades. It reflects increasing attention to the significant role of English language proficiency in academic and professional contexts. Despite convergent points, such investigations have been carried out in diverse educational settings, employing how different research approaches to understand how GE courses meet learners' linguistic and communicative needs.

Recent international research on the GE courses has explored pedagogical innovations and affective dimensions of language learning. Shirvan and Taherian (2021) examined the longitudinal dynamics of foreign language enjoyment (FLE) and foreign language classroom anxiety (FLCA) among 367 undergraduates enrolled in a GE course. Using latent growth curve modeling and qualitative triangulation, they reported a significant increase in FLE and a decrease in FLCA over a semester, with shifts in the correlation between these variables suggesting complex emotional trajectories. In a complementary line of inquiry, Zarrinfard et al. (2021) investigated comparative effects of mid-tech and low-tech flipped instruction, blended learning, and conventional teaching on the reading proficiency of 100 first-year university students in the on-campus GE classes. Their findings indicated that mid-tech flipped instruction, where digital materials were integrated into an educational website, yielded the most substantial gains in reading performance. Qualitative insights further suggested generally positive learner attitudes toward flipped approaches, alongside calls for adjustments in teacher roles, learner responsibilities, and instructional material design. Arguably, these two studies underscore the multifaceted nature of GE learning, valuing the connection among instructional design, learner engagement, and emotional experiences in shaping learning outcomes.

Within Vietnamese contexts, investigations into GE courses have spanned from examining policy enactment to assessing curriculum quality and workplace relevance. Le and Pham (2019), using a sequential mixed-method design, explored how GE teachers perceived the top-down adoption of the CEFR for non-English majors at a university. Although participants valued the CEFR framework and acknowledged its necessity, they also pointed out several practical barriers, including limited instructional time, inadequate access to suitable materials, and a considerable gap between learners' starting proficiency and the mandated outcomes. Similarly addressing policy-related issues, Pham (2022) analyzed curriculum documents from one public and one private university, focusing on intended learning outcomes, course structures, and assessment methods. The findings showcased that while the MoET's regulations, they adapted their curricula and assessment approaches differently to fit institutional contexts. From a graduate perspective, Tran et al. (2023) surveyed 159 employees about their previous GE courses. Respondents appreciated foundational skills gained for workplace English use and reported shortcomings (e.g., overly theoretical content, a shortage of authentic tasks in both instruction and assessment, limited formative feedback, and insufficient guidance for independent learning). Shifting to quality assurance, Quach and Nguyen (2024) adopted a mixed-method approach to investigate how GE programs were monitored and evaluated across higher education settings. Feedback from both instructors and program leaders indicated that current measures were generally effective and aligned with recognized standards; nevertheless, there was a call for a more comprehensive system that integrates student needs and expectations, along with more systematic curriculum evaluation. Arguably, these studies have revealed progress in GE provision in Vietnam and subsequently highlighted enduring challenges in aligning policy goals, teaching practices, and learner needs.

While existing studies have enriched the understanding of GE course provision, very few have delved into learners' lived experiences, particularly via a qualitative inquiry that can capture the multifaceted nature of GE learning. Addressing this gap is crucial for developing pedagogical practices that are responsive to learners' needs and aligned with the intended CEFR proficiency outcomes. To achieve this aim, the study was informed by the sole question: *How do Vietnamese EFL tertiary students experience their learning in GE classrooms?*

Research Methodology

Pedagogical Setting & Research Design

This study adopted a retrospective narrative inquiry approach, which enabled the exploration of learners' lived experiences in the GE courses after the completion of their studies. As Holland et al. (1998) defined people's experiences within their narratives, "people tell others who they are, but even more importantly, they tell themselves, and they try to act as though they are who they say they are" (p. 3). As Phan and Tran (2025) further evaluated, this methodological approach facilitates "a close examination of the [people's] lived experiences as expressed through personal stories" (p. 7). Accordingly, this qualitative design was deemed appropriate for capturing participants' personal stories and reflections, thereby offering more comprehensive insights into how they experienced and made sense of GE learning journeys.

The current study was situated in a disciplinary public university in the Mekong Delta region of Vietnam. Its GE curriculum, administered by the General English Department of the School of Foreign Languages, served approximately 1,200 learners annually. At the time of the study, the curriculum comprised three sequential courses including GE 1 (four credits), GE 2 (three credits), and GE 3 (three credits), totaling 150 teaching hours. The required exit level for the program was B1 (VSTEP Level 3) stipulated by the NFLP 2020. Following the GE curriculum, learners could proceed to discipline-specific ESP courses offered by other faculties; nonetheless, these courses were not mandatory to some disciplines in this university. Methodologically, this GE curriculum was carefully selected for the investigation because it (1) constitutes a core component of tertiary education in Vietnamese universities, (2) enrolls learners from a wide range of disciplines, and (3) has received limited research attention, particularly regarding students' lived experiences and perceived relevance to their academic and professional development.

Participants

This study involved five undergraduate students who had completed all three levels (Levels 1-3) of the researched GE curriculum. They were purposefully selected since their full completion of the GE sequence indicated that they could provide comprehensive reflections on their entire learning journey. Subsequently, these participants majored in different academic disciplines, ensuring that multiple perspectives and experiences were represented. The following table presents participants' demographic information.

Table 1. *Demographic information of participants*

Participant	Gender	Age	Year of Study	**Academic discipline**
P1	Female	19	Year 2	Business Administration
P2	Male	20	Year 3	Information Technology
P3	Female	21	Year 3	Tourism and Hospitality Management
P4	Male	22	Year 4	Mechanical Engineering
P5	Female	20	Year 2	Bio-technology

Notably, all participants had successfully passed each GE level as part of the university's English curriculum. Their varied majors and academic years allowed for a broader view of how the GE program was perceived and experienced by students from different disciplines.

Data Collection & Analysis

Data for this study were qualitatively collected narrative-framed semi-structured interviews. Each participant was kindly asked to partake in one individual face-to-face interview conducted in a quiet room at the university to ensure privacy and comfort. All interviews were scheduled at

mutually convenient times, typically in the late afternoon after their classes, and lasted between 60 and 75 minutes. The interview protocol included a set of open-ended questions that encouraged participants to reflect on their learning trajectories, challenges, achievements, and perceptions of how the GE courses had transformed their English proficiency and study habits. Simultaneously, probing questions were asked when necessary to prompt deeper narratives and richer descriptions of lived experiences. Data saturation was reached when no new themes or insights emerged from additional interviews. In this study, saturation occurred by the fourth participant, with the fifth confirming existing patterns. This was determined through iterative coding and theme comparison, ensuring the key dimensions of participants' lived experiences were fully captured.

All interviews were conducted in Vietnamese to allow them to express their experiences fully without language barriers. With participants' informed consent, each interview was audio-recorded using a high-quality digital recorder. Immediately after each session, the recordings were transcribed verbatim in Vietnamese and translated into English. To ensure semantic accuracy and cultural sensitivity, the researchers cross-checked all transcripts against the original recordings. Additionally, a second trained teacher-researcher reviewed a sample of transcripts for verification and subsequent modifications.

The data were analyzed utilizing narrative-framed thematic analysis (Clandinin & Connelly, 2022). This approach combines elements of simultaneous narrative inquiry and thematic coding. The analysis followed several iterative steps. First and foremost, the researchers read each transcript multiple times to become deeply familiar with the participants' stories and the contexts in which they unfolded themselves. Secondly, initial narrative segments were identified, representing key episodes in each participant's learning journey. Thirdly, these narrative segments were coded inductively to generate preliminary themes related to motivation, learning strategies, emotional experiences, and perceived outcomes. Fourthly, the themes were refined and organized into broader categories that captured commonalities and contrasts across participants while preserving the integrity of individual narratives. Finally, the researchers constructed composite narratives that wove together these themes, allowing findings to reflect collective patterns and unique voices. To avoid bias of the researchers' analysis, the coding framework and thematic interpretations were peer debriefed by a teacher-researcher with experience in qualitative language education research. Discrepancies in coding or interpretation were discussed until consensus was reached. All analytical decisions were documented in an audit trail to ensure transparency and rigor in the research process.

Ethical Considerations & Trustworthiness

This study adhered to the ethical guidelines for research involving human participants established by the host institution. Participants received detailed information sheets outlining the study's purpose, procedures, potential risks, and benefits. Written informed consent was obtained, with participants informed of their right to withdraw at any time without subsequent penalty. To protect confidentiality, participants were assigned numerical codes, and all identifiable information was removed from transcripts and reports. Audio recordings, transcripts, and related documents were stored on password-protected devices accessible only to the research team and will be deleted after three years.

To establish trustworthiness, the study followed Lincoln and Guba's (1985) four criteria: credibility, transferability, dependability, and confirmability. Credibility was enhanced through prolonged engagement, rapport building, and member checking. Transferability was supported by providing rich contextual descriptions enabling readers to assess applicability to other settings. Dependability was ensured through maintaining an audit trail and peer review of research processes.

Confirmability was addressed through reflexivity practices, including reflective journaling to acknowledge and minimize potential researcher bias throughout the study.

Results

From Expectations to Initial Engagement with GE

Participants' narratives depicted the beginning of their GE classes as a blend of necessity and cautious optimism. P1 admitted that the first motivation to enroll in these courses was "just to pass the exam and move on". However, the underneath was an unspoken hope of "being able to talk with foreign customers someday". P4 entered further confessing this ambition:

> "English was never my strength. At first, I felt GE was just an obstacle I had to overcome for graduation. But part of me also wondered if I could fix my weaknesses here."

Initial expectations reflected participants' mental images of "ideal" English learning. Upon starting GE classes, nonetheless, most were met with larger-than-expected class sizes, varied teaching styles, and a curriculum that leaned toward grammar and reading comprehension in early stages. P3 expressed this dissonance:

> "I imagined we would speak a lot in class because my future job needs that skill. But in the first semester, the lessons were mainly about grammar rules and reading passages. It was different from what I thought, so I had to change my approach."

Some encountered a gap between their entry-level proficiency and the curriculum's demands. P5 additionally recalled feeling "left behind" during the first month because "the pace was faster than in high school".

Faced with these realities, participants entered a phase of recalibration. The alteration was not only about adjusting study strategies but also about reshaping their attitudes toward the GE program. P2 described this turning point vividly:

> "At first, I just studied to pass. Then I realized I could use English to understand technical documents better. That was I started to care more about improving my English."

This adjustment involved integrating outside resources, resulting in the re-conceptualization of GE courses as a foundation and acknowledging that real improvement required learning beyond the classroom.

Collectively, participants entered their GE courses with varied expectations, encountered mismatches, and responded by adjusting their learning strategies and perceptions of the program's value. This adaptive process laid the groundwork for the coping mechanisms and classroom strategies.

Classroom Day-to-Day Realities and Learner Adaptive Strategies

After the initial adjustment of expectations, participants' narratives altered toward the day-to-day realities of learning in GE classrooms. These realities encompassed a complex mix of instructional practices, classroom dynamics, and learning resource availability. These contributors collectively influenced how students engaged with the courses.

One common challenge was the large class size, which possibly limited opportunities for individual participation and made it harder for students to receive personalized feedback. P3 recalled her early frustration:

> "Sometimes, I wanted to answer or ask something, but there were too many students raising their hands at the same time, so the teacher just picked a few. It made me feel like I was not really part of the lesson and just someone sitting in the background."

This scarcity of interaction made some feel less engaged, especially those who thrived in smaller, more communicative settings. Another challenge was that the curriculum's emphasis on grammar and reading comprehension. In fact, it left less space for speaking and listening activities that many participants considered essential for their language development. P1 reflected on this trade-off:

> "I learned grammar better than before, especially tenses and sentence patterns I used to mix up. Yet, I realized my speaking did not improve much unless I practiced outside. I wished there were more role-plays or discussions. I could try utilizing what I learned right away."

Additionally, teaching styles varied across lecturers, from more traditional lecture-based approaches to interactive task-based lessons. While some appreciated the structure of teacher-centered classes, others found them less motivating.

P2 described how different teaching styles affected his engagement:

> "When the teacher explained everything from the textbook and gave us exercises to do quietly, I tended to lose focus after a while. But when the teacher used games, short videos, or small group discussions, I felt more involved. I remember one time we had to role-play as tourists and tour guides. It was so much fun that I forgot I was practicing English."

Similarly, prescribed textbooks and materials were viewed as outdated or even disconnected from real-life language usage. This mismatch encouraged students to seek authentic materials by themselves.

P4 noted down,

> "Some topics in the book felt a bit far from what we actually need. I wanted to learn how to talk about my hobbies or my major in English, not just read about random topics."

Across narratives, peer collaboration stood out as a key coping strategy. Study groups, informal speaking sessions, and shared exam preparation materials not only supported learning but also fostered a sense of community.

P5 explained how working with classmates eased her stress:

> "Studying with friends was the best part in these courses. We met at the library or even at a coffee shop to review lessons together. Sometimes we just chatted in English for fun, making mistakes but correcting each other. It felt safe to practice because no one judged."

Arguably, through these strategies, participants gradually found their own solutions to navigate the constraints of the GE classrooms. These adaptive behaviors could be perceived as central to sustaining motivation and making tangible progress despite perceived limitations in the formal learning environment.

Progress, Reflection, and Unmet Needs in the GE Learning Experiences

As participants moved through their GE courses, they recognized many positive outcomes alongside deeper shifts in their language abilities and self-perceptions. They expressed pride in mastering key skills while becoming increasingly aware of areas requiring further development, particularly in spontaneous speaking-listening comprehension.

P5 recalled a memorable moment:

> "When I first started, I could not even introduce myself properly in English. After completing the courses, I found I could hold simple conversations with foreign visitors on campus. It was not perfect; I still hesitated and forgot words at times, but crossing that invisible line from silence to speaking made me feel very proud. Even though I am still learning, that moment gave me strength to keep going."

Many clearly noticed progress in their reading and writing skills. At the same time, they admitted that speaking fluently and understanding real-time conversations continued to present significant challenges.

P1 shared her lived experiences:

> "I can write essays and understand reading texts much better than before, but when I speak, especially with native speakers, my mind sometimes goes blank. I know the words in my head, but I cannot say them quickly enough. Speaking remains the skill I need to improve the most."

Beyond skill development, participants described a remarkable change in their attitude toward learning English. What began as a mandatory subject gradually became a valuable tool for study, work, and building social connections.

Participant 4 reflected:

> "After taking these courses, I realized it was a key that could open many doors such as academic opportunities, jobs, even friendships. I am not fluent yet, but I no longer feel afraid to try."

While celebrating these achievements, participants also recognized some limitations of the program. The curriculum emphasized grammar and reading comprehension, leaving limited time for authentic speaking practice. Textbooks sometimes felt disconnected from real-life needs, which constrained their learning experience. P3 expressed:

> "I wished for more chances to speak, to make mistakes in a safe environment, and to learn from them."

Such limitations encouraged participants to take initiative in seeking additional learning resources and practice outside of class. P2 recalled:

> "I realized that real progress would not come from the classroom alone. I started watching English videos and practicing with friends. This helped me feel more confident and better prepared."

In brief, these experiences illustrated a layered learning journey marked by meaningful gains, profound self-awareness, and a positive attitude toward overcoming ongoing challenges. The participants' candid voices create a vivid picture of the balance between achievement and aspiration, showing that GE courses formed a solid foundation and motivation for continued effort toward language proficiency.

Discussion

The findings of this study provided more comprehensive insights into the multifaceted lived experiences of participants navigating GE courses within higher education. Initially driven by extrinsic motivations, participants continuously exhibited alterations towards intrinsic motivations characterized by the recognition of English as a critical communicative tool for future professional contexts despite some critical challenges during the learning journey.

Participants entered GE courses with expectations ingrained in communicative competence enhancement and interactive practices. However, the prevailing realities of large class sizes, predominantly grammar- and receptive skill-focused practices, and heterogeneous teaching styles disrupted these expectations. This resulted in necessitating a process of cognitive and affective recalibration. Such a recalibration involved adjustments in learning strategies and transformations in attitudes towards the program, reflecting learners' adaptive agency in confronting systemic

constraints. This highlighted the significance of individualizing learning, since "for each learner there is an optimal level at which they should be working. It allows teaching to focus on the strengths and weaknesses which are helping or hindering learners" (University of Cambridge, 2011, p. 13). Moreover, participants' initiatives to engage in self-directed learning and peer collaboration emphasized their proactive role in compensating for perceived curricular inadequacies. These findings also implied the persistent challenges associated with developing oral communicative competence within the formal GE. Despite demonstrable progress in receptive skills, learners reported ongoing difficulties with spontaneous productive skill comprehension. This imbalance thereby suggested a structural prioritization within the curriculum that favors receptive over productive skills, which may inadvertently hinder holistic language proficiency development. Moreover, the limited opportunities for authentic oral practices in large class settings negatively impacted learner engagement and self-efficacy. When viewed alongside the CEFR-VN standards and the objectives of the Refreshed NFLP 2020, which emphasize balanced development of all language skills and real-world communication, these findings illustrated areas of alignment as well as the opportunities for enhancement. While learners' progress was notable in receptive skills, learners expressed a desire for more opportunities to develop other skills, reflecting the need to further support communicative competence consistent with national frameworks. Significantly, affective factors emerged as salient determinants in the language learning trajectory. Participants' progression from initial apprehension and disengagement to increased confidence and positive self-perceptions aligns with established theories on the affective filter and its influence on language acquisition. The social support derived from peer networks functioned as an essential mechanism for fostering a psychologically safe space conducive to risk-taking and error correction, thereby reinforcing motivation and sustaining learner persistence. While the GE curriculum served as an vital foundation, participants articulated a more evident need for continued language development beyond formal instruction. Their reliance on external resources and autonomous learning practices signals the importance of fostering learner autonomy within and outside classrooms.

Such findings also resonate with international research that highlights the complex interplay of instructional design, learner motivation, and affective experiences in GE contexts. Zarrinfard et al. (2021) demonstrated the efficacy of mid-technology flipped instruction in enhancing reading proficiency and learner attitudes. This study extends these insights by foregrounding participants' unmet communicative needs, particularly in spontaneous speaking, advocating for more balanced skill development within GE curricula. In terms of affective dimensions, emotional trajectories identified in this study complement Shirvan and Taherian's (2021) analysis of foreign language enjoyment and anxiety. Both studies reaffirmed the significance of affective factors in shaping learners' engagement and achievement, shedding light on the need for pedagogical practices nurturing positive emotional experiences alongside cognitive growth. Within the Vietnamese context, the complications identified corroborate findings from Le and Pham (2019) and Pham (2022), who highlighted the complexities of implementing the CEFR-VN curricula amid contextual limitations. Additionally, Quach and Nguyen's (2024) exploration of quality assurance echoes the need for more comprehensive and learner-responsive evaluation frameworks to bridge policy and practice. From a graduate employability perspective, this study aligns with Tran et al. (2023), whose participants acknowledged foundational benefits of GE courses yet critiqued their insufficient focus on authentic communicative tasks and formative feedback. These studies revealed a persistent gap between curricular intentions and workplace language demands. This entails the necessity for pedagogical reforms that integrate real-world applicability and learner autonomy.

Conclusion and Implications

Participants' learning experiences in the GE courses showed a journey of evolving motivation and personal growth. Although practical goals such as passing exams initially drove their participation, many developed stronger ambitions related to real-life communication and professional development. This process encouraged learners to reflect on their abilities and adjust their approaches, which fostered academic resilience and self-awareness. The dynamic nature of their learning journey thereby highlighted the critical importance of nurturing adaptability and sustainability to achieve long-term success.

These findings generate some important pedagogical considerations for the design and delivery of GE courses. First and foremost, lecturers should strive to create a more balanced integration of both receptive and productive language skills. This integration should be with particular emphasis on developing oral proficiency and spontaneous communication. Secondly, pedagogical approaches that foster small-group interactions, peer collaboration, and task-based learning can mitigate these constraints by providing more opportunities for meaningful language practice and feedback. The cultivation of psychologically safe environments wherein learners feel comfortable taking risks is essential for encouraging oral participation and reducing language anxiety. Thirdly, the proactive role which learners took in supplementing formal instruction through autonomous study and peer support values the importance of fostering learner autonomy within the pedagogical framework. Lecturers can scaffold this by explicitly integrating adaptive strategies for independent learning. Embedding meta-cognitive awareness and self-regulation techniques into the curriculum can empower learners to take ownership of their progress beyond the classroom. Lastly, the affective dimension of language learning, evidenced by the shift from apprehension to self-confidence, should inform teaching practices. Lecturers should recognize and address the emotional factors that influence engagement and achievement by adopting empathetic, learner-centered approaches. Providing consistent, formative feedback and celebrating incremental progress can also reinforce learners' resilience.

Beyond its practical implications, this study contributes to the theoretical insights of how learners motivational shifts and adaptive strategies unfold within the structural and pedagogical realities of General English programs in EFL contexts. By linking these individual trajectories to curriculum design, class dynamics, and skill balance, the findings extend existing models of learner agency, autonomy and affective engagement. It informs both policy and classroom practice by highlighting the importance of communicative, learner- centered instruction, a balanced focus on all language skills in compliance with CEFR-VN targets, class sizes and authentic communication to narrow the gap between policy and the real-world of teaching method application.

Limitations and Recommendations

Although this study offers valuable insights into how GE courses are locally experienced, its focus on a single institutional context probably limits the generalizability of these findings. The reliance on learners' self-reported experiences without extensive data triangulation from stakeholders and objective performance data suggests the needs for future research to incorporate multiple stakeholder perspectives, including teachers and curriculum developers, to achieve a more comprehensive understanding of the GE course effectiveness. Additionally, longitudinal studies tracking learners' progress over time would provide deeper insights into the sustainability of motivational changes and skill development.

Funding: This research is not funded.
Acknowledgments: Not applicable.
Conflicts of Interest: The authors declare no conflict of interest.
Authenticity: This manuscript is an original work.
Artificial Intelligence Statement: During the preparation of this work, the authors used ChatGPT-4 to improve the clarity, coherence, and conciseness of English expression. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

About the Author:

Bui Thi Thuc Quyen (quyen.btt@ou.edu.vn) is a lecturer at Ho Chi Minh City Open University, Vietnam. Her research interests include English language teaching methodology, curriculum and syllabus design, curriculum and materials development, assessment for learning, and teachers' professional development. Her ORCID is https://orcid.org/0009-0006-1946-7497

Phan Nhat Hao (phannhathao85@gmail.com) obtained his Master of Education in Didactics from Can Tho University, Vietnam. His research interests cover English language teaching, teacher-researcher identity construction, task-based language teaching, and professional development in the context of internationalization. His ORCID is https://orcid.org/0009-0003-4069-6735

Nguyen Thi Phuong Hong (ntphong@ctu.edu.vn) is a lecturer at Can Tho University, Vietnam, specializing in English language assessment and English for Specific Purposes (ESP). Her research interests include second language assessment, language policy in higher education, and technology-enhanced language learning. Her OCID is https://orcid.org/0009-0006-8054-3401

Pham Xuan Binh (pxbinh@ctu.edu.vn) is a lecturer at Can Tho University, Vietnam. His research explores the intersection of pedagogical innovation, learner engagement, and task-based instruction. He is interested in project-based learning, teacher training, and the role of intercultural competence in language education. His ORCID is https://orcid.org/0009-0007-0797-4659

Huynh Anh Huy (hahuy@ctu.edu.vn) is a lecturer in Physics at Can Tho University, Vietnam. His research interests include atomic, molecular, and chemical physics, condensed matter physics (solid-state physics and superconductivity), and plasma and fluid physics (surface physics). His ORCID is https://orcid.org/0000-0002-4047-2331

References

Albright, J. (Ed.). (2019). *English tertiary education in Vietnam*. Routledge.

Bui, T., Nguyen, T. T. T., & Nguyen, D. A. (2019). Vietnamese higher education language planning and university students' career development. In J. Albright (Ed.), *English tertiary education in Vietnam* (pp. 54–67). Routledge.

Campion, G. C. (2016). "The learning never ends": Exploring teachers' views on the transition from General English to EAP. *Journal of English for Academic Purposes, 23*, 59–70. https://doi.org/10.1016/j.jeap.2016.06.003

Clandinin, D. J., & Connelly, F. M. (2022). *Narrative inquiry: Experience and story in qualitative research*. Routledge.

Council of Europe. (2001). *Common European Framework of Reference for languages: Learning, teaching, assessment*. Press Syndicate of the University of Cambridge. http://www.coe.int/t/Dg4/Linguistic/Source/Framework_EN.pdf

Doan, L. D., & Hamid, M. O. (2019). Economics of English: Examining the demand for English proficiency in the Vietnamese job market. *RELC Journal*, Advance online publication. https://doi.org/10.1177/0033688219863166

Far, M. M. (2008). On the relationship between ESP & EGP: A general perspective. *English for Specific Purposes World, 7*(1), 1-11.

Guilbert, J. J. (1987). *Educational handbook for health personnel* (6th ed.). WHO.

Holland, J., Ramazanoglu, C., Sharpe, S., & Thomson, R. (1998). *The male in the head: Young people, heterosexuality and power*. Tufnell Press.

Johnson, R. C., & Tweedie, M. G. (2021). "IELTS-out/TOEFL-out": Is the end of general English for academic purposes near? Tertiary student achievement across standardized tests and general EAP. *Interchange, 52*(1), 101–113. https://doi.org/10.1007/s10780-021-09416-6

Lam, L. T. (2018). General English for non-majors in higher education. In J. Albright (Ed.), *English tertiary education in Vietnam* (pp. 86–101). Routledge.

Lam, T. L. H., & Albright, J. (2019). Vietnamese foreign language policy in higher education. In J. Albright (Ed.), *English tertiary education in Vietnam* (pp. 1–15). Routledge.

Le, T. T. H., & Pham, T. H. N. (2019). Implementing the CEFR at a Vietnamese university: General English language teachers' perceptions. *CEFR Journal - Research and Practice, 1*, 41–57. https://doi.org/10.37546/JALTSIG.CEFR1-4

Le, T. T., & Chen, S. (2019). Globalization and Vietnamese foreign language education. In J. Albright (Ed.), *English tertiary education in Vietnam* (pp. 16–27). Routledge.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry* (1st ed.). Sage Publications.

MoET. (2008). *Quyet dinh ve viec phe duyet de an "Day va hoc ngoai ngu trong he thong giao duc quoc dan giai doan 2008–2020. No. 1400/QD-TTg. [Decision No. 1400/QD-TTg on the approval of the project "Teaching and learning foreign language in the national education system in the period 2008–2020]*. Hanoi, Vietnam: Vietnam National Politic Publications.

MoET. (2017). *Quyet dinh ve viec phe duyet dieu chinh, bo sung de an day va hoc ngoai ngu trong he thong giao duc quoc dan giai doan 2017–2025. No. 2080/QDTTg [Decision No. 2080/QD-TTg on the approval, adjustment and supplement of the project "Teaching and learning foreign language in the national education system in the period 2008–2020]*. Hanoi, Vietnam: Vietnam National Politic Publications.

Nguyen, V. H., & Hamid, M. O. (2015). Educational policy borrowing in a globalized world: A case study of Common European Framework of Reference for languages in a Vietnamese university. *English Teaching: Practice & Critique, 14*(1), 60-74. https://doi.org/10.1108/ETPC-02-2015-0014

Pham, C. (2022). General English programs in tertiary education in Vietnam. *Language Related Research, 13*(5), 169–196. Language Related Research. https://doi.org/10.52547/LRR.13.5.4

Phan, N. H., & Tran, Q. H. (2025). A retrospective narrative inquiry into ecological influences on a high school EFL teacher's teacher-researcher identity negotiation. *Educational Point, 2*(2), e124. https://doi.org/10.71176/edup/16728

Phan, N. H., Trinh, Q. L., & Le, C. T. (2025). Factors hindering EFL Master's students' motivation in research engagement: A self-determination perspective. *Pakistan Journal of Life and Social Sciences, 23*(1), 1782-1795. https://doi.org/10.57239/PJLSS-2025-23.1.00138

Phan, T. T. H. (2009). Impacts of Vietnam's social context on learners' attitudes towards foreign languages and English language learning: Implications for teaching and learning. *Asian EFL Journal, 11*(4), 169-188. https://usnpendbing.wordpress.com/wp-content/uploads/2015/03/december-2009.pdf#page=169

Quach, T. T. N., & Nguyen, L. (2024). Using accreditation standards for internal quality assurance for tertiary general English program in Vietnam. *International Journal of TESOL & Education, 4*(3), 1–19. https://doi.org/10.54855/ijte.24431

Selçuk, M. (2025). Exploring the validity of a CEFR-aligned placement test through norm-referenced and criterion-referenced approaches. *Edu 7: Yeditepe Üniversitesi Eğitim Fakültesi Dergisi, 15*(17), 48–70. https://dergipark.org.tr/en/download/article-file/4915868

Shirvan, M. E., & Taherian, T. (2021). Longitudinal examination of university students' foreign language enjoyment and foreign language classroom anxiety in the course of general English: Latent growth curve modeling. *International Journal of Bilingual Education and Bilingualism, 24(1), 31–49.* https://doi.org/10.1080/13670050.2018.1441804

Tran, L. H. N., Nguyen, P. A., Nguyen, A. T., & Nguyen, T. M. D. (2023). Graduates' feedback on the contribution of general English courses to their English skills for work purposes. In N. H. L. Tran, M. T. Ngo, & L. T. Tran (Eds.), English language education for graduate employability in Vietnam (pp. 207–232). Springer.

Trinh, Q. L., Nguyen, T. H., Le, C. T. & Ngo, H. H. N. (2024). EFL teachers' perceptions and practices of reflective teaching as a means for professional development. *Arab World English Journal, 15*(4), 400-417. https://dx.doi.org/10.24093/awej/vol15no4.25

Trinh, Q. L., Phan, N. H., Le, T. N., Luu, V. H. N., & Le, C.T. (2025). Investigating factors affecting the completion of EFL teachers' graduate studies from Social Cognitive perspectives. *Arab World English Journal, 16*(2), 262-275. https://dx.doi.org/10.24093/awej/vol16no2.15

University of Cambridge. (2011). *Using the CEFR: Principles of good practice*. Retrieved from www.cambridgeenglish.org/images/126011-using-cefr-principles-of-good-practice.pdf

Uygun, E., & Cesur, K. (2025). Teaching English for general purposes through a MOOC in higher education: A quasi-experimental study to explore its effects on learner achievement and satisfaction. *Education and Information Technologies, 30*(3), 3389-3409. https://doi.org/10.1007/s10639-024-12948-w

Waluyo, B. (2020). Learning outcomes of a general English course implementing multiple e-learning technologies and active learning concepts. *Journal of Asia TEFL, 17*(1), 160-181. https://doi.org/10.18823/asiatefl.2020.17.1.10.160

Warnby, M. (2025). Relating academic reading with academic vocabulary and general English proficiency to assess standards of students' university-preparedness: The case of IELTS and CEFR B2. *Scandinavian Journal of Educational Research, 69*(3), 506-523. https://doi.org/10.1080/00313831.2024.2318434

Zarrinfard, S., Rahimi, M., & Mohseny, A. (2021). Flipping an on-campus general English course: A focus on technology complexity of instruction and learners' levels of impulsivity. *International Journal of Educational Technology in Higher Education, 18*(1), 1-25. https://doi.org/10.1186/s41239-021-00255-7

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp.224-238
DOI: https://dx.doi.org/10.24093/awej/vol16no4.13

Conversational Analysis of Learner – AI Chatbot Interactions in Developing Spoken Fluency

[1]Anna Matiienko-Silnytska, [2]Nana Mikava, [2]Iryna Savranchuk, [2]Neonila Tkhor, [3]Hanna Poliakova

[2]Odesa I.I. Mechnikov National University, Odesa, Ukraine
[3]Odesa National Maritime University, Odesa, Ukraine
[1]Corresponding Author: departmentofforeignlanguages4@gmail.com

Received: 08/13/2025 Accepted:11/19/2025 Published: 12/10/2025

Abstract

This study investigates interactions between AI chatbots and EFL learners and their effect on spoken fluency and communicative competence, focusing on micro-level features of discourse such as turn-taking management, discourse-marker use, hesitation phenomena, and repair mechanisms. Thirty-first- and second-year undergraduates (CEFR B1–B2) at Odesa I. I. Mechnikov National University completed two counterbalanced 20-minute speaking sessions: a voice-based chatbot conversation and a matched human–human peer task. Audio data were transcribed and annotated for turn-entry timing, discourse markers, hesitation, and repair sequences. Fluency indices included speech rate, articulation rate, mean length of run, pause profiles, and repair density. The findings indicate that chatbot interaction yielded modest but reliable improvements in speech rate and mean length of run, alongside a reduction in very long pauses, while articulation rate remained essentially unchanged. Self-repair density increased, with most repairs being self-initiated and quickly resolved, often triggered by lexical access challenges or automatic-speech-recognition misrecognitions; these sequences supported conversational progressivity rather than disrupting flow. Gains were most evident in information-gap and problem-solving tasks, whereas opinion-exchange tasks showed weaker effects due to limited backchanneling and occasional system latency. Learners perceived chatbot practice as useful, less anxiety-provoking than peer interaction, and supportive of willingness to communicate, while noting limitations in discourse naturalness and prosodic feedback. The significance of the study lies in linking interactional micro-features to fluency outcomes, offering empirical evidence for AI-mediated speaking practice design. Pedagogically, results support blended implementation: chatbots for warm-up, rehearsal, and decision-focused tasks, complemented by human interaction to develop pragmatic nuance and prosody.

Keywords: Conversational AI, conversation analysis, EFL speaking, spoken fluency, speech rate, pauses

How to cite: Matiienko-Silnytska, A., Mikava, N., Savranchuk, I., Tkhor, N,, & Poliakova, H. (2025). Conversational Analysis of Learner – AI Chatbot Interactions in Developing Spoken Fluency. *Arab World English Journal*, *16*(4): 224-238. https://dx.doi.org/10.24093/awej/vol16no4.13

Introduction

In the last few years, the use of artificial intelligence to learn foreign languages has radically transformed how English as a Foreign Language (EFL) is taught and learnt. Of these innovations, conversational AI chatbots have been significant agents that simulate actual communication and give learners the opportunity to interact meaningfully. As a difference from traditional language learning courses that take a form of mechanical drills, AI chatbots can generate spontaneous context-based conversation and hence give learners a space to exercise fluency in an interactive environment. This shift is in line with communicative and interactionist models of language learning, where real-life input, meaning negotiation, and building communicative competence are emphasized. Recent research underscores this trend. For instance, Koç and Savaş (2025) report that AI-enhanced conversational practice can increase learners' fluency-related behaviors, particularly in turn-taking and output quantity. In contrast, Markus and Kirner-Ludwig (2024) note that AI-driven linguistic interfaces may influence pragmatic choices and lexical variation, pointing to evolving forms of interactional competence. Yet despite these emerging insights, empirical evidence on how AI-mediated conversation shapes spoken fluency remains limited, particularly regarding the micro-pragmatic mechanisms through which learners manage turn-taking, deploy discourse markers, negotiate breakdowns, and execute repair. The present study addresses this gap by applying conversation-analytic principles to investigate learner–chatbot interaction as a site of talk-in-interaction. Through this lens, the research seeks to illuminate how interactional practices with AI facilitate or constrain the development of natural fluency, and how they compare to peer communication in shaping communicative performance.

The aim of the paper is to investigate how interactions in discourse between AI chatbots and EFL learners affect spoken fluency and communicative competence. The present study is focused on the micro-level features of learner–chatbot interaction, including turn-taking management, use of discourse markers, hesitation phenomena, and repair mechanisms. Based on the analysis of these features, the study will help to identify the potential advantages and disadvantages of AI-based interaction for facilitating natural communicative practice.

For achieving this goal, the study sets the following tasks: to explore the structural make-up of learner–AI chatbot interaction from a conversational analysis perspective; to identify the linguistic and pragmatic means learners draw upon when co-constructing meaning through chatbot-mediated interaction; to ascertain the extent to which chatbot interaction supports learners' speech fluency, with consideration for indicators of fluency like speech rate, pauses, and self-repair; to compare students' attitudes towards chatbot-based speaking practice with conventional communicative practice.

Literature Review

Recent policy and practice debates frame conversational AI as both opportunity and risk in higher education, highlighting the need for mechanism-level evidence in speaking pedagogy. Commentaries on academic integrity emphasize that generative systems challenge authorship and

assessment but can support transparency and learning when coupled with appropriate norms (Stanko, 2024). Surveys of academics similarly report perceived efficiencies and personalization alongside concerns about reliability, bias, and workload, underscoring the importance of training, guidelines, and thoughtful task redesign. Within language education, practitioner accounts note accessibility and engagement benefits while stressing principled integration and assessment guardrails (Patty, 2024). Collectively, these perspectives establish institutional relevance but operate at a macro level, offering limited insight into how AI-mediated talk shapes real-time interaction and spoken fluency.

Research in human–computer interaction narrows this gap by demonstrating that AI mediation can alter micro-pragmatic communication features such as grounding and ambiguity resolution (Mieczkowski et al., 2021). These findings foreground constructs central to Conversation Analysis (CA)—turn-entry timing, clarification sequences, and repair—and suggest that any fluency effects of chatbot practice likely arise through such interactional pathways. Work in lexicography further highlights the importance of language resources and transparency in AI systems (Markus & Kirner-Ludwig, 2024), a concern that translates in chatbot contexts to lexical coverage, dialect sensitivity, and interpretable candidate understandings.

Parallel developments in technology-enhanced language learning emphasize the communicative potential of dialogic AI systems. Meta-syntheses report positive effects on communication, affect, and participation, while also noting substantial variation across chatbot modality, task alignment, and assessment practices (Du & Daniel, 2024; Koç & Savaş, 2025). Primary studies show converging evidence that chatbots can improve speaking performance (Tai & Chen, 2024), increase willingness to communicate, and reduce anxiety (Wang et al., 2024; Wu et al., 2025), including in self-directed settings (Kang & Sung, 2024). A PRISMA-guided review adds that benefits appear strongest when chatbots support informal, CEFR-aligned practice, though existing studies remain short, small-scale, and methodologically heterogeneous (Klímová & Ibna Seraj, 2023).

Conversation-analytic work helps specify relevant interactional mechanisms. Task-based chatbot research shows high engagement and successful turn-taking in young EFL learners (Yang et al., 2022). Human–computer interaction studies identify disclosure, questioning, empathy markers, and humor as strategies shaping participation and reciprocity (Croes et al., 2023). Recent CA-informed analyses of virtual assistants foreground repair initiation and resolution as key to maintaining progressivity in AI-mediated dialogue (Pütz, 2024; Galbraith et al., 2024).

However, most studies assess “speaking performance” via global ratings rather than linking micro-interactional behavior (e.g., turn-entry, overlap, repair) to established fluency indices such as speech rate and pausing (Du & Daniel, 2024; Koç & Savaş, 2025; Wang et al., 2024; Tai & Chen, 2024). Few directly compare learner–AI vs. peer interaction under matched tasks.

The present study addresses these gaps by applying CA to learner–chatbot interaction and connecting turn-taking, discourse-marker use, hesitation, and repair to quantitative fluency

metrics. This integrated approach extends the field from broad advocacy or caution toward a process-based account of how fluency develops during AI-mediated conversation.

Methodology

This study used a mixed-method, within-subjects crossover design incorporating conversation analysis and quantitative fluency testing. Two matched speaking sessions each lasting 20 minutes, spaced one week apart, were administered by each participant: a session with an AI-chatbot voice interface and a human–human peer session, ordered randomly to minimize practice effects; post-session perception questionnaires triangulated.

Participants

Participants were 30 first- and second-year ONU students (18 female, 12 male; age range 18–21, Mage ≈ 19). Criteria for inclusion were current enrolment, no reported hearing or speech impediment, and consent for audio recording. Participation was voluntary, though recruitment occurred through course enrolment, and students provided written informed consent. Screening for competence via the university placement process aligned to CEFR ensured a B1–B2 level appropriate for spontaneous conversation tasks. The recordings took place in quiet laboratory environments with closed-back headsets using boom microphones.

Research Procedures

The chatbot was a large-language-model system with automatic speech recognition and text-to-speech aligned with B1–B2 conversational scaffolding, topic restrictions, safety censorship, and session logging (turn number and time stamps) **(**LLM engine comparable in scale and function to GPT-4 with integrated ASR/TTS components). The two conditions both used an identical trio of communicative tasks—an information-gap planning task under explicit restriction, an opinion-exchange discussion on regular academic or student-life matters, and a problem-solving scenario (e.g., traveling or registration issues with target specifications). For the chatbot condition, prompts were delivered in conversational form; peer dyads received the same printed prompts.

Participants completed a brief orientation and sound check, then completed the first 20-minute session and 6-minute perception survey; they completed the alternate condition with the same topics and survey one week later. An invigilator inspected timing and technical stability.

Audio was recorded at 44.1 kHz/16-bit mono. Manual corrections of machine-generated draft transcripts were carried out by trained assistants following a protocol for intonation unit and turn-level utterance segmentation and for the annotation of overlaps, truncations, filled pauses, and repairs. Double-checked samples of the 15% stratified sample per condition and task were obtained to a minimum of 98% word-level accuracy. The final corpus contains approximately 20 hours of dialogue (approximately 10 hours per condition) with turn-by-turn aligned timestamps.

Two annotators coded interactional features in ELAN/EXMARaLDA using a fine-grained codebook: turn-taking (turn length in words and syllables, turn-entry latency in ms, overlap onsets, backchannel frequency); sequence organization (adjacency pairs such as question–answer, request–response, assessment–agreement or disagreement, and their expansions); discourse markers distinguished as interactive (well, so, you know, I mean, okay) versus stance (maybe, actually, basically, I think) with positional and functional coding; hesitation phenomena (filled pauses, counts and mean duration of silent pauses with a 250-ms threshold, repetitions, truncations); and repair (self-initiated self-repair, self-initiated other-repair, other-initiated self-repair), with locus at the trouble source and outcome as replacement, insertion, or deletion. After being trained with 90 minutes of data, inter-annotator agreement was Cohen's κ of at least .80 for categorical codes and ICC(2,k) of at least .85 for continuous measures, with residual disagreements adjudicated.

Primary fluency measures from audio and annotations were speech rate (syllables per second and total duration), articulation rate (syllables per second in phonation only), mean length of run in syllables between pauses longer than 250 ms, ratio of phonation time, pause frequency and mean pause duration, repair density per 100 words, and discourse-marker density per 100 words with interactive and stance markers distinguished. Metrics were quantified per task and aggregated for modeling.

Perceptions from the learners were gathered immediately after each session through a 24-item seven-point Likert scale indexing perceived authenticity, enjoyment, cognitive load, usefulness for speaking, willingness to communicate, and speaking anxiety, followed by two open-ended questions about perceived strengths and weaknesses of chatbot vs. peer practice. The instrument adapted items from validated willingness-to-communicate and foreign-language anxiety scales; items were piloted with a small group (n=5) to ensure clarity and timing.

Statistical analyses were conducted in R. For repeated measures outcomes, linear mixed models possessed fixed effects of session, order, condition, and task type, participant random intercepts, and by-participant random slopes in condition with convergence; mixed models were selected because they account for individual differences across repeated measures and increase statistical efficiency in within-subject designs. Overdispersed count outcomes such as repair and discourse-marker densities were modeled via negative-binomial mixed models. Assumptions were checked with residual diagnostics, p-values used Satterthwaite approximation, and the Benjamini–Hochberg procedure controlled the false discovery rate across outcome families. Exploratory mediation probed whether turn-taking mechanisms such as turn-entry latency and backchannel frequency mediated condition effects on fluency. Qualitative survey comments were analyzed using reflexive thematic analysis to contextualize quantitative trends.

A priori, within-subjects design simulations indicated that a sample size of 30 has at least 80% power to detect medium-sized within-person effects on target fluency measures at $\alpha = .05$, assuming within-person correlations are moderate; crossover design increases precision by lessening between-person variance. The resulting dataset includes 30 ONU undergraduates in

years 1–2, two sessions per participant in peer and chatbot conditions with order randomized, around 20 hours of fully transcribed and time-aligned audio, fine-grained interactional annotation, objective fluency measures, and complementary perception data.

Ethical Considerations

The protocol was approved by the ONU ethics committee. Written informed consent was guaranteed; pseudonyms replaced names in all files; logs and audio were stored on encrypted servers and identifiers removed before analysis. Anonymized transcripts, codebook, and analysis scripts will be deposited on an open repository at publication, with audio access subject to further ethical clearance. The participation was voluntary and withdrawal was allowed anytime.

Results

The Structural Organization of Learner–AI Chatbot Dialogues

Drawing on conversation-analytic principles and the annotated corpus, learner–AI exchanges display a recognizable architecture encompassing openings, sequence development, repair, topic management, and closings, with several features shaped by ASR/TTS latency and the chatbot's facilitative role. Transcription follows standard CA cues: overlaps [ ], micropause (.), timed gaps in seconds (0.4), lengthening :::, cut-off -, and capitals for emphasis.

Openings commonly begin with a summons–answer sequence and a brief "readiness check," after which the chatbot issues a topic-initial elicitor that frames the activity phase. In voice mode, a technical pre-opening (hearing check) frequently precedes greetings and task launch.

Extract 1: summons–answer + task framing
1 AI: Hi there, can you hear me clearly?
2 S1: yes (.) I can
3 AI: great (.) today let's plan a weekend event together (.) are you free on Saturday?

Turn construction and turn-taking align with the turn-constructional unit and transition-relevance place system. Learners build turns with clause-based TCUs and frequent turn-initial discourse markers (well, so, okay) that project forthcoming action and secure the floor. Chatbot turns are typically one TCU ending in a question, installing conditional relevance for the next learner move. Gaps at TRPs are longer than in peer talk and are treated by learners as "system processing," delaying self-selection; overlaps are rarer and occur mainly as launch overlaps when learners anticipate the chatbot's question type.

Extract 2: delayed self-selection and projected question
1 AI: we could go to the city museum (.) or a concert (.) which sounds better to you?
2 (0.6)
3 S2: uh the museum (.) because I have never been there

Sequence organization is dominated by request–information, proposal–acceptance or rejection, assessment–agreement or disagreement, and question–answer pairs. Insert-expansions

arise when the chatbot seeks clarification or proposes a candidate understanding; post-expansions often summarize an agreed next step, functioning as sequence-closing thirds.

Extract 3: proposal–acceptance with insert expansion

1 S3: maybe we meet at eleven?

2 AI: eleven a.m., right?

3 S3: yes (.) a.m.

4 AI: Okay (.) then we can book tickets for noon

Preference organization exhibits classic asymmetries. Preferred responses (acceptances, agreements) are brief and prompt, while dispreferreds show delay and mitigation. Because the chatbot's policy is to sustain progressivity, it frequently reframes rejections as softened alternatives, preserving forward movement.

Extract 4: dispreferred rejection with account

1 AI: Shall we invite more classmates?

2 (0.4)

3 S4: Well, maybe not (.) because it's difficult to organize a big group

Repair is a salient locus of interactional work. Learners engage in self-initiated self-repair near lexical access difficulties and following misrecognitions; the chatbot regularly initiates repair through open-class initiators ("sorry?") and candidate understandings that enable quick resolution. When ASR misrecognition persists, extended repair chains may appear, with measurable consequences for pause profiles.

Extract 5: other-initiated repair with candidate understanding

1 S5: we can go to the plan- pla[neta-]

2 AI: [planetarium?]

3 S5: yes (.) the planetarium (.) near the park

4 AI: got it (.) the planetarium near the central park

Discourse marking and hesitation resources support action projection and stance management. Turn-initial markers (well, so, okay) launch plans and proposals; stance markers (I think, maybe, actually) mitigate disagreement and calibrate epistemic stance. Filled pauses (uh/um) cluster at topic shifts and during planning portions of problem-solving tasks. The chatbot provides relatively sparse continuers (mm-hm, right); in their absence, learners often self-prompt to sustain extended runs.

Extract 6: stance and planning

1 S6: Okay, so I think first we should check prices (.) then we decide the time

Topic management is largely chatbot-led through agenda-setting, periodic summaries, and boundary markers that segment sub-tasks. Summary turns commonly operate as sequence-closing thirds and as pre-closings for activity phases, ratifying shared decisions and projecting transition.

Extract 7: summary as sequence-closing third

1 S7: Tickets are cheaper online

2 AI: good point (.) so, we meet at eleven and buy online tickets now?
3 S7: yes (.) let's do that

Closings typically feature canonical pre-closing tokens (so/okay/anyway) and are often initiated by the chatbot via a wrap-up plus an invitation to add final items; learners accept and frequently offer a brief positive assessment before terminal exchange.

Extract 8: pre-closing and terminal exchange
1 AI: okay (.) we made a plan (.) anything else to add?
2 (0.3)
3 S8: no, that's all
4 AI: great (.) talk to you next time
5 S8: thanks bye

Relative to matched peer dialogues, chatbot interactions exhibit fewer overlaps, longer TRP gaps traceable to processing latency, more proposal–acceptance cycles with summary thirds, and fewer listener continuers. These structural properties shape opportunities for production: candidate understandings and clarification inserts create additional, short turns for the learner, whereas scarce continuers can shorten runs unless learners self-prompt. The principal structural loci linking organization to fluency are turn-entry latency, the density of insert-expansions, and the distribution of repair. Shorter latencies and the presence of receipt tokens support longer mean length of run; candidate understandings can convert stalls into productive self-repair; repeated ASR-triggered repair inflates pause duration and depresses articulation rate. These observations motivate the quantitative modeling of turn-taking metrics and repair density alongside speech rate, mean length of run, and pause measures reported in the results. Quantitatively, learners exhibited modest but reliable fluency advantages in the chatbot condition. Mean speech rate was higher during learner–AI interaction (≈ 3.6 syll/s) than in peer conversation (≈ 3.3 syll/s), and mean length of run likewise increased (≈ 6.8 vs. ≈ 6.1 syllables between pauses > 250 ms). The frequency of long pauses (> 1 s) was reduced, while articulation rate remained stable across conditions, indicating that gains reflected smoother delivery rather than faster articulation alone. Repair density was higher in chatbot dialogue (≈ 6.2 vs. ≈ 4.9 repairs per 100 words), but most repairs were self-initiated and rapidly resolved, often triggered by recognition issues rather than communicative breakdown. Mixed-effects models confirmed significant condition effects for speech rate and mean length of run, with pause behavior and repair patterns conditioned by turn-entry latency and ASR performance.

Linguistic and pragmatic strategies learners employ when negotiating meaning in chatbot-mediated conversations.

Students employ a routine set of practices to resolve trouble, query common understanding, and facilitate task goals in chatbot-mediated dialogue. At the lexical–semantic level, circumlocution and lexical approximation are used where target vocabulary is absent; students specify characteristics or functions in a bid to obtain recognition, typically followed by quick confirmation and uptake. Paraphrasing and reformulation operate to maintain intersubjectivity

after alignment has been first achieved, while option-listing frames candidate sets for the chatbot to select from, lowering risk of misrecognition. Pragmatically, confirmation checks, clarification requests, and comprehension checks frame the negotiation sequence, typically bundled with stance-softening and hedging. Turn-holding resources—filled pauses, discourse markers, and incremental planning frames—allow learners to hold the floor through ASR/TTS latency and leave open a formulating development until an acceptable lexical resolution is reached. Repair routines combine these resources: self-initiated self-repair is the default option, with restarts and replacements closing most repair episodes; where required, learners solicit other-initiated repair from the chatbot through the use of open-class questions or by presenting partial forms for completion.

Extract 1: circumlocution and candidate uptake
1 S1: we can go to the place where you watch the stars inside
2 AI: the planetarium?
3 S1: yes the planetarium (.) near the park

In addition to circumlocution, learners frequently perform confirmation checks that test the chatbot's understanding of key parameters, especially numbers, times, and constraints. These checks often appear as increments to a prior proposal and are delivered with rising intonation to mark conditionality; once confirmed, learners proceed to elaborate the plan.

Extract 2: confirmation check and elaboration
1 S2: maybe we meet at eleven (.) eleven a.m. right?
2 AI: yes eleven a.m.
3 S2: okay then first we buy online tickets and after that we go

Clarification requests surface when the chatbot introduces unfamiliar lexicon or compressed proposals. Learners use open-class initiators ("sorry?"), wh-questions targeting a term or step, or meta-communicative prompts that ask the chatbot to slow down or rephrase. These moves reopen the sequence while preserving progressivity and often yield a simplified reissue or a candidate paraphrase that the learner can accept.

Extract 3: clarification request and simplified reissue
1 AI: we could take the shuttle to the venue
2 S3: sorry what is shuttle
3 AI: a small bus that takes people between places
4 S3: okay the small bus (.) we can take that

Negotiation also proceeds through option-listing and contrastive reformulation. Learners present minimal pairs or short sets to elicit a preference, thereby distributing the selection task to the chatbot and securing a next action without heavy lexical search.

Extract 4: option-listing and preference elicitation
1 S4: do you prefer morning or afternoon (.) because tickets are cheaper in morning
2 AI: morning works
3 S4: good then we choose the ten o'clock time

ISSN: 2229-9327

Hedging and stance management mitigate potential face-threat in dispreferred responses. Learners preface rejections with delay and softeners, then supply accounts; this preserves rapport and reduces the likelihood of adversarial sequences. The chatbot's candidate alternatives facilitate acceptance of the revised trajectory.

Extract 5: mitigated rejection with account
1 AI: shall we invite more classmates
2 (0.5)
3 S5: maybe not (.) I think it's hard to organize many people
4 AI: okay we keep it small then

Self-repair is the primary engine of form–meaning adjustment on the learner side. Restarts, replacements, and insertions cluster at lexical access points and after suspected misrecognitions; learners frequently align to the chatbot's candidate understanding and incorporate it into the developing turn, producing a quick closing of the repair sequence.

Extract 6: self-initiated repair with incorporation
1 S6: we should res— reserve the tick— tickets now
2 AI: reserve the tickets now
3 S6: yes reserve the tickets now online

Where lexical access fails, learners solicit collaborative completion, either explicitly asking for a word or offering an initial fragment that invites the chatbot's candidate completion. This strategy reduces stall time, limits long silent pauses, and converts potential breakdown into jointly produced progress.

Extract 7: collaborative completion
1 S7: how do you say when the train is not on time (.) it is
2 AI: delayed
3 S7: delayed yes (.) if it's delayed we need plan B

Discourse markers organize plans and keep turns open during multi-step reasoning. Turn-initial okay/so/well project action, while enumerations and "first–then–finally" frames allow the learner to hold the floor and display emerging structure even before all lexemes are available. This incrementalism supports longer runs and higher articulation time despite occasional hesitation.

Extract 8: incremental planning and turn-holding
1 S8: okay so first we check prices (.) then we choose time (.) and finally we buy

Learners also perform comprehension checks addressed to the chatbot to ensure downstream relevance. These checks appear as yes/no interrogatives or tag questions and function as sequence closers that ratify shared understanding before transition, thereby reducing backtracking later in the task.

Extract 9: comprehension check as sequence closer
1 S9: so we meet at the main gate at eleven (.) right
2 AI: right
3 S9: great then I will bring the student cards

ISSN: 2229-9327

Occasional L1 insertions appear as temporary scaffolds for unknown lexemes, typically followed by an immediate request for translation or a self-initiated replacement. In the voice interface these are short and functional, minimizing derailment while preserving task momentum.

Extract 10: L1 scaffold and replacement

1 S10: we should take the marshrutka (.) the small bus

2 AI: a minibus

3 S10: yes the minibus to the museum

Through these actions, meaning negotiation is directed towards progressivity and mutual intelligibility rather than exhaustive accuracy. Option-set and incremental planning strategies anticipating action, discourse marker and filler strategies winning the floor, and collaborative completion and simplified reissue strategies enlisting the chatbot's candidate understandings are associated in our corpus with longer mean length of run and fewer disruptive pauses. Recurrent misrecognition followed by retardant repair rather prolongs pause length and shortens runs. These results invite the inclusion of turn-holding and repair processes, along with speech rate and mean length of run, into quantitative models of fluency effects considered below.

Discussion

The study set out to determine whether and how chatbot-assisted interaction improves spoken fluency and to identify the interactional mechanisms underpinning any gains. The findings indicate that chatbot interaction supported spoken fluency by enabling extended production, reducing disruptive pausing, and transforming potential stalling into productive self-repair. Compared with peer talk, learners produced faster speech and longer uninterrupted runs, while articulation rate remained stable. Pauses shifted toward shorter durations with fewer long silences, and self-initiated repairs increased, mainly driven by lexical access and occasional ASR misrecognition. These repairs were typically brief and flow-preserving, enabling learners to maintain progressivity rather than lose the turn. Gains were most pronounced in information-gap and problem-solving tasks where clear next actions and system prompts scaffolded turn continuation; in opinion-exchange tasks, limited backchanneling constrained extended talk unless learners self-prompted with discourse markers. Turn-entry latency emerged as a key determinant: timely recognitional uptake and candidate completions facilitated fluency, while repeated ASR trouble prolonged pauses. Learners at the lower end of the B1–B2 band appeared to benefit slightly more, showing more frequent collaborative completions and reissues to sustain turns.

These results align with research demonstrating that dialogic AI can scaffold participation and maintain conversational momentum in L2 contexts (Yang et al., 2022; Tai & Chen, 2024; Wu et al., 2025). The observed increases in speech rate, run length, and efficient repair parallel findings from meta-level reviews reporting fluency benefits under well-designed conversational conditions (Du & Daniel, 2024; Koç & Savaş, 2025). The prominence of self-initiated repair and rapid closure is consistent with CA and HCI accounts of repair as a resource for managing human-machine dialogue and sustaining interaction (Croes et al., 2023; Galbraith & Gómez i Martínez, 2023; Pütz

& Esposito, 2024). Task-dependent patterns also reflect studies showing greater gains when chatbot turns provide clear action trajectories and structured scaffolding (Yang et al., 2022; Tai & Chen, 2024).

Learner perceptions converged with performance outcomes: chatbot practice felt useful, less anxiety-provoking, and conducive to willingness to communicate, echoing evidence that conversational AI can lower affective barriers and enhance WTC when cognitive load is manageable (Wang et al., 2024; Wu et al., 2025; Klímová & Ibna Seraj, 2023). However, reservations about discourse naturalness, backchannel scarcity, and occasional recognition errors match cautions regarding system latency, prosodic limitations, and design heterogeneity (Du & Daniel, 2024; Koç & Savaş, 2025; Pütz & Esposito, 2024).

Taken together, the results suggest that chatbot-mediated speaking yields modest but meaningful fluency gains when tasks offer clear next actions and system timing remains responsive. Optimal learning conditions involve minimizing latency, providing gentle clarification scaffolds, and pairing chatbot practice with human interaction to support pragmatic nuance and prosodic development (Kang & Sung, 2024; Wang et al., 2024; Wu et al., 2025).

Conclusion

This study shows that dialogue-based AI can support EFL spoken fluency when tasks and system performance are well aligned. Addressing the main aim (how learner–AI interaction shapes fluency and through which mechanisms) the findings indicate that learners spoke faster, produced longer uninterrupted runs, and paused less during chatbot sessions compared with peer talk. Increased self-repair, mostly self-initiated and quickly resolved around lexical access or misrecognition, helped maintain progressivity rather than disrupt speech. Benefits were strongest in information-gap and problem-solving tasks with precise next steps and chatbot clarification moves; gains were smaller in opinion-exchange exchanges where limited backchanneling shortened turns unless learners self-prompted.

CA evidence linked turn-entry timing, discourse-marker use, hesitation, and repair to fluency indices, showing how interactional resources converted potential stalls into productive sequences. Learners perceived chatbot practice as beneficial, realistic enough for rehearsal, and less anxiety-inducing than peer interaction, though they noted limits in naturalness, prosody, and occasional latency.

Pedagogically, chatbots fit well for warm-up and rehearsal phases, with human interaction still essential for pragmatic nuance and prosody. Future work should test longer-term effects, varied proficiency levels, and system design features such as backchannel density and repair affordances.

Funding: This research is not funded.
Acknowledgments: Not applicable.
Conflicts of Interest: The authors declare no conflict of interest.

Authenticity: This manuscript is an original work

Artificial Intelligence Statement: The authors confirm that all AI-generated content was carefully reviewed, edited, and integrated to ensure academic integrity, originality, and compliance with scholarly standards. During the preparation of this work, the authors used AI to check the grammar and translate several passages from Ukrainian into English. After using these tools, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

About the Authors:

Anna Matiienko-Silnytska, candidate of philological sciences, associate professor, associate professor of the Department of Theoretical and Applied Phonetics of the English language, Faculty of Romance and Germanic Philology, Odesa I.I. Mechnikov National University. Her research interests are Language Teaching and Learning, English dialectology, Sociolinguistics.
ORCID ID: https://orcid.org/0000-0001-5097-0302

Nana Mikava, candidate of philological sciences, associate professor, associate professor of the Department of Foreign Languages for Specific Purposes, Faculty of Romance and Germanic Philology, Odesa I.I. Mechnikov National University. Her research interests are Language Teaching and Learning, English for Specific Purposes, and Phonetics. E-mail: mikava.nani84@gmail.com ORCID ID: https://orcid.org/0000-0001-8526-7676

Iryna Savranchuk, candidate of philological sciences, associate professor, associate professor of the Department of Lexicology and stylistics of the English language, Faculty of Romance and Germanic Philology, Odesa I.I. Mechnikov National University. Her research interests are Language Teaching and Learning, English for Specific Purposes, Phonetics, Lexicology. E-mail: irinapopik77@gmail.com ORCID ID: https://orcid.org/0000-0001-6191-2298

Neonila Tkhor, candidate of philological sciences, associate professor, associate professor of the Department of Lexicology and stylistics of the English language, Faculty of Romance and Germanic Philology, Odesa I.I. Mechnikov National University. Her research interests are Language Teaching and Learning, English for Specific Purposes, Phonetics, Stylistics.
E-mail: departmentofforeignlanguages4@gmail.com ORCID ID: https://orcid.org/0000-0001-6191-2298

Hanna Poliakova, candidate of pedagogical science, assistant professor, Odesa National Maritime University, Department of Philology, Odesa, Ukraine. Scientific interests include: methodics of teaching foreign languages, written translation and interpretation.
departmentofforeignlanguages4@gmail.com ORCID https://orcid.org/0000-0001-6734-501X

References

Croes, E. A. J., Antheunis, M. L., Goudbeek, M. B., & Wildman, N. W. (2023). "I am in your computer while we talk to each other": A content analysis on the use of language-based strategies by humans and a social chatbot in initial human–chatbot interactions. *International Journal of Human–Computer Interaction, 39*(10), 2155–2173. https://doi.org/10.1080/10447318.2022.2075574

Du, J., & Daniel, B. K. (2024). Transforming language education: A systematic review of AI-powered chatbots for English as a foreign language speaking practice. *Computers and Education: Artificial Intelligence, 6*, 100230. https://doi.org/10.1016/j.caeai.2024.100230

Galbraith, M. C., & Gómez i Martínez, M. (2023). *An analysis of dialogue repair in virtual voice assistants* (arXiv:2307.07076). arXiv. https://arxiv.org/abs/2307.07076

Kang, S., & Sung, Y. (2024). EFL students' self-directed learning of conversation skills with AI chatbots: A case-study approach. *Language Learning & Technology, 28*(3), 1–22. https://doi.org/10.10125/73477

Klímová, B., & Ibna Seraj, P. M. (2023). The use of chatbots in university EFL settings: Research trends and pedagogical implications. *Frontiers in Psychology, 14*, 1131506. https://doi.org/10.3389/fpsyg.2023.1131506

Koç, F. Ş., & Savaş, P. (2025). The use of artificially intelligent chatbots in English language learning: A systematic meta-synthesis study of articles published between 2010 and 2024. *ReCALL, 37*(1), 4–21. https://doi.org/10.1017/S0958344024000168

Markus, M., Kirner-Ludwig, M. (2024). A Philologist's Perspective on Artificial Intelligence – a Case Study into *English Dialect Dictionary Online 4.0*, *International Journal of Lexicography*, *37*(2), 226–252 https://doi.org/10.1093/ijl/ecae001

Mieczkowski, H., Hancock, J. T., Naaman, M., Malte Jung, & Hohenstein, J. (2021). AImediated communication: Language use and interpersonal effects in a referential communication task. *Proceedings of the ACM on Human-Computer Interaction, 17*(1), 17:14. https://doi.org/10.1145/3449091

Patty, J. (2024). The use of AI in language learning: what you need to know. *Jurnal Review Pendidikan dan Pengajaran, 7*, 642-654. DOI:10.31004/jrpp.v7i1.24609

Pütz, O., & Esposito, E. (2024). Performance without understanding: How ChatGPT relies on humans to repair conversational trouble. *Discourse & Communication, 18*(6), 859–868. https://doi.org/10.1177/17504813241271492

Stanko, D. (2024). Integrating Artificial Intelligence in Naruto Fan Fiction Writing: a Case Study. *Arab World English Journal (AWEJ)* Special Issue on ChatGPT, 156-168. DOI: https://dx.doi.org/10.24093/awej/ChatGPT.10

Tai, T.-Y., & Chen, H. H.-J. (2024). Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions. *Computers & Education, 220*, 105112. https://doi.org/10.1016/j.compedu.2024.105112.

Wang, C., Zou, B., Du, Y., & Wang, Z. (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.

Wu, T.-T., Hapsari, I. P., & Huang, Y.-M. (2025). Effects of incorporating AI chatbots into think–pair–share activities on EFL speaking anxiety, language enjoyment, and speaking performance. *Computer Assisted Language Learning*, 1–39. https://doi.org/10.1080/09588221.2025.2478271.

Yang, H., Kim, H., Lee, J. H., & Shin, D. (2022). Implementation of an AI chatbot as an English conversation partner in EFL speaking classes. *ReCALL*, *34*(3), 327–343. https://doi.org/10.1017/S0958344022000039

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp.239-256
DOI: https://dx.doi.org/10.24093/awej/vol16no4.14

From Problems to Progress: Evaluating Dalton Plan-Based Instruction in Chinese College English Reading

Tianxiu Qi
School of Foreign Languages
Northwest Minzu University, Lanzhou, China
Corresponding Author: qtx077@163.com

Intan Safinas Mohd Ariff Albakri
Faculty of Languages and Communication
Universiti Pendidikan Sultan Idris, Perak, Malaysia
Email: safinas@fbk.upsi.edu.my

Mazlin Mohamed Mokhtar
Faculty of Languages and Communication
Universiti Pendidikan Sultan Idris, Perak, Malaysia
Email: mazlin@fbk.upsi.edu.my

Received:06/21/2025 Accepted:11/10/2025 Published:12/10/2025

Abstract
The Dalton Plan, an educational approach centered on student autonomy through assignment contracts and self-paced learning, offers a potential alternative to traditional teacher-center instruction. This study examines its application within the challenging context of Chinese College English (CE) reading courses, where students often struggle with academic reading comprehension and lack self-regulation skills. A quasi-experimental design was employed to evaluate the impact of Dalton Plan-based instruction on the reading performance and learner autonomy of 31 second-year medical students. Participants were divided into a control group (n=15) receiving conventional instruction and an experimental group (n=16) receiving Dalton Plan-based instruction, which emphasized independent reading assignments with tailored teacher guidance. A mixed-methods approach collected quantitative data (reading tests, Autonomous Learning Scale) and qualitative data (thematic analysis of student self-reports). Results indicated that Dalton Plan-based instruction significantly enhanced learner autonomy; however, its effect on reading test scores was not statistically significant. Thematic analysis of self-reports revealed a developmental trajectory from initial challenges—such as time management and vocabulary inadequacy—to later progress in reading interest, strategy use, and persistence. These findings not only provide pedagogical insights for innovating reading instruction in Chinese EFL (English as foreign language) contexts but also highlight the complex relationship between fostering autonomy and achieving measurable gains in reading proficiency.

***Keywords*:** College English reading, Dalton Plan-based instruction, English as a foreign language, learner autonomy, reading performance

How to cite: Qi, T., Albakri, T. S. M. A., & Mokhtar, M. M. (2025). From Problems to Progress: Evaluating Dalton Plan-Based Instruction in Chinese College English Reading. *Arab World English Journal, 16*(4):239-256. tps://dx.doi.org/10.24093/awej/vol16no4.14

Introduction

Studies in English as a second language (ESL) or / English as a foreign language (EFL) field have consistently highlighted the importance of integrating 21st-century skills into teaching and learning practices (Mokhtar, 2021; Tahir et al., 2020). Among these skills, learner autonomy has emerged as a critical focus in contemporary education. Additionally, assessing students' behaviors and individual areas of interest has been found to significantly contribute to the enhancement of their language proficiency (Tahir et al., 2020; Zulkepli, 2024). One notable instructional approach that supports learner autonomy is the Dalton Plan, which has been implemented in China to foster greater student independence and self-directed learning.

The Dalton Plan is designed to transform traditional classroom teaching. The fundamental distinction between traditional teaching and the Dalton Plan lies in the locus of control: while the former emphasizes direct instruction, the latter prioritizes student autonomy through self-directed assignments. Parkhurst (1922) argued that traditional teaching methods often rely excessively on teacher-led instruction, which may hinder the development of students' creativity, initiative, and sense of responsibility. Under the Dalton Plan, students and teachers establish a learning contract, granting students autonomy in making decisions regarding their learning process. Students determine what tasks to complete, where and when to complete them, and at what pace. The teacher's primary role is to track students' progress, provide individualized guidance, and facilitate collaborative discussions through classroom conferences. This shift from teacher-driven instruction to student-centred learning allows students to take ownership of their educational progress, fostering greater independence and accountability.

Comparative studies have demonstrated that the Dalton Plan is more effective than traditional instructional methods in promoting learner autonomy, self-regulation, and collaboration (Babuščáková, 2008; Moosavi, 2019). This effectiveness can be attributed to the Plan's emphasis on student initiative, which helps sustain prolonged motivation. According to Van der Ploeg (2013), this motivation largely stems from "the work as work" (p. 84). He elaborates that the assignment is perceived as "a real job" and experienced as "my work"—with an emphasis on "my"—which in turn fosters motivation, concentration, and discipline. Furthermore, Tian (2017) argued that implementing the Dalton Plan is especially pertinent in the" Internet +" era, as it encourages students to engage in self-directed learning while still benefiting from structured guidance.

The Dalton Plan has gained substantial traction in China's educational landscape since its introduction (Huang, 2022). Many schools nationwide have implemented Dalton-inspired educational reforms, leading to notable advancements (Tian, 2017). Zhang (2018) suggested that although the Dalton Plan has traditionally been applied in middle and high school settings, college students who exhibit greater psychological maturity stand to benefit even more from its emphasis on independent learning. The principles of freedom and cooperation embedded in the Dalton Plan align well with the broader goals of higher education, which aim to cultivate inquiry-driven, responsible individuals. One domain where learner autonomy is particularly critical is the College English (CE) course, a mandatory program for non-English major students in Chinese universities. The course, which spans two years (four semesters), aims to develop students' proficiency in listening, speaking, reading, writing, and translation. However, with limited instructional time (3-4 hours per week), achieving these objectives remains a challenge. Recognizing this limitation, *The College English Teaching Guide* (2020) has emphasized the importance of academic reading, particularly its integration with students' respective disciplines. The guide encourages students to

engage with English-language scholarly materials to enhance their expertise in their fields. Consequently, reading proficiency plays a crucial role in the CE curriculum.

A major challenge in the CE course is students' difficulty in academic reading comprehension. According to Cai (2016, 2017, 2018), many university graduates in China struggle to read English-language academic literature related to their disciplines. This issue suggests that reading comprehension remains a persistent challenge for CE students. Zhang and Seepho (2013) found that Chinese EFL learners often lack essential self-regulation skills, such as self-planning, self-monitoring, and self-evaluation, which impede their reading performance. Lian (2017) further noted that the traditional teacher-centred grammar-translation approach which focuses on sentence-by-sentence explanations tends to make students passive learners, ultimately diminishing their interest in reading. In contrast, reading is an inherently cognitive and interactive process that requires active engagement for optimal comprehension.

Given these challenges, this study investigates the effects of Dalton Plan-based classroom instruction on learner autonomy and reading performance in the CE course. Additionally, it explores the challenges students face when implementing Dalton Plan-based instruction and how they progress in reading under this instructional approach. The study aims to answer the following research questions:

1. Does Dalton Plan-based classroom instruction improve students' learner autonomy in reading in the CE course?
2. Does Dalton Plan-based classroom instruction enhance students' reading performance in the CE course?
3. What challenges and progress related to reading performance do students report under Dalton Plan-based classroom instruction?

Literature Review

The Dalton Plan, an educational approach developed by Helen Parkhurst in the early 20th century, is founded on the core principles of autonomy, responsibility, and collaboration. Its relevance endures in modern pedagogy, which prioritizes learner-centred models aimed at fostering independent thinking and self-regulation (Moškvanová, 2009). Research consistently highlights its strength in fostering learner autonomy, a critical foundation for successful language acquisition. Babuščáková's (2008) comparative study of English lessons in the Czech Republic found that the Dalton Plan's use of individualized work, pair and group tasks, and a shift in the teacher's role from "controller" to "facilitator" were instrumental in cultivating a learning environment where students took greater initiative and responsibility for their learning outcomes. This development of autonomy is closely linked to the improvement of metacognitive skills. A longitudinal study by Říčan and Lanková (2019) demonstrated that pupils in Czech Dalton schools showed statistically significant improvements not only in reading comprehension but also in metacognitive monitoring, indicating an enhanced ability to plan, monitor, and evaluate their own learning processes.

Beyond general autonomy, the plan has shown efficacy in developing specific language competencies. Geraskevich and Nurgalieva (2021) implemented the Dalton Plan in a Russian high school context, focusing on foreign language communicative competence. Their findings revealed that the plan's structure of self-paced assignments, coupled with reflective practices like self-analysis, led to significant student improvement in communication skills. Notably, students reported high levels of interest and enthusiasm, underscoring the method's positive impact on

motivation. Similarly, at the primary school level, Kwella (2020) integrated the principles of independence, responsibility, and cooperation into English lessons in Poland, observing marked improvements in students' self-esteem, confidence, and collaborative problem-solving abilities in the target language. However, the implementation of the Dalton Plan in language classrooms is not without its difficulties. A recurring theme across studies is the time-consuming nature of the method for both teachers in preparing individualized assignments and for students in managing their learning (Friedl, 2012; Geraskevich & Nurgalieva, 2021). Furthermore, the success of the approach can be variable. Friedl's (2012) investigation of a Dalton-inspired model (COOL) in Austrian vocational schools, while noting gains in student independence and cooperation, found that only 60% of students felt it directly helped their English learning, suggesting that the subjective experience of progress does not always align uniformly with pedagogical objectives.

The plan's versatility is evidenced by its successful adaptation beyond language learning. For instance, Yang et al. (2023) integrated it into high school aerobics using learning contracts to foster ownership and teamwork, while Stary & Weichhart (2017) applied its principles to digital design education, using "learning contracts" and feedback systems to scaffold self-regulation. A key contemporary development is the integration of the Dalton Plan with digital tools. Moosavi (2019) theorized that the plan provides a solid pedagogical framework for "Bring Your Own Device" (BYOD) initiatives. This has been operationalized in studies like that of Weichhart et al. (2018), who developed a digital learning environment featuring "Dalton Plan Editors" and "Feedback Graphs," thereby translating the plan's core mechanics of assignment management and progress tracking into a digital format to support self-regulated learning.

Despite this body of evidence, critical gaps remain, particularly concerning the Chinese tertiary EFL context. First, while studies like Říčan and Lanková's (2019) focus on reading in primary schools, there is a scarcity of research investigating the Dalton Plan's effect on the academic reading proficiency of university-level EFL learners. Second, the existing literature often notes challenges like student adaptation to self-direction, but few studies qualitatively explore the specific, evolving problems and progress learners report throughout the process. Finally, there is a lack of empirical investigation into how the plan functions within educational cultures characterized by more teacher-centred traditions, such as that of China. Therefore, this present study is designed to address these gaps by empirically evaluating Dalton Plan-based instruction in the Chinese College English reading course. It will quantitatively measure its effects on learner autonomy and reading performance, while qualitatively tracing the challenges and progress students experience. By doing so, it aims to provide a nuanced model of how autonomy-focused frameworks can be effectively adapted to a specific and widespread EFL context, moving from documented problems to tangible progress.

Methods

This study employed the quasi-experimental research design using quantitative and qualitative data collection approaches.

Participants

This study employed convenience sampling due to its practical accessibility to the researcher. This non-random sampling method involves selecting participants based on ease of access, geographic proximity, availability, or willingness to participate (Dörnyei & Griffee, 2007). The study focused on second-year medical students enrolled in two CE course classes organized by the university’s

Section of Academic Affairs. Class one, the control group, consisted of 15 students: nine majoring in Clinical Medicine and six in Nursing. Class two, the experimental group, comprised 16 students: nine majoring in Stomatology and seven in Medical Laboratory Technology.It has to be pointed out that initially each class had about 60 students, but many were exempted from attendance and exams after passing CET4 with scores of 425 or above, therefore, the participants involved in the study were those who had not passed CET4 and had a relatively lower language proficiency compared with their peers who had passed CET4 already.

Intervention Plan

This study implemented two distinct instructional approaches over a three-month period: Dalton Plan-based instruction for the experimental group and traditional classroom instruction for the control group.

Dalton Plan-Based Instruction (Experimental Group)

The experimental group followed a Dalton Plan-based classroom approach, comprising three key stages. First, students, in consultation with the researcher, identified their reading objectives to ensure alignment with their individual learning needs. Second, students and the researcher collaboratively selected monthly reading assignments from three available options (maximum, medium, or minimum workload), allowing for personalized workload adjustments. Once finalized, students signed assignment contracts and completed their reading tasks both during and outside class hours. Lastly, students worked on their reading assignments with instructional support provided through classroom sessions and individual guidance.

Teacher instruction was delivered in two formats. Classroom sessions focused on developing reading skills and analyzing complex sentences within reading materials. In contrast, individual guidance was conducted outside class time to address non-reading-related challenges such as time management and emotional regulation, based on students' weekly self-reports submitted online. Each week, students allocated approximately one-third of their three weekly 45-minute sessions to teacher-led instruction, while the remaining two-thirds were dedicated to independent reading practice. Adjustments were made based on students' self-reported challenges, ensuring that instructional support was targeted and responsive to their needs.

Traditional Classroom Instruction (Control Group)

The control group followed a conventional teacher-centred approach, which differed in several key aspects. Time allocation was a significant distinction. While the experimental group spent most of their time on independent reading, the control group followed a lecture-based model, with two-thirds of class time dedicated to teacher-led instruction and only one-third allocated for individual reading practice. During their reading practice sessions, most students focused on materials related to College English Test Four (CET4) preparation.

Another key difference was the instructional approach. Unlike the Dalton Plan-based model, which was tailored to students' self-reported needs, the control group followed a standardized curriculum based on textbook content and teacher discretion. Instruction primarily focused on analyzing text structure, genre, and main ideas, explaining difficult vocabulary and sentence structures, completing textbook exercises, and providing general guidance on reading strategies and proficiency assessment.

Research Instruments

This study utilized three main types of reading materials: course book texts from *New Progressive College English Integrated Course (Book Four)*, CET4-related reading materials, and students' self-selected materials, which included both leisure reading and discipline-specific literature relevant to their majors. The course book materials comprised 18 reading passages organized into six units, with each unit containing three passages centred on the same topic. These passages were designated for intensive reading, requiring students to thoroughly comprehend the text, including vocabulary, sentence structures, and accompanying exercises. Intensive reading has been shown to enhance student engagement, motivation, and vocabulary acquisition through contextual memorization (Andrés, 2020).

The CET4-related reading materials consisted of reading comprehension passages sourced from authentic CET4 test papers and model tests. These passages were categorized into blank-filling, paragraph-matching, and multiple-choice questions of varying lengths. They were assigned as extensive reading tasks, designed to help students develop an overall understanding of the main ideas rather than requiring a detailed analysis. After reading, students completed exercises to reinforce comprehension. Extensive reading has been found to counteract passive learning tendencies commonly associated with traditional classroom instruction (Mede, 2013).

The students' self-selected materials included two categories: leisure reading materials and academic literature related to their fields of study. Leisure reading materials were accessible through the WeLearn online teaching platform, which supports CE courses. CE teachers could create online classes corresponding to the course books used in their lessons. In this study, the researcher provided reading passages through the platform's teacher interface, allowing both experimental and control groups to select and read them.

For discipline-specific literature, students independently selected reading materials, often using academic databases like PubMed, as recommended by their professional course instructors. These materials were also assigned as extensive reading tasks, aiming to cultivate students' interest in voluntary reading. Prior research by Fatimah (2019) highlights that extensive reading (ER) is closely linked to "pleasure reading" or "voluntary reading", emphasizing that students engage in extensive reading only when genuinely interested in the material.

Data Collection Methods

Autonomous Learning Scale

The Autonomous Learning Scale (ALS), developed by Macaskill and Taylor (2010), was administered online to both the control and experimental classes before and after the experiment. The purpose was to determine if there were differences in learner autonomy between the two groups. This scale consists of 12 items, each rated on a 5-point Likert scale ranging from one (very unlike me) to five (very like me). Higher scores indicate higher levels of autonomy in learning. It is important to note that all items, except for item 2, require a positive response. For item 2 ("I frequently find excuses for not getting down to work."), a negative response indicates higher autonomy. To ensure consistency in responses, the researchers reversed the scoring for item 2 using the data transformation function in the SPSS statistical package. The survey was conducted during class sessions, with an emphasis on the confidentiality of responses and their irrelevance to students' CE examination scores to encourage authentic responses from the students.

Reading test paper

The reading test, developed by the researcher, featured passages from the WeLearn online platform for CE courses, with a difficulty level comparable to CET4. The test included four passages: a vocabulary-focused passage (200-250 words) with ten fill-in-the-blank items; a long passage (1,000 words) requiring sentence-to-paragraph matching for main ideas; and two shorter passages (300-350 words each) with five comprehension questions each. Both pre- and post-tests were identical, with minor adjustments to question order in the post-test to avoid recall bias. Reliability analysis using Cronbach's Alpha yielded a coefficient of 0.84, indicating acceptable internal consistency. The tests were administered to control and experimental groups before and after the intervention, each lasting 40 minutes during a class session, with printed papers and answer sheets provided. Scores were calculated after both tests to evaluate reading progress.

Self-reports

The experimental class students' self-reports detailing their feelings and challenges in completing the reading assignments were collected. The self-report format was adapted from progress graphs used by Parkhurst (2005). It comprised three sections. The first section recorded basic information such as the student's name, major, and type of assignment contract. The second section tracked the percentage of the reading assignment completed each week. The third section invited students to describe any challenges and emotions experienced while working on the reading assignment; notably, the word "feelings" was used instead of "progress" to avoid prompting students to report progress directly. This section resembled open-ended questions, allowing students to express their thoughts freely. If students had nothing to share, they had the option to leave this part blank. Since the study spanned three months, students submitted weekly self-reports, resulting in a total of 12 reports per student. Data involved in this study was students' report of their feelings and problems.

Results

Autonomous Learning Scale

Table 1. *Comparison between two classes on learner autonomy*

		Pre			Post		
Group	N	M	SD	Sig(two-tailed)	M	SD	Sig(two-tailed)
Experimental Group	16	38.69	3,807	.642	42.06	5.792	.043
Control Group	15	37.87	3.255	(P>0.05)	39.60	3.225	(P<0.05)

Table 1 presents the results of the ALS (Autonomy of Learning Scale) for both the experimental and control groups. Before the experiment, the mean pre-test scores for the experimental group were 38.69 and for the control group were 37.87. After the experiment, the mean post-test scores increased to 42.06 for the experimental group and 39.60 for the control group. This indicates that both groups showed improvement in learner autonomy from before to after the experiment.

Initially, there was no statistically significant difference between the two groups as determined by the pre-test ALS ($p = 0.642$, which is greater than 0.05). However, three months later, the post-test ALS showed a significant difference ($p = 0.043$, $p < 0.05$), indicating that students who received Dalton Plan-based instruction demonstrated significantly greater improvement in learner autonomy compared to those in conventional classroom instruction.

Reading Test

Table 2. *Comparison between two classes on reading scores*

		Pre			Post		
Group	N	M	SD	Sig(two-tailed)	M	SD	Sig(two-tailed)
Experimental Group	16	140.80	19.821	.691	145.93	14.119	.0712
Control Group	15	138.00	18.389	(P>0.05)	143.33	23.011	(P>0.05)

Table 2 shows the descriptive and inferred results of a reading test conducted on both the experimental and control groups. Before the experiment began, the average pre-test scores were 140.80 for the experimental group and 138.00 for the control group. After the experiment concluded, the average post-test scores rose to 145.93 for the experimental group and 143.33 for the control group. This indicates that both groups demonstrated improvement in their reading test scores from before to after the experiment.

Prior to the experiment, there was no statistically significant difference between the two groups based on their pre-test reading scores ($p = 0.691$, which is greater than 0.05). Similarly, three months later, the post-test reading scores also showed no significant difference ($p = 0.0712$, $p>0.05$). This suggests that students who received the Dalton Plan-based classroom instruction did not exhibit significantly greater improvement in reading proficiency compared to those taught using conventional classroom methods.

Self-Reported Problems

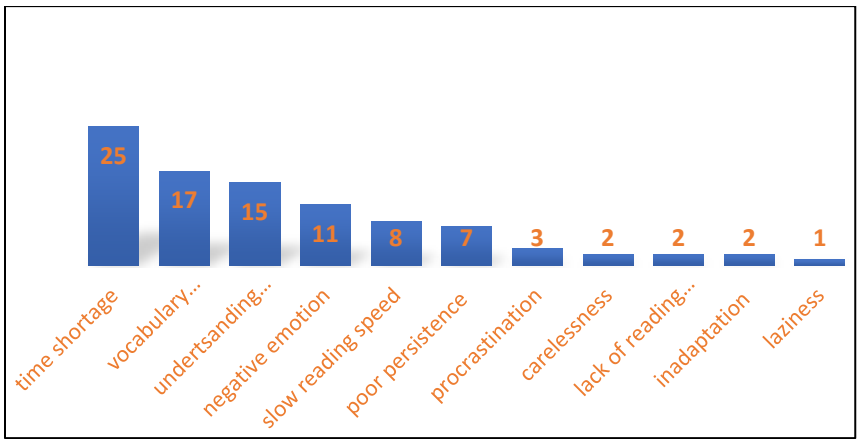


Figure 1. Problems and frequency

Figure 1 above illustrates that a total of 11 types of problems were reported across 12 instances of self-report. These issues, listed in order of frequency from highest to lowest, include time shortage, vocabulary inadequacy, understanding difficulty, negative emotions, slow reading speed, poor persistence, procrastination, carelessness, lack of reading strategies, inadaptation to new learning modes, and laziness.

Time Shortage

The issue of time shortage was the most reported problem. Almost all participants (except for S9 and S13) mentioned that they did not have enough time to complete their reading assignments. This issue was largely influenced by the students' majors; all participants in the study

were majoring in medicine, which meant they had more classes and homework than other students. Below are some excerpts from students:
"*At the beginning of the plan, since I myself had to arrange the assignment, inefficiency and inaccuracy appeared, and sometimes I felt time was not enough to finish the assignment.*" (S3)
"*Because of the mid-term examination, I do not have enough time to review all subjects, so I have to give up English reading most of time, in the next following two weeks, I will make up for what I have missed.*" (S6)

Vocabulary Inadequacy

Vocabulary presents the most significant challenge for foreign language learners worldwide (Hellekjær, 2009). In this study, insufficient vocabulary was identified as the second most prevalent issue. A total of 11 students (S1, S2, S4, S6, S7, S8, S10, S13, S14, S15, S16) mentioned that their vocabulary was inadequate for completing reading assignments efficiently and effectively.

"*I need to repeatedly memorize the new words again and again, still I forget many of them, so I cannot understand the reading passages well and sometimes it takes much time to finish one.*" (S8)
"*In the aspect of reading, I have a problem with words. If there are too many new words in a passage and I cannot understand most part of it, I just give it up.*" (S7)

Understanding Difficulty

The issue with the third highest percentage was inefficiency and inaccuracy in understanding texts. In total, 12 students (S1, S2, S3, S5, S6, S7, S8, S9, S10, S13, S14, S15) in this study reported difficulty in understanding, mainly due to long and complex sentences and paragraphs. The following excerpts are from students' reports.
"*I can understand the main idea of the passage, but when I do the exercises after the passage, I make mistakes, that means I do not have a complete understanding of it.*" (S10)
"*In reading, I do not understand some long and difficult sentences, especially when there are unknown words in in them, which makes it harder to understand.*" (S5)

Negative Emotion

Various negative emotions related to reading were reported by students in the study, with impatience and depression being the most common. Students typically felt these negative emotions when they struggled to comprehend what they were reading, particularly in long passages containing many unfamiliar words. Below are some excerpts from students:

"*I feel impatient when I read, the reason for this is that I do not understand the meaning since there are many new words, then I feel panic.*" (S15)
"*When I read the long passages, I feel unwell and confused, I just couldn't concentrate on it.*" (S6)
"*This week, I felt especially restless and couldn't calm down, so my reading plan didn't go well. I'll make adjustments as soon as possible next week.*" (S8)

Slow Reading Speed

Reading speed is considered one of the indicators of reading proficiency (Khasawneh, 2021). It was commonly reported that it usually took them much more time than expected to complete a passage. Below are some responses from students:
"*I do not finish the assignment this week because I read very slowly due to many unknown words in it.*" (S4)
"*My problem in reading is that it takes me more time compared with other students, I cannot understand the meaning of the whole passage that quickly.*" (S2)

Poor Persistence

Seven students (S1, S2, S6, S7, S10, S12, S16) reported experiencing poor persistence during their reading process in accomplishing the reading assignment. According to their feedback, this issue did not arise when they read material they could easily comprehend without much additional effort. However, when faced with longer and more complex passages, they reported lacking the determination and patience needed to overcome obstacles and fully understand the content. The following are some excerpts from the students:

"*In accomplishing the reading assignment, I would like to memorize as many new words as I can, it's convenient and easy to do, but as for reading, I have no patience to continue since I make many mistakes in understanding.*" (S9)
"*Sometimes I just give up reading long and difficult passages since it's time-consuming, usually two or three hours to finish one passage. I know I do not have the patience to analyse them carefully.*" (S12)

Other problems

Procrastination (S5, S13, S16), carelessness (S1, S7), lack of reading strategy (S10, S14), inadaptation to the new learning mode (S5, S7), and laziness (S16) are other problems reported with low frequency in the study. Below are some responses from the students:

Progress

Figure 2 illustrates sixteen aspects of progress reported by students in their self-reports conducted 12 times. Among these aspects, the first nine improvements are considered common due to their relatively higher percentage of reporting. In descending order of frequency, these common improvements are increased reading interest, understanding improvement, cultivation of reading habits, improvement in using strategies, better emotion control, increased reading speed, more persistence, vocabulary enhancement, timely completion of assignments, and better time management. The remaining aspects of improvement such as enhancement of language sense, increased self-discipline, and increased confidence, were reported less frequently.

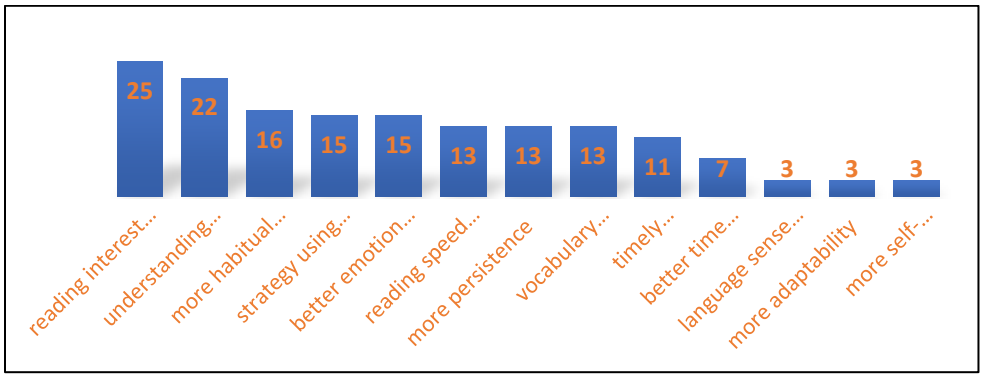


Figure 2. Progress and frequency

Reading Interest Increase

All students except for one (S6) expressed their interest in reading English texts increased. Students' interest in reading increased mainly by the sense of achievement in completion of reading assignment given in the assignment contract. The sense of achievement strengthened reading pleasure. The following are some excerpts from students:

"*……, practices make perfect, I find the correct way in reading just by reading a lot, when I verify the method by doing the exercises and find the accuracy improved, I feel a sense of achievement, and I can't help doing more reading, this is a very interesting process.*" (S3)

"*Since the college entrance examination, I have had a negative attitude to learning English, the reason for this was that I had been bored and tired of being driven by teacher every day in high school. But this time the teacher arranged us to read on our own independently, I was not used to it at the first, but after trying and trying, I began to have a little interest, although I may make many mistakes in reading exercises after the passage, I am relieved to have some improvements.*" (S7)

Understanding Improvement

All participants reported this aspect of improvement. Among them, four students (S2, S5, S10) reported twice, while one student(S11) reported four times. Most accounts of this progress were too general to pinpoint the specific aspect. Many students expressed improvements in aspects such as "reading" and" understanding," or "become smoother in reading." Below are some excerpts from students:

"*There has been progress compared to the beginning stage, and I'm gradually able to understand the reading passages. Unknown words are having less of an impact on me, and my accuracy in answering questions has also improved.*" (S15)

"*Reading is becoming smoother, and without looking up words, I can roughly understand the meaning of complex and difficult passages.*" (S5)

More Habitual Reading

The second most reported improvement was the cultivation of more habitual reading habits, mentioned by 14 students (S1, S2, S3, S5, S6, S7, S8, S9, S10, S11, S12, S13, S14) in their self-reports. From the self-reports, it is evident that many students tried to read daily. Below are some excerpts from students:

"*Reading is becoming more regular. I strive to do a reading passage every day. Building up vocabulary is important, and repeatedly memorizing words greatly helps in answering reading questions.*" (S9)
"*This new learning process has guided us in developing good reading habits and improving our autonomy. In the midst of our busy daily lives, English reading has become a way to take a little break.*" (S1)

Improved Reading Strategy

Almost all students (except for S9 and S15) expressed improvement in their reading strategies. Among them, two students (S3 and S12) reported improvements twice. These improvements are primarily evident in strategies such as skimming for the main idea, controlling reading speed, analyzing the structure of complex sentences. Below are some responses from the students' self-reports:
"*I use skimming more often than before; it is a very good way to make yourself concentrated on the main idea and ignore the unknown words.*" (S11).
"*Now I am able to guess the meaning of the new words in the passages, not like before, I looked up each unknown word I met.*" (S8)
"*One important strategy I have learned is that the analysis of complex sentences. If I can know the structure of these sentences, I can understand their meaning no matter how complicated they seem."* (S13)

Better Emotion Control

There was a gradual process, as many students initially displayed impatience and restlessness. However, as the experiment progressed, students reported varying degrees of progress in overcoming this issue. A total of 10 students (S2, S3, S5, S6, S7, S9, S10, S13, S14, S15) reported this improvement. The time it took for them to make some improvements ranged from one week to four weeks (one month). The following are excerpts from some students:
"*Now I can calm down when I am reading and I read very carefully, I try to understand the meaning inside the articles, and also with a sense of pleasure.*" (S2)
"*Generally, I have made progress, now I can settle down more quickly to reading*." (S15)

Reading Speed Improvement

Ten students (S1, S2, S3, S5, S6, S9, S10, S12, S15, S16) mentioned this improvement in reading speed in their self-reports. This study found that reading speed was closely related to the vocabulary students mastered, the type of passages, and the structure of the passages. Below are some excerpts from the students:
"*My reading speed is much faster than before, and I take an interest in reading (major-related) literature, but the present state does not reach my expectations, it's somewhat confusing. I will keep trying with harder work.*" (S3)
"*I feel my vocabulary expanded to some extent, and my reading speed improved accordingly, slowly the accuracy rate in understanding was enhanced.*" (S10)

More persistence

A total of 11 students (S1, S2, S3, S5, S6, S7, S11, S12, S14, S15, S16) mentioned improvements in this area. The enhanced persistence was primarily evident in the aspect of

continuing to read even when faced with difficulties. Below are some excerpts from the students' responses:
"*Now I can go on reading with the unknown words in the passages, they (the unknown words) do not affect me too much, not as they did before, and I can read passages of more genres than before.*" (S14)
"*I try not to read the Chinses translation given in the answer sheet, even there is much difficulty in understanding the meaning. I look up every unknown word bit by bit, trying to catch the meaning with my efforts.*" (S16)

Vocabulary Increases

The improvement in vocabulary was reported by 11 students (S1, S2, S4, S6, S8, S9, S11, S12, S13, S14, S15). One of the components of the assignment contract was the vocabulary memorization task, which was included upon the students' request. Most of the students had some form of vocabulary memorization plan before the experiment, but some may not have practiced it effectively in the past. The following are some responses from the students:
"*This month I have made relatively big improvement in my vocabulary quantity, I recite 25 new words each day in good time, I also have made slight improvement in reading speed.*" (S16)
"*I have improved in vocabulary accumulation, which is very important to reading comprehension, and I have developed a little language instinct in reading.*" (S9)

Timely Completion of Assignment

A total of 11 students (S1, S2, S3, S4, S6, S7, S9, S10, S12, S14, S15) reported this` improvement. Most students noted this improvement in comparison to the initial stage when they struggled with the assignments. Below are some responses from the students:
"*This week the undertaking of the assignment is relatively smooth, sometimes when I do not finish the amount of reading I assigned myself, I would add the amount the next day, trying to reach the goal required.*" (S3)
"*The general completion status is good, and I have gradually developed the habit of completing the reading assignment on time.*" (S14)

Better Time Management

Altogether six students (S1, S5, S8, S10, S12, S15) mentioned this improvement. Efficiency and flexibility in management of time are commonly reported by students. One of the essences of the Dalton Plan which distinguishes it from other teaching forms is the students' budget of time. Students had to use their time judiciously to accomplish their assignments on time under the Dalton Plan-based instruction. The following are some excerpts from students' response.
"*Sometimes when time is ample and I am in a good state, I can do more reading, I even do the task set for next week in advance.*" (S1)
"*Slowly I can make use of some fragmented time to do reading, not like the before, I had to have some time set specially for reading. And the amount of reading I have finished in one week is larger than before.*" (S5)

Other Progress

Some additional improvements were observed, though reported with lower frequency. These include enhanced language sense, increased adaptability, and greater self-discipline. Specifically,

ISSN: 2229-9327

the enhancement of language sense was noted by three students (S2, S8, S9), adaptability was mentioned by three students (S1, S4, S11), and increased self-discipline was reported by another three students (S3, S12, S14).

Discussion

The implementation of Dalton Plan-based classroom instruction has demonstrated a significant positive impact on learner autonomy in reading among CE students. However, despite these improvements in autonomy, the anticipated gains in reading proficiency were not achieved. Early in the program, students reported common challenges, such as time constraints, vocabulary limitations, difficulties in comprehension, negative emotions, slow reading speed, and low persistence. In the later stages, students self-reported notable progress in overcoming these challenges, including increased interest in reading, enhanced comprehension, development of reading habits, better strategy use, improved emotional regulation, faster reading speeds, and more persistence. This shift from initial "problems" to later "progress" reflects a pattern of gradual improvement in areas where students initially struggled, supporting a "from problems to progress" model of development.

Interestingly, while students' self-reports reflected these subjective improvements, they did not correspond to gains on standardized reading assessments. This discrepancy suggests that while students felt more engaged and autonomous in their reading practices, these subjective improvements did not directly translate to measurable gains in reading proficiency as assessed by standardized tests. A possible reason for this disconnect may lie in the differences between self-perceived progress and skills that are quantifiable through traditional assessment methods, which may not fully capture the nuanced, qualitative aspects of autonomy and engagement developed under the Dalton Plan-based classroom instruction. These tests may thus overlook the more gradual, foundational shifts in students' reading habits, persistence, and strategy use.

The observed increase in learner autonomy aligns with Babuščáková's (2008) research, which found that Dalton Plan-based lessons enhance autonomy more effectively than non-Dalton methods. This may be attributed to the Dalton Plan's emphasis on assignment contracts, which place responsibility on students, encouraging them to manage their learning independently. Responsibility is widely regarded as a foundational element of learner autonomy and is essential for successful language acquisition (Benson & Voller, 2014; Holec, 1981; Myartawan & Latief, 2013). Under this model, teachers guide progress and provide feedback rather than dominate instruction, giving students time and space for independent learning, the teachers' less dominant role might contribute to the improvements in learner autonomy. In Parkhurst's words, "freedom and responsibility together perform the miracle"(Parkhurst, 1922, p.29).This study adds to the body of evidence supporting the adaptation of the Dalton Plan in college settings, where the maturity of students could make them more receptive to learner-centred approaches (Zhang, 2018). In this study, autonomy was fostered through choices in reading assignments, flexible completion rates, and the inclusion of extracurricular materials, illustrating the potential of the Dalton Plan to promote autonomy in college students.

The lack of improvement in reading proficiency could be explained by two main factors. First, the participants were students who had not passed the CET-4 exam, indicating relatively low language proficiency. Prior research (Qi et al., 2024) suggests that the Dalton Plan may be more effective for students with higher language proficiency, who may better utilize autonomous learning techniques to improve their reading skills. Second, although essential Dalton Plan

elements like learning contracts and teacher-guided progress tracking were integrated, traditional classroom schedules were retained, limiting flexibility in time management. This rigid structure may have restricted the benefits of independent reading activities, contributing to the lack of improvement in reading test scores.

The "problems to progress" pattern observed in students' self-reports highlights that problem-targeted instruction effectively addressed specific challenges. For instance, impatience and restlessness were frequent issues during the early phase, particularly when students encountered long, difficult passages. Over time, students reported progress in managing these emotions. They were encouraged to persevere even when confronted with unfamiliar words and advised to avoid looking up each unknown word immediately. Instead, they were taught to infer meanings from context, an approach modeled during instruction. This strategy helped students focus on understanding the overall text rather than becoming fixated on individual words. Additionally, the Dalton Plan-based self-report mechanism encouraged students to reflect on their learning processes and identify specific areas for improvement. Recognizing one's own challenges is a critical step toward addressing them, as self-reflection has been shown to foster strategic learning and self-regulation (Schunk & Zimmerman, 2008).

The mismatch between students' self-reported improvements and their reading test results suggests that while the Dalton Plan-based instruction effectively fosters a sense of autonomy and engagement, there is gap between subjective improvements and actual progress. Research (Dunning et al., 2003; Schunk & Zimmerman, 2008) indicates these subjective improvements may not directly translate into quantifiable gains on standardized assessments. Several additional factors may have contributed to this disconnect. Students reported consistent challenges with time management and vocabulary, which, despite some progress, likely affected their reading test performance. The demanding course load for medical majors may have limited the time students could dedicate to reading, heightening the sense of difficulty with complex texts. Vocabulary, a fundamental component of reading comprehension, was frequently mentioned as a barrier throughout the experiment, which would have further hindered test performance.

Conclusion

In summary, the Dalton Plan-based classroom instruction appears to foster a more autonomous learning environment that empowers students to take responsibility for their reading habits and promotes positive attitudes toward learning. However, while qualitative data suggest significant gains in motivation and engagement, the quantitative data from reading tests do not reflect similar improvements. Future studies may involve the development of alternative assessments that more accurately capture the skills fostered through Dalton Plan-based classroom instruction. Additionally, integrating focused strategies for vocabulary development and comprehension support could further enhance students' reading proficiency. Ultimately, this study contributes to the understanding of how learner autonomy can be cultivated in EFL contexts and underscores the complexities of translating such autonomy into measurable academic outcomes.

Funding: This research is not funded.
Acknowledgments: No applicable.
Conflicts of Interest: The authors declare no conflict of interest.
Authenticity: This manuscript is an original work.

Artificial Intelligence Statement: The authors used DeepSeek to check and correct some possible mistakes to assist with improving language clarity and readability. The authors took full responsibility for the content of the publication.

About the Author:
Tianxiu Qi is a lecturer at the School of Foreign Languages, Northwest Minzu University, Lanzhou, China. She is also a Ph.D. candidate in TESL in Universiti Pendidikan Sultan Idris (UPSI), Malaysia. Her research focuses on studies on Instruction and Educational Pedagogy in reading in EFL/ESL field. ORCID ID: https://orcid.org/0000-0002-5724-4677

Dr.Intan Safinas Mohd Ariff Albakri is an Associate Professor at the Faculty of Languages and Communication, Universiti Pendidikan Sultan Idris (UPSI), Malaysia. Her areas of specialization are Course and Syllabus Design, Educational Technology, Materials Development in ELT and Teacher cognition. ORCID ID: https://orcid.org/0000-0001-5499-1709

Dr. Mazlin Mohamed Mokhtar is a senior lecturer at the Faculty of Languages and Communication, Universiti Pendidikan Sultan Idris (UPSI), Malaysia. Her main research interests lie in the areas of Education, TESL, Sociolinguistics, and ICT in Education.
ORCID ID: https://orcid.org/0000-0002-8231-2678

References

Andrés, I. C. (2020). Enhancing Reading Comprehension through an Intensive Reading Approach. *HOW*, *27*(1), 69–82. https://doi.org/10.19183/how.27.1.518

Babuščáková, D. (2008). The Dalton Plan in ELT: the role of teacher and the role of student.

Benson, P., & Voller, P. (2014). Introduction: Autonomy and independence in language learning. In *Autonomy and independence in language learning* (pp. 1-12). Routledge.

Cai, J. G. (2016). A study of the efficiency of College English teaching from the perspective of the economics of language. *Fudan Education Forum 2016*, *14*, 86–92. https://doi.org/10.3969/j.issn.1672-0059.2016.02.013

Cai, J. G. (2017). Review of Chinese tertiary English education: Failure and lessons. *Journal of Northeast Normal University* (*Philosophy and Social Sciences*), 1–7.

Cai, J. G. (2018). Teaching English as a second language in the China EFL context: An exploration of a paradigm shift. *Journal of Northeast Normal University* (*Philosophy and Social Sciences*), *295*, 12–17.

Dörnyei, Z., & Griffee, D. T. (2007). Research methods in applied linguistics. Oxford university press.

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, *12*(3), 83–87. https://doi.org/10.1111/1467-8721.01235

Fatimah, A. S. (2019). Portraying Learner's Autonomy in Extensive Reading Classroom. *OKARA: Jurnal Bahasa Dan Sastra*, *13*(1), 35–46. https://doi.org/10.19105/ojbs.v13i1.2228

Foreign Language Teaching Advisory Committee of the Ministry of Education. (2020). The College English Teaching Guide. Beijing: Higher Education Press.

Friedl, A. (2012). EFL between tradition and innovation: focus on COOL – cooperative open learning. In *Conference Proceedings: The Future of Education* (2nd ed.)

Geraskevich, N., & Nurgalieva, V. (2021). Dalton-Plan Technology in the Structure of Development of Foreign Language Communicative Competence Among Senior School Children. *IOSR Journal of Research & Method in Education (IOSR-JRME)*, *11*(2), 67–74.

Hellekjær, G. O. (2009). Academic English reading proficiency at the university level: A Norwegian case study. *Reading in a Foreign Language*, *21*(2), 198–222.

Holec, H. (1981). Autonomy and foreign language learning. Oxford: Pergamon Press.

Huang, Z. (2022). Introduction to 'The Shanghai Model.' *Educational Philosophy and Theory*, *54*(10), 1560–1562.

Khasawneh, M. A. S. (2021). The use of reading speed strategy in promoting reading comprehension among EFL students with learning disabilities. *Al-Lisan: Jurnal Bahasa (e-Journal)*, *6*(2), 225-235. https://doi.org/10.30603/al.v7i2.2135

Kwella, M. N. (2020). Teaching English in the Dalton Way, i. e. the Dalton Plan in English Lessons. *Elementary Education in Theory and Practice*, *15*(1(55)), 117–130. https://doi.org/10.35765/eetp.2020.1555.07

Lian, Y. (2017). Problems and strategies in reading in CE course. Think Tank Era.

Macaskill, A., & Taylor, E. (2010). The development of a brief measure of learner autonomy in university students. *Studies in Higher Education*, *35*(3), 351–359. https://doi.org/10.1080/03075070903502703

Mede, E., İnceçay, G., & İnceçay, V. (2013). Fostering Learner Autonomy through Extensive Reading: The Case of Oral Book Reports. *ELT Research Journal*, *2*(1), 16–25.

Mokhtar, M. M., Hashim, H., Khalid, P.Z.M., Albakri, I.S. M. S., & Jobar, N. A. (2021). A comparative study of boosters between genders in the introduction section. *Arab World English Journal*, *12* (1) 515-526. https://doi.org/10.24093/awej/vol12no1.33

Moosavi, Z. (2019). 21st century education on the Dalton Plan. *International Journal of Innovation and Research in Educational Sciences*, *6*(4), 546–551.

Moškvanová, V. (2009). *Dalton Elements in English Language Teaching* (Doctoral dissertation, Masaryk University, Faculty of Education).

Myartawan, I. P. N. W., & Latief, M. A. (2013). The correlation between learner autonomy and English proficiency of Indonesian EFL college learners. *Teflin Journal*, *24*(1), 63–81.

Parkhurst, H. (1922). Education On Dalton Plan. E.P. DUTTON.

Parkhurst, H. (2005). Education on the Dalton Plan (Chen.Jin-fang, & Zhao.Yulin, Trans) (2nd ed.). Peiking University Press.

Qi, T., Mohd Ariff Albakri, I. S., & Mokhtar, M. M. (2024). The role of the Dalton Plan-based instruction in enhancing learner autonomy and reading proficiency for Chinese college English students. *English Language Teaching*, *17*(8), 52–64. https://doi.org/10.5539/elt.v17n8p52

Říčan, J., & Lanková, B. (2019). Dalton Plan Pupils' Progress in Reading Comprehension and Metacognitive Monitoring: Longitudinal Study. In *ICERI2019 Proceedings* (pp. 442-446). IATED. https://doi.org/10.21125/iceri.2019.0144

Schunk, D. H., & Zimmerman, B. J. (Eds.). (2008). Motivation and self-regulated learning: Theory, research, and applications. Routledge.

Stary, C., & Weichhart, G. (2017). Enabling Digital Craftsmanship Capacity Building: A Digital Dalton Plan Approach. In *Proceedings of the European Conference on Cognitive Ergonomics* (pp. 43–50). ACM. https://doi.org/10.1145/3121283.3121287

Tahir, M.H.M., Albakri, I.S.M.A., Adnan, A.H.M., Shaq, M.S.Y., Shah, D.S.M. (2020). The application of visual vocabulary for ESL students' vocabulary learning. *Arab World English Language Journal (AWEJ)*, *11*(2), 323-338. https://doi.org/10.24093/awej/vol11no2.23

Tian, A. (2017). Research on classroom reform based on the Dalton Plan in era of "Internet +." Modern Education Technology, 26–30.

Yang, J., Sha, M., Sun, T., Pang, J., & Cao, G. (2023). " Dalton Plan " high school aerobics curriculum construction and implementation path. *International Journal of Educational Science and Theory*, *2*(1), 9–13.

Zhang, J. (2018). Enlightment of freedom and cooperation in the Dalton Plan to college classroom teaching. *Literature Education*, June, 146–147. https://doi.org/10.16692/j.cnki.wxjyx.2018.06.066

Zhang, L., & Seepho, S. (2013). Metacognitive strategy use and academic reading achievement: Insights from a Chinese context. *Electronic Journal of Foreign Language Teaching*, *10*(1), 54–69.https://e-flt.nus.edu.sg/wp-content/uploads/2020/09/v10n12013/zhang.pdf

Zulkepli, N., Albakri, I. S. M., Khalid, P. Z. M., & Ishak, M. F. F. (2024). Using the ADDIE model to develop an online professional development program for non-specialist ESL teachers. *Arab World English Journal (AWEJ) Special Issue on CALL (10)*, 302-313. https://dx.doi.org/10.24093/awej/call10.19

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp.257-272
DOI: https://dx.doi.org/10.24093/awej/vol16no4.15

From Silent to Speaking: Reducing Speaking Anxiety in Low-Proficiency ESL Learners Through Game-Based Learning

Mafarhanatul Akmal Ahmad Kamal
Academy of Language Studies,
Universiti Teknologi MARA, 40450 Shah Alam, Selangor Darul Ehsan, Malaysia.
Email: mafarhanatulakmal@uitm.edu.my

Mohd Fadhli Shah Khaidzir
Pusat Pengajian Citra Universiti
Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor Darul Ehsan, Malaysia.
Corresponding Author Email: cali.khaidzir@ukm.edu.my

Received:08/11/2025 Accepted:11/25/2025 Published:12/10/2025

Abstract
Language connects people, as it is one of the important tools for communication. Having one shared language to communicate to one another is essential particularly in today's a fast-paced borderless world. In Malaysia, despite English being learnt formally in academic settings for an average of 11 years for each Malaysian student, their language proficiency is still concerning. Thus, this study explores the potential of using game-based learning activities in reducing speaking anxiety among low-proficiency English as a second language (ESL) learners at the tertiary level. This study employed a qualitative method, involving 18 participants answering an open-ended survey on their perceptions of using game-based learning activities for their English language class at the end of their semester. The findings revealed that these low-proficiency students feel comfortable using the language after being exposed to game-based learning activities throughout the semester. The game-based learning activities promote a conducive learning environment, increase motivation, and reduce their anxiety, making them feel confident using the language. Consequently, game-based learning activities effectively reduce speaking anxiety in using the English language among low-proficiency ESL learners by providing a supportive and enjoyable learning environment. This approach holds significant potential for enhancing communicative confidence among low-proficiency ESL learners in Malaysian tertiary institutions.
Keywords: communicative confidence, English as a second language, game-based learning, low proficiency, speaking anxiety

Cite as: Mafarhanatul Akmal Ahmad Kamal & Mohd Fadhli Shah Khaidzir (2025). From Silent to Speaking: Reducing Speaking Anxiety in Low-Proficiency ESL Learners Through Game-Based Learning. *Arab World English Journal (AWEJ), 16*(4).257-272. https://dx.doi.org/10.24093/awej/vol16no4.15

Introduction

Proficiency in the English language is vital for learners to engage effectively in the academic, professional, and social environments, especially in a progressively interconnected world. In Malaysia, the English language is taught for an average of 11 years in formal education; nonetheless, numerous students, particularly in higher education, continue to encounter difficulties with oral communication skills (Mohd Fadhli Shah et al., 2024). For Malaysian students, their Sijil Pelajaran Malaysia (SPM) examination taken at the end of their secondary school serves as the benchmark for their academic performance before entering tertiary education. This examination includes some mandatory subjects, including English, which is aligned with the Common European Framework of Reference for Languages (CEFR), allowing their proficiency results to be recognised internationally (Chua, 2024). The CEFR scale categorises language learners into six proficiency levels which are basic users at A1 and A2, independent users at B1 and B2, and proficient users at C1 and C2 (Azurawati & Ramiaida, 2021). An A2 CEFR level indicates that a learner is a basic user who can understand and use simple, common sentence structures, especially in everyday situations, but may find complex and spontaneous communication challenging. Low-proficiency learners of English as a Second Language (ESL) frequently experience speaking anxiety, which is characterised by fear, apprehension, or hesitancy when compelled to communicate in the target language. This worry serves as both an emotional and cognitive impediment, obstructing active engagement, diminishing the desire to speak, and constraining possibilities for language practice. The challenging situation is made worse by conventional language instruction techniques that prioritise grammar, writing, and reading, frequently neglecting the development of speaking skills. Therefore, innovative, learner-centred methodologies are essential to mitigate the affective filters that hinder language acquisition and establish secure, engaging environments for learners to practise speaking.

Although a significant amount of research has been conducted on the application of game-based learning (GBL) in language education, most of the current literature focuses on its effects on motivation, vocabulary acquisition, and general engagement rather than its potential to reduce speaking anxiety, particularly among low-proficiency ESL learners. Studies that do investigate anxiety frequently concentrate on test anxiety or general language learning anxiety, with a limited number of studies specifically addressing the communicative challenges encountered by low-proficiency learners. Moreover, the direct applicability of the findings to tertiary-level ESL learners in Malaysia is restricted by the fact that a substantial number of these studies are conducted in primary or secondary school contexts or in non-Malaysian settings, where sociocultural factors and classroom dynamics may differ significantly.

In addition, while recent research has indicated that GBL can reduce affective barriers and promote communicative risk-taking, there is a lack of empirical evidence that systematically investigates its effectiveness in reducing speaking anxiety among English learners with low proficiency. The sustained impact of GBL when integrated throughout an entire semester is a space that has been left unexplored by numerous prior investigations, which concentrate on short-term interventions or one-off activities. The truancy of context-specific and longitudinal research requires studies that not only examine the effects of GBL on speaking anxiety but also record learners' perceptions of how these activities influence their confidence and willingness to communicate. By closing this gap, educators and curriculum designers can gain valuable insights that will allow them to establish more inclusive and supportive language learning environments.

Students at tertiary institutions should be equipped with essential skills to survive, such as the ability to communicate effectively and access to learning materials. Most educational resources and the medium of instruction are in English, making it crucial for learners to have high proficiency in the language. Alarmingly, many Malaysian university students still struggle with speaking proficiency despite having exposure to the language. Despite years of formal English learning, they frequently avoid speaking opportunities due to low self-confidence, negative peer evaluation, and fear of making mistakes. This anxiety can be worsened by traditional teaching techniques that focus more on grammatical accuracy over fluency, which results in a high-pressure environment that discourages active participation and restricts communication development. Thus, GBL has the potential to resolve this issue by offering participative, interactive, and low-stress activities that shift the focus from error avoidance to meaningful communication. However, there is insufficient empirical evidence regarding its efficacy in reducing speaking anxiety among low-proficiency ESL learners in Malaysia making it crucial to comprehend the role of GBL in the establishment of supportive and pleasurable speaking environments in order to develop innovative teaching strategies intending to enhance the communicative confidence of learners. Thus, this article aims to explore the use of GBL activities and examine their effectiveness in reducing speaking anxiety among A2 learners, who are categorised as low-proficiency ESL learners according to the CEFR.

Literature Review

GBL has emerged as a dynamic and learner-centred approach in language education, with the goal of enhancing the motivational power of games to promote communication and ease learning apprehensions. Theoretical frameworks like Krashen's (1982) suggest that reducing anxiety and increasing motivation can meaningfully improve language acquisition. GBL frequently accomplishes its objective by implementing authentic, entertaining contexts that prioritise meaningful interaction and minimise the repercussions of errors (Chadafi, 2023), redirecting the learners' focus to communicating rather than accuracy. Furthermore, GBL uses the principles of self-determination theory and flow theory, providing defined objectives, immediate feedback, autonomy, and a sense of competence, promoting motivation and reduced anxiety, all of which foster sustained language use and engagement (Adipat et al., 2021; Hou & Keng, 2020).

Past research shows positive empirical evidence on the effects of GBL in ESL contexts, particularly in helping students' language learning using GBL activities through collaborative tasks, scaffolding, and immediate feedback. For example, Wong and Melor (2023) found that online board games significantly improved learners' speaking skills, including their fluency and engagement in classroom activities. Chan and Lo (2024) reported that students have the passion and desire to participate in GBL environments, as they are motivated both intrinsically and extrinsically by the collaborative and competitive nature of learning. Collaborative digital game activities in tertiary ESL classes significantly enhanced learners' oral participation and significantly decreased their speaking apprehension compared to conventional drill-based methods. Mohd Fadhli Shah and Mafarhanatul Akmal (2023) asserted that in non-digital GBL, game elements like challenges and rewards, as well as immediate feedback from classmates, create interactive and immersive environments, improving students' participation and engagement. Tian and Irfan Naufal (2025) further contend that gamified elements like points, badges, and scoreboard facilitate effective language learning by fostering independent learning, which consequently minimises language anxiety. In addition, interactive learning tasks in GBL activities, such as discussions, storytelling, and authentic interactive elements, boost learners' confidence and

engagement. Similarly, the integration of group-based collaborative tasks like narrative role-play games and debate or group discussions into speaking tasks facilitated the reduction of affective barriers and the enhancement of fluency in low-proficiency learners over time (Alzubi et al., 2024; Siti Fatimah et al., 2020). These findings are consistent with systematic evaluations that recognise improvements in motivation, willingness to communicate, and speaking proficiency that result from well-designed GBL interventions (Zolfaghari et al., 2025).

However, academicians suggest that the success of GBL is dependent upon the meticulous design of the instructional materials. Game tasks must be consistent with language learning objectives, rather than solely serving as entertainment. Learners may engage in play instead of learning when scaffolding and a distinct linguistic focus are absent (Reinders & Wattana, 2015). Additionally, teacher mediation is essential, ensuring learners understand the procedures and the expected outcomes of the activities and providing reflective discussion after the activities (Sun et al., 2023). In Malaysian tertiary contexts, contextualisation is essential, as cultural norms regarding saving face and students' reticence due to fear of making mistakes and keeping a positive outlook are highly culturally embedded, influencing their mannerisms and engagement in learning as well as participation in language class activities (Badrasawi et al., 2020; Le, 2024). Hence, the integration of GBL is highly promising; however, it necessitates careful execution with instructors' mediation.

Learner Motivation and Engagement

Motivation is a critical factor in the acquisition of a second language (L2), as it influences the strategies that learners employ, the amount of effort they invest, and their persistence in the face of challenges. Motivation is a multifaceted construct that encompasses intrinsic and extrinsic factors, as well as integrative and instrumental orientations, in L2 learning (Alizadeh, 2016; Dörnyei & Ushioda, 2011). Alsawaier (2018) asserted that there are strong relationships between motivation and learning engagement. Low motivation can result in disengagement and stagnation, while high motivation can support learners throughout the frequently protracted and demanding process of acquiring communicative competence. Engagement, which is closely associated with motivation, is evidenced by the degree to which learners are actively engaged in language tasks, including cognitive, affective, and behavioural components (Ariyurek & Yurtseven, 2024; Fredricks et al., 2004). This is connection is further emphasised by Allen-Abulhassan (2025), who asserts that "student engagement and motivation also play a significant role in learning outcomes" (p. 330). In the event that learners are both motivated and engaged, they are more inclined to pursue opportunities to use the language beyond the classroom, practise speaking, and take risks.

GBL has been demonstrated to have a positive impact on both motivation and engagement in L2 contexts by offering interactive, pleasant, and goal-oriented activities. GBL can fulfil learners' intrinsic motivational requirements for autonomy, competence, and relatedness by integrating elements such as rewards, collaboration, and challenges, as suggested in Self-Determination Theory (Deci & Ryan, 1985). Research has demonstrated that learners are more inclined to engage, spend more time on tasks, and exhibit greater persistence when language learning is integrated into game environments (Reinders & Wattana, 2015; Zhou, 2024). Lin et al. (2025) further contend that motivation is crucial in facilitating language learning success, particularly through intrinsic interest, as learners show greater engagement when they enjoy acquiring the language through interactive platform. Additionally, GBL redirects the emphasis from language as an academic discipline to language as a tool for accomplishing significant

objectives, thereby enhancing both cognitive investment and affective engagement. This dual enhancement of motivation and engagement is especially advantageous for low-proficiency ESL learners, as it can establish a supportive environment that motivates them to surmount speaking anxiety and cultivate communicative confidence.

Language Anxiety

Language anxiety is one of the most common affective barriers in second language (L2) acquisition, particularly among learners with low proficiency. It is frequently accompanied by avoidance behaviours and limited participation, as well as feelings of apprehension and anxiety or nervousness when compelled to communicate in the target language (Horwitz et al., 1986). Speaking anxiety may result from the dread of making errors, negative peer evaluation, and the pressure to deliver precise speech in real time in the context of ESL. Learners' fluency and overall communicative competence are also impacted by this anxiety, which also affects their willingness to communicate. The affective filter is heightened by high levels of anxiety, which restricts the processing of language input and reduces opportunities for output, both of which are crucial for the development of oral language (Krashen, 1982).

Even among students who have completed years of formal English instruction, speaking anxiety remains a persistent challenge in Malaysian tertiary institutions. The hierarchical nature of classroom interactions and cultural factors such as "saving face" may prevent further learners from speaking up (Reinders & Wattana, 2015). Past research has shown that without targeted, supportive, and engaging interventions, anxious learners often remain silent, missing critical opportunities to practice and refine their communication skills (Goodarzi & Namaziandost, 2025; Liu & Jackson, 2008). Nor Aida et al. (2025) further corroborate that language anxiety among Malaysian ESL undergraduates is often influenced by their unwillingness to communicate using the language due to fear of negative evaluation, communication anxiety, and test anxiety. This emphasises the significance of establishing a learning environment that is supportive and low-stress and fosters risk-taking, while also framing errors as a natural part of the learning process. Interventions like GBL have the potential to reduce apprehension by redirecting attention from linguistic form to communicative meaning, thereby allowing learners to engage more confidently in speaking tasks.

Despite the expanding corpus of literature on GBL in language education, current research has not adequately addressed the unique challenges that A2 CEFR learners encounter, including a restricted vocabulary range, diminished confidence, and elevated levels of language learning anxiety. Despite past studies concentrating on intermediate and advanced users, beginner-level learners necessitate teaching interventions that are solely designed to adhere to their linguistic limitations and affective requirements. This inconsistency is especially important because beginner learners may show a distinct response to GBL in comparison to their more advanced classmates, particularly in the context of aligning cognitive demands with motivational support. Moreover, research on GBL for this demographic is still fragmented, with a small number of studies evaluating both the linguistic and affective aspects of learning within a single framework. Additionally, current past research primarily emphasises descriptive analysis by having students complete scaled questionnaires, which results in limited findings, especially for beginner learners who have limited linguistic abilities. The findings using an open-ended survey may provide an in-

depth view of how GBL influence students' language learning, including motivation, confidence, engagement, and speaking anxiety.

Methods

This study uses a qualitative research approach, through the use of an open-ended survey to explore the effectiveness of GBL activities for language learning among low-proficiency ESL learners. Creswell and Creswell (2017) noted that a qualitative approach enables in-depth exploration of the phenomenon to be examined, making it relevant for this study. The focus of this was to explore the actual perceptions of learners towards the use of GBL activities, through a qualitative approach which results in a richer and more profound understanding of the context (Akyildiz & Ahmed, 2021).

Research Participants

The participants were selected using purposive sampling, looking into how beginner language users view the effectiveness of integrating GBL activities in class. A total of 18 students were selected from the intended cohort of 47 students. The participants in this study were pre-diploma students who had a CEFR A2 level for their SPM examination and were taking English as part of their semester study plan.

Research Instrument

The open-ended survey collected more extensive responses than the interview, further strengthening the findings. Jain (2021) asserts that survey questions are common and have long been used in qualitative research, as they are easier to conduct and allow access to a larger group of responses online. The survey was divided into two sections, each of which contained demographic information and four open-ended questions regarding their perceptions of GBL activities. The questions were designed based on the previous literature to thoroughly examine the effectiveness of GBL activities for low-proficiency students.

Research Procedures

Prior to the data collection, ethics approval was granted by the Universiti Teknologi MARA (UiTM) Research Ethics Committee on January 26, 2025, adhering to ICH Good Clinical Practice Guidelines, Malaysian Good Clinical Practice Guidelines and the Declaration of Helsinki with the reference number REC/02/2025 (ST/MR/33). The participants were briefed and informed about the study before participating in it. They were informed of their right and allowed to withdraw from the study at any moment without penalty. Their responses were assured that they would be kept confidential. This procedure was done during their final class at the end of the semester by giving them the open-ended survey through Google Forms, as it was convenient for the participants.

Data Analysis

The open-ended survey was analysed using thematic analysis to identify recurring themes aligned with the objective of this study. To proceed with the analysis, data cleaning was done by omitting the outliers, particularly those students who had CEFR above the A2 level. This procedure was to ensure that the findings aligned with the objectives and scope of the study, leaving only 18 students to participate in it.

Results

This study explored how low-proficiency students perceive the use of GBL activities for language classroom. There were 18 participants with an A2 CEFR level based on their English results in the Sijil Pelajaran Malaysia (SPM), with 13 male and five female respondents. The thematic analysis presents the following findings, organised according to the four open-ended questions.

Table 1. *Learner motivation through GBL activities*

Participants	Responses
Confidence Gained Through Learning	
P3	*"With game, i can improve my confidence"*
P5	*"the game teach me how to communicate with confident"*
P6	*"can improve more confident"*
P18	*"That help me to be confident more"*
Enjoyment and Fun in Learning	
P2	*"fun"*
P4	*"Games can be a great way to learn English because they make the learning process more interactive, fun, and engaging"*
P11	*"game looks easier"*
P14	*"make me feel enjoy learning English"*
P15	*"Learning in an interactive and fun way"*
P17	*"the mission makes me want to know how to complete it"*
Improved Oral Communication Skills	
P1	*"speaking a lot"*
P7	*"Improving grammar which is many story and role-playing games use mostly proper sentences ,helping you to improve speaking"*
P8	*"practices listening and speaking skills"*
P9	*"i more understand English when play games"*
P12	*"because the subtitles we use in the game are in English"*

Table 1 presents the impacts of GBL activities on learners' motivation in using the language. The three themes for this question include learners' confidence gained through learning, enjoyment and fun in learning, as well as how GBL activities improved oral communication skills. Several participants expressed that games helped them gain confidence in using the language as reflected in statements such as *"With game, i can improve my confidence"* (P3) and *"the game teach me how to communicate with confident"* (P5). P6 and P18 also stated that they gained confidence when exposed to games in language learning, noting that games *"can improve more confident"* and *"That help me to be confident more,"* respectively. In addition, many participants asserted that GBL activities were enjoyable and engaging. P4 and P15 found that the learning process was more interactive and enjoyable, while P2 simply described the activities as *"fun."* Including games as part of language learning also made the lessons feel less complicated as agreed by P11 noting that *"game looks easier"* and P14 noting that *"make me feel enjoy learning English"*. This sense of ease thereby influenced the learners' curiosity and spark their interest in

completing the tasks, suggesting that intrinsic motivation, as stated by P17 *"the mission makes me want to know how to complete it"*. Lastly, several learners recognised improvements in their communication skills, owing to constant speaking, listening, and exposure to language through games. P1 highlighted the importance of *"speaking a lot,"* whereas P12 emphasised the benefit of using English subtitles throughout the game. P7 also agreed that using games helped learners to improve their speaking skills and P9 emphasised the impact of games in making better understanding of the language. The findings indicate that enjoyment, confidence, and meaningful language use in a GBL setting are significant motivators for low-proficiency ESL learners.

Table 2. *Increased confidence in speaking English through GBL activities*

Participants	Responses
Positive Emotions	
P14	*"Yes! Playing games can make speaking English feel more natural and less stressful."*
P5	*"yes i feel better when i playing games"*
P7	*"Yes, I'm feeling good"*
P12	*"yes i feel good"*
Reduced Pressure	
P14	*"Since games create a fun and relaxed environment, many people feel more confident using English while playing."*
P17	*"yes I am so interested to learn English"*

Table 2 illustrates how exposure to GBL activities enhances confidence in language learning. There are two themes that emerged for this question, which are positive emotions and reduced pressure. Some students expressed feeling emotionally positive and at ease while communicating in English during the GBL activities. P5, P7, and P12 shared that they *"...feel better…"* and *"...feel good…"* when engaging in games. P14 asserted further that games made the learning process and speaking English *"more natural and less stressful."* These responses indicate that GBL activities promote a more relaxed and emotionally comfortable atmosphere, supporting learners to participate in speaking without fear of judgement or failure. In addition to the positive emotional responses, participants admitted that they felt less pressure to use the language when playing games. P14 emphasised that games create a *"...fun and relaxed environment..."* which enhances learners' confidence in using English. Similarly, P17 stated that *"yes I am so interested to learn English"*, agreeing that less pressure in using language correctly helps learners to have high interest and encourages their involvement in the language learning process. These responses indicate that a supportive and enjoyable classroom, free from fear of making mistakes or being judged by their friends, facilitates learners' confidence in speaking English.

Table 3. *Learner engagement during GBL activities*

Participants	Responses
Collaborative Peer Interaction	
P3	*"Because i can practice speaking with my friends"*
P6	*"because it's fun to be able to work together with friends"*
P10	*"Because we can communicate with our classmates the most"*

Supportive Learning Environment	
P4	*"Playing games in English class makes learning more fun, interactive, and effective"*
P5	*"can get energise in class because can communicate with friends"*
P9	*"yes because i enjoy the class without feel pressure"*
P12	*"because i like playing with my friend"*
P16	*"fun and chill"*
P18	*"My hobby is playing game and thats make me easy to learn more"*
Competitive Engagement	
P2	*"challenge because want to beat my friends mark"*
P7	*"I can learn some new vocabulary and practice for my grammar and try to speak"*
P8	*"I like playing games in English class because it can improve my speaking."*
P13	*"because it's more fun when i play with my friend and make me want to try again if i make mistakes"*
P15	*"Improving listening and reading comprehension"*

Table 3 illustrates learners' engagement while playing games in the language classroom. There are three major themes analysed in the responses, such as collaborative peer interaction, supportive learning environments, and competitive engagement. Participants emphasised the importance of collaborative peer interactions during games, as accounted for by P3, who noted that *"Because i can practice speaking with my friends"*, and P6, who expressed, *"because it's fun to be able to work together with friends."* Other than that, peer communication is crucial in sustaining participation in language learning, as implied by P10, who stated that "Because we can communicate with our classmates the most." A supportive learning environment is also prominent in participants' responses, as can be seen from P4, who stated that "Playing games in English class makes learning more fun, interactive, and effective", and P5 explained that "can get energise in class because can communicate with friends." P9 described the classroom environment as enjoyable and free of stress, stating, *"...because i enjoy the class without feel pressure"*, while P12 and P16 regarded the experience playing games with their friends as *"fun and relaxed."* Moreover, GBL activities, which correspond closely with individual learner interests, significantly boost student engagement in the classroom. P18 affirmed this by stating that "My hobby is playing game and thats make me easy to learn more." In addition to a supportive learning environment, the theme of competitive engagement was common among participants. P2 enjoyed the challenge, noting that *"challenge because want to beat my friends mark"*, expressing a desire to surpass their friend's record. P7 and P8 stated that the competitive environment helps them to learn and use the language, noting that playing games *"...learn some new vocabulary and practice for my grammar and try to speak"* and *"...playing games in English class because it can improve my speaking"*, respectively. P13 found motivation in attempting again after errors *"...when i play with my friend and make me want to try again if i make mistakes"*, while P15 indicated that games facilitated enhancements in listening and reading comprehension. All of the themes demonstrated that GBL encourages learners to use the language by promoting interaction with others, providing an enjoyable and conducive learning environment, and incorporating elements of challenge and repetition.

Table 4. *Reduction of speaking anxiety through GBL activities*

Participants	Responses
Anxiety Reduction Through Fun Context	
P4	*"Yes, games can make speaking English less scary because they create a relaxed and enjoyable environment for learning"*
P5	*"yes can less scary. We can play more activities"*
P6	*"with games can make conversation in English become fluent because it is fun to do with friends"*
P8	*"speaking English is less scary because we can learn and collaborate with classmates during game-based activities."*
P9	*"games makes learning English fun"*
P12	*"because understanding English is very fun"*
P15	*"create a more relaxed and fun environment."*
Confidence Through Repetition	
P1	*"game with speaking English , we get more confident"*
P2	*"We can play more activities"*
P3	*"Always practice"*
P7	*"encourage people to build their confidence to speak also be brave"*
P13	*"in games we can practice again and again"*
P14	*"when speaking English while playing game makes me feel more better"*

Table 4 explores the usefulness of GBL activities in reducing speaking anxiety among low-proficiency ESL learners. The responses to this question reveal two major themes, which are anxiety reduction through a fun context and confidence building through repetition. Many participants expressed that games facilitate an engaging and enjoyable learning environment, making speaking English less intimidating. P4 agreed to this by stating that *"Yes, games can make speaking English less scary because they create a relaxed and enjoyable environment for learning"*, while P15 similarly highlighted that they *"create a more relaxed and fun environment."* P6 expressed that games enhance the ability to speak using English, *"...as it is fun to do with friends"*, while P8 agreed that speaking English was less intimidating by stating that it was *"less scary because we can learn and collaborate with classmates during game-based activities."* It can also be observed that the other participants agreed that speaking English while playing games was less intimidating as compared to a conventional language learning classroom, as playing games reduces the fear of speaking and provides the need to communicate using the language. Another significant theme regarding how GBL activities alleviate speaking anxiety is the development of confidence through repetition. Participants expressed that using games exposes them to practising using the language outside their comfort zone, thereby developing language confidence. P1 mentioned that playing games required them to speak in English, which made them *"...more confident"*, whereas P3 merely stated, *"Always practise."* P7 highlighted that games *"encourage people to build their confidence to speak also be brave"*, and P13 added that *"in games we can practice again and again"*. P14 echoed that confidence can be gained through repetition by stating, *"when speaking English while playing game makes me feel more better"* as repetition makes the learners communicate better, hence making them feel better about themselves. Overall, the responses indicate that GBL reduces learners' emotional barriers to speaking English by making

language learning activities more enjoyable and allowing learners to use language repetitively without fear of making mistakes.

Discussions

This study explores the use of GBL activities for low-proficiency ESL learners at tertiary institutions, as students are expected to excel or be able to communicate using the language for daily activities since they are generally exposed to English language for 11 years of formal education. However, Mohd Fadhli Shah et al. (2024) asserted that, despite prolonged exposure to English and its status as the second language in Malaysia, students' proficiency remains worrying. This is further concerning, as this study involved 18 pre-diploma students, who were classified with low proficiency in English, as indicated by their A2 CEFR level in the national SPM examination. Offering these students alternative instructional approaches such as using language games may lead to better outcomes, as these methods are not limited to traditional language learning. The findings revealed that there are four major domains of responses on learners' perceptions of GBL, which include learners' motivation, confidence, engagement, and anxiety reduction.

Firstly, engaging with GBL activities in classrooms promotes higher motivation to use English. The learners were interested in learning and using the language, as their learning experiences were enjoyable and engaging. Providing unrestricted platforms for the students to use and explore the language increased their motivation to participate in the learning session actively. They also highlighted that they were willing to engage and participate in the activity, as games made learning more enjoyable and less challenging. These findings align with Vygotsky's Sociocultural Theory of Learning, who emphasises the role of social interaction in acquiring knowledge (Lantolf, 2000), suggesting that learners are highly motivated to engage in meaningful and socially interactive learning environments. The findings also support Chan and Lo (2024), which states that collaborative games significantly enhance oral participation and motivation among ESL learners.

Secondly, the findings highlight learners' increased confidence when speaking English during GBL activities because the environment was supportive and there was less pressure. They expressed feeling positive, less anxious, and more natural in speaking English, as they were not judged for making mistakes. This finding corresponds with Krashen's Affective Filter Hypothesis, that emotional variables such as motivation, anxiety, and confidence act as filters that may help or hinder language learning (Krashen, 1982). Language learning is most effective when anxiety is minimal and motivation is maximised. As such, engaging learners in intriguing and enjoyable environments enables them to communicate using the language in natural settings, not having to worry about making grammatically sound sentences as well as being judged by their classmates and instructors.

Thirdly, strong and positive engagement in language learning is imperative. The findings demonstrated that learners had strong engagement in GBL activities, stating that interacting with other students, having a supportive learning atmosphere, and friendly competition were motivating factors. Participants' interest and involvement were maintained through teaming up with classmates, enjoying class sessions, and overcoming speaking challenges through games. Despite having limited linguistic ability, participants expressed that their willingness to actively participate and challenge their friends influenced them to complete tasks due to the supportive environments. According to Self-Determination Theory (Deci & Ryan, 1985), learners are more likely to engage

in learning environments when their psychological needs for autonomy, competence, and relatedness are met. Thus, the findings were consistent with this notion, as the participants demonstrated a strong level of engagement while playing language games with their classmates, overcoming communication barriers, and enjoying the supportive GBL environment.

Finally, the findings revealed that GBL activities reduced learners' anxiety in using the language, facilitating their language learning experiences. Participants noted that they willingly and naturally used the language as learning sessions were engaging and not too intimidating. Nor Aida et al. (2025) supported that less formal and intimidating lessons with little emphasis on grammatical correctness or testing encourage learners to participate in communicative activities. Consequently, when participants were motivated to participate in the GBL activities, the participants eventually repetitively used the language. This scenario helped build their courage to speak and fluency using the language, thereby encouraging them to feel at ease in speaking and lowering their speaking anxiety. GBL activities not only provide participants with engaging and enjoyable experiences but also allow low-proficiency language learners to engage in effective learning environments without the fear of making mistakes (Mohd Fadhli Shah & Mafarhanatul Akmal, 2023).

Conclusion

Having a proficient grasp of the English language is essential and highly demanded to connect with others, especially in today's fast-paced borderless society. This is crucial for Malaysian ESL learners in tertiary institutions as part of their survival and preparation for the real working world, as English provides common access to educational materials and career opportunities. This study demonstrates GBL activities significantly enhancing low-proficiency learners' ability to communicate in English. This approach can help them reduce their language apprehension by reducing their speaking anxiety, building confidence, and fostering learning engagement. Learners are prone to participate in communicative activities in comfortable and conducive environments, as they enjoy the process of learning. GBL activities lower affective filters, hence motivating learners to use and practice English repetitively without fear of making mistakes. With the integration of both digital and non-digital GBL activities, students who have prior exposure to the language but lack confidence can benefit from them, as they use it naturally while playing the games. Peer collaboration, friendly competition, and purposeful communicative tasks enhance engagement and transform learners' emotional connection with English from apprehension to active participation.

The study emphasises the importance of GBL in ESL language classrooms, especially for low-proficiency learners who may feel intimidated and restricted in using the language. Being confined to a conventional classroom that emphasises accuracy and completes the syllabus may limit opportunities for learners to feel connected and engage in authentic learning environments. By integrating active learning that involves learners participating in communicative tasks as playful, low-stakes challenges and collaborative learning, the classroom can be shifted into a safe space for them to communicate, allowing learners to transform their silence into confident speech in the language. The findings presented in this study contribute to the field of language pedagogy by demonstrating the efficacy of GBL in reducing speaking anxiety among low-proficiency ESL learners. The findings provide insights for educators to integrate GBL in their language learning to facilitate proficiency and classroom engagement, providing interactive, low-stress environments that promote risk-taking and meaningful communication.

The study suggests the integration of structured GBL activities in speaking-focused modules and language instruction to meet the learners' cognitive and affective demands. The research also demonstrates the benefits of GBL in enhancing learners' communicative confidence by fostering supportive learning environments, making them confident enough to use the language outside classroom settings. It is recommended that future research explore larger and diverse CEFR proficiency levels and the effects of both digital and non-digital GBL activities on students' communicative competence. GBL has the potential not only to create inclusive and authentic learning environments but also to equip learners with the skills and confidence to use the language effectively in real-world situations.

Funding: This research was not funded.
Acknowledgement: The authors express gratitude to Universiti Teknologi MARA for their support in data collection and execution of this research, as well as to all individuals who contributed to the completion of this study.
Conflicts of Interest: The authors declare no conflict of interest.
Authenticity: This manuscript is an original work.
Ethical Statement: The Research Ethics Committee, Universiti Teknologi MARA (UiTM) has granted ethical approval for this research with the code reference REC/02/2025 (ST/MR/33).
Artificial Intelligence Statement: AI-assisted technologies were used to improve readability and language. The authors take responsibility for the content of the publication.

About Authors

Mafarhanatul Akmal Ahmad Kamal is a lecturer at the Department of English Language and Linguistics, Universiti Teknologi MARA (UiTM) Shah Alam. She received her Master of Arts in Applied Linguistics with TESOL from The University of Sheffield. Her research focuses on language and linguistics, bilingualism, and English language teaching. ORCID: https://orcid.org/0000-0001-9216-478X

Mohd Fadhli Shah Khaidzir is a Senior Lecturer at CITRA UKM, Universiti Kebangsaan Malaysia. He holds a PhD in Postcolonial Literature and Psychogeography. His expertise includes psychogeography, travel literature, and sense of place, with current projects on augmented reality, heritage, visual communication, and new media. ORCID: https://orcid.org/0000-0001-7873-7000

References

Adipat, S., Laksana, K., Busayanon, K., Ausawasowan, A., & Adipat, B. (2021). Engaging students in the learning process with game-based learning: The fundamental concepts. *International Journal of Technology in Education*, *4*(3), 542–552. https://doi.org/10.46328/ijte.169

Akyildiz, S. T., & Ahmed, K. H. (2021). An overview of qualitative research and focus group discussion. *International Journal of Academic Research in Education*, *7*(1), 1–15. https://doi.org/10.17985/ijare.866762

Alizadeh, M. (2016). The impact of motivation on English language learning. *International Journal of Research in English Education*, *1*(1), 11-15.

Allen-Abulhassan, T. (2025). Confidence Challenges in English Language learning: Student perceptions. *Arab World English Journal, 16*(2), 328–341. https://doi.org/10.24093/awej/vol16no2.18

Alsawaier, R. S. (2018). The effect of gamification on motivation and engagement. *The International Journal of Information and Learning Technology, 35*(1), 56-79.

Alzubi, A. A., Nazim, M., & Ahamad, J. (2024). Examining the effect of a collaborative learning intervention on EFL students' English learning and social interaction. *Journal of Pedagogical Research, 8*(2), 26-46. https://doi.org/10.33902/JPR.202425541

Ariyurek, E. A., & Yurtseven, N. (2024). School engagement and learning responsibility in middle school students. *International Journal of Curriculum and Instructional Studies, 14*(2), 220-250. https://doi.org/10.31704/1500085

Azurawati Wok Zaki & Ramiaida Darmi. (2021). CEFR: Education towards 21st Century of Learning. Why Matters? *Journal of Social Science and Humanities, 4*(2), 14–20. https://doi.org/10.26666/rmp.jssh.2021.2.3

Badrasawi, K. J., Solihu, A., & Ahmad, T. B. T. (2020). Second Language Speaking Anxiety among Malaysian Postgraduate Students at a Faculty of Education. *International Journal of Education and Literacy Studies, 8*(2), 54-61.

Chadafi, M. (2023). The Use of game-based learning on student's speaking skill and speaking anxiety: An experimental research at seventh grade of MTs YASMU. *JEET, Journal of English Education and Technology, 4*(01), 43 - 57.

Chan, S., & Lo, N. (2024). Enhancing EFL/ESL instruction through gamification: A comprehensive review of empirical evidence. *Frontiers in Education, 9*:1395155. https://doi.org/10.3389/feduc.2024.1395155

Chua, R. (2024, February 4). *Priority to improve standard of English*. *The Star*. https://www.thestar.com.my/news/nation/2024/02/04/priority-to-improve-standard-of-english

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Springer.

Dörnyei, Z. and Ushioda, E. (2011). *Teaching and Researching Motivation*. 2nd ed., Pearson Education Limited, New York, NY.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59-109.

Goodarzi, A., & Namaziandost, E. (2025). A mixed-methods study on the effects of online dynamic assessment on L2 learners' motivation and speaking anxiety: Beyond the classroom perspectives. *Computers in Human Behavior Reports, 17*, 100584. https://doi.org/10.1016/j.chbr.2024.100584

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

Hou, H., & Keng, S. (2020). A Dual-Scaffolding framework integrating Peer-Scaffolding and Cognitive-Scaffolding for an augmented Reality-Based educational board game: an analysis of learners' collective flow state and collaborative learning behavioral patterns.

Journal of Educational Computing Research, *59*(3), 547–573. https://doi.org/10.1177/0735633120969409

Jain, N. (2021). Survey versus Interviews: Comparing data collection Tools for Exploratory research. *The Qualitative Report*. https://doi.org/10.46743/2160-3715/2021.4492

Krashen, S. D. (1982). *Principles and practice in second language acquisition.* Pergamon Press.

Lantolf, J. P. (2000). Introducing sociocultural theory. *Sociocultural theory and second language learning*, *1*, 1-26. Oxford University Press.

Le, H. V. (2024). Silence in the classroom: Unraveling the cultural dynamics affecting participation and critical thinking in Malaysian and Vietnamese engineering education. *Cogent Education*, *11*(1). https://doi.org/10.1080/2331186X.2024.2404780

Lin, Y., Mohamad, Maslawati Mohamad & Mohd Izwan Mahmud. (2025). Exploring the relationship between English-Speaking Self-Efficacy and English learning motivation among Polytechnic EFL learners. *Arab World English Journal*, *16*(1), 55–69. https://doi.org/10.24093/awej/vol16no1.4

Liu, M., & Jackson, J. (2008). An exploration of Chinese EFL learners' unwillingness to communicate and foreign language anxiety. *The Modern Language Journal*, *92*(1), 71–86. https://doi.org/10.1111/j.1540-4781.2008.00687.x

Mohd Fadhli Shah Khaidzir & Mafarhanatul Akmal Ahmad Kamal. (2023). Sampan the survival game: Basic communication strategies for ESL learners. *International Journal of Academic Research in Progressive Education and Development*, *12*(4). https://doi.org/10.6007/ijarped/v12-i4/19812

Mohd Fadhli Shah Khaidzir, Mafarhanatul Akmal Ahmad Kamal, Ramachandran, D., Siti Aisyah Mohd Adaha, Nurul Nazihah Nuraddin & Maryam Fakhrul Anuar. (2024). The relationship between English proficiency and public speaking anxiety among second language learners. *International Journal of Academic Research in Business and Social Sciences*, *14*(11). https://doi.org/10.6007/ijarbss/v14-i11/23209

Nor Aida Sapuan, Fariha Diyana Awang Ali, Arif Bahari Mohamed Musli, & Rashidin Idris. (2025). Conquering the Fear: Navigating second Language speaking anxiety among ESL learners in Selangor Private University. *International Journal of Research and Innovation in Social Science, IX*(IIIS), 5146–5152. https://doi.org/10.47772/ijriss.2025.903sedu0371

Reinders, H., & Wattana, S. (2015). Affect and willingness to communicate in digital game-based learning. *ReCALL*, *27*(1), 38–57. https://doi.org/10.1017/s0958344014000226

Siti Fatimah Omar, Haslinda Sutan Ahmad Nawi, Tengku Shahrom Tengku Shahdan, Mee, R. W. M., Pek, L. S., & Fatin Syamilah Che Yob. (2020). Interactive language learning activities for learners' communicative ability. I*nternational Journal of Evaluation and Research in Education (IJERE)*, *9*(4), 1010. https://doi.org/10.11591/ijere.v9i4.20605

Sun, L., Kangas, M., Ruokamo, H., & Siklander, S. (2023). A systematic literature review of teacher scaffolding in game-based learning in primary education. *Educational Research Review*, *40*, 100546. https://doi.org/10.1016/j.edurev.2023.100546

Tian, F., & Irfan Naufal Umar. (2025). Research Trends of Digital Game-based Learning for Technical Vocabulary Acquisition: A Systematic Decennial Literature Review (2014-2024). *Arab World English Journal*, *16*(3), 60–88. https://doi.org/10.24093/awej/vol16no3.4

Wong, C. H. T., & Melor Md Yunu. (2023). Let "Flippity" speak: Using online board game to improve speaking skills among elementary pupils. *European Journal of Educational Research*, *12*(2), 1085–1096. https://doi.org/10.12973/eu-jer.12.2.1085

Zhou, S. (2024). Gamifying language education: The impact of digital game-based learning on Chinese EFL learners. *Humanities and Social Sciences Communications*, *11*(1). https://doi.org/10.1057/s41599-024-04073-3

Zolfaghari, Z., Karimian, Z., Zarifsanaiey, N., & Farahmandi, A. Y. (2025). A scoping review of gamified applications in English language teaching: a comparative discussion with medical education. *BMC Medical Education*, *25*(1), 274.

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp. 273 -288
DOI: https://dx.doi.org/10.24093/awej/vol16no4.16

Teaching ESP Vocabulary with the Application of Corpus-Based Materials at Universities: A Rapid Review of Literature

Marta Lacková
Institute of Mediamatics and Cultural Heritage
University of Žilina, Slovakia
Email: marta.lackova@umkd.uniza.sk

Received:08/29/2025 Accepted:11/07/2025 Published:12/10/2025

Abstract
This rapid review of the literature examines the effectiveness of ESP vocabulary teaching, with a focus on the application of corpus-based materials. This review aims to critically analyse recent studies on the corpus method while teaching ESP vocabulary at universities. The review also highlights the challenges associated with corpus-based materials in ESP vocabulary teaching and their perceptions among ESP students. We searched six research databases, generating 136 studies. During the evaluation process, studies not written in English, published before 2010, lacking evidence-based support, non-peer-reviewed, without abstracts, and not focused on ESP vocabulary, corpus-based approaches, or university education were excluded. Forty-five articles were included after full-text screening; we subsequently coded them according to the data they presented. The overall evidence from this rapid literature review suggests that corpora can be beneficial for many learners across various contexts. When used thoughtfully and appropriately, they become a valuable resource for both ESP teachers and students. More specific results from our research show that a corpus-based approach to ESP teaching improves linguistic skills by supplying comprehensive data on authentic vocabulary and syntax patterns, enabling insights into frequency, connotations, nuanced meaning differences, and proper collocations. The research also shows that utilizing available resources and seeking guidance from experts can help overcome challenges connected with corpus-based materials. As this is the first rapid review in the field of corpus-based ESP vocabulary learning at universities, its findings offer valuable insights for shaping effective academic policies across diverse ESP disciplines.
Keywords: authentic materials, corpus, corpus-based materials, ESP, rapid review, teaching, vocabulary

How to cite: Lacková, M. (2025). Teaching ESP Vocabulary with the Application of Corpus-Based Materials at Universities: A Rapid Review of Literature. *Arab World English Journal, 16*(4): 273 -288. https://dx.doi.org/10.24093/awej/vol16no4.16

Introduction

The majority of students studying English in higher education are enrolled in fields other than English itself, such as economics, medicine, engineering, science, or interdisciplinary areas like specialized terminology and translation. Enhancing the relevance of tertiary courses in English for Specific and/or Academic Purposes (ESP/EAP) to meet the professional needs of these students—particularly in mastering the specialized vocabulary pertinent to their disciplines—can be significantly improved through a corpus approach (Jitpainboon & Rungswang, 2022). The corpus approach offers an empirically grounded understanding of the discourse and language used within specific professional contexts, thereby equipping students with greater independence in performing their professional tasks (Boulton et al., 2012).

The widespread use of corpora has been monitored for research purposes in areas such as lexicography within specialized registers and materials (Flowerdew, 2005; Jaafar, 2022; Martínez, Beck, & Panza, 2009), as well as in pedagogical approaches to ESP courses, which involve students' engagement in the learning process. Yet, before 2010, ESP, in this regard, resembled English as a Foreign Language (EFL), as both faced a recurring theme in the studies: a lack of effective integration between corpus research and teaching practice in EFL and Second Language Acquisition (SLA) (Aijmer, 2009; Sutarsyah et al., 1994).

Recently, the outlook seems more promising at the graduate and postgraduate levels. Since English has become the primary language of research publication across many academic fields, numerous studies employing corpus-based methods for the needs of ESP vocabulary learning have been conducted in recent years (Coxhead, 2025; Curry, N., McEnery, 2025; Flowerdew, 2015; Huang & Wang, 2023; Ryan, G., 2012; Vyatkina & Boulton, 2017). The emphasis is placed on exploring effective ways to integrate ESP vocabulary teaching and corpus technologies, proposing new models of corpus-based vocabulary teaching for professionals (Charles, 2013; Huang & Wang, 2023; Römer, 2011).

Literature Review

Corpus linguistics has also had an indirect impact on the development of ESP teaching materials (Chen & Liu, 2020; Lelakova & Lackova, 2015), integrating specialized corpora into university courses for non-native English speakers. Within the curriculum design context, scholars strive to develop ESP vocabulary lists (Hsiao-I, 2014; Lei, Liu, 2016; Muñoz, 2015; Tongpoon-Patanasorn, 2017). The reason is that the vocabulary varies significantly across individual disciplines in terms of meaning, frequency, range, and collocation. Laosrirattanachai & Ruangjaroon (2021) highlight that one of the primary benefits of the created word lists is that they prepare learners for their future careers, as the corpora used to develop these lists are mainly derived from authentic sources like websites. In contrast to other constructed word lists, which often rely on textbooks or research articles, these lists are more reflective of actual language usage in the real world. The scholars have advocated for the application of these lists in teaching ESP vocabulary, both in class and autonomously as supplementary material (Laosrirattanachai & Ruangjaroon, 2021).

Considering the elaboration of ESP word lists, Tangpijaikul (2014) combined intuition-based and corpus-based approaches in developing a list of ESP words from the field of business; the researcher suggests that future research should aim to explore the most effective methods for teaching these words within the ESP classroom, including ways to raise students' awareness of

their collocational patterns. Additionally, studies should investigate which teaching strategies best support learners in retaining these words in their long-term memory.

Gardner, D., & Davies, M. (2014) introduced a new approach to creating corpus-based vocabulary academic lists. The crucial point is that lemmas, rather than word families, are used to generate the initial counts and analyses. Subsequently, word families are created from these lemmas to fulfill specific academic requirements.

In our research sample, we encountered corpus-based elaboration of ESP vocabulary from the following scientific disciplines: education, humanities, history, social science, medicine and health, economics and finance, law and political science, agriculture, chemistry, theology, automobile industry, tourism, hotel and airline business, philosophy, religion, psychology, and psychiatry.

Furthermore, some studies typically use corpora to examine attributes of lexis connected with various ESP disciplines. Examples include a corpus-based analysis of a word's collocational patterns, which reveal its multiple meanings or uses and determine if it carries any negative connotations or associated prosody. This information is highly valuable to lexicographers, materials developers, ESP teachers, and ESP learners (Walker, 2011). It has also been proven that specialised corpora exhibit a higher frequency of content words (Hsiao-I, 2014).

Based on the preceding literature review and in concord with Jitpainboon and Rungswang (2022, p.147), who claim that "the pedagogical application of corpus linguistics in the ESP domain appears to be underresearched", there are no existing rapid reviews examining the effectiveness of corpus-based materials for teaching ESP vocabulary at the tertiary level. The available systematic papers do not focus on the impact of corpus-based materials on the development of ESP vocabulary; instead, they systematically analyse teaching ESP vocabulary in accordance with the guidelines set by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). Additionally, these reviews do not explore the effectiveness of corpus-based approaches within the context of teaching ESP vocabulary in higher education.

Thus, this review seeks to address this gap by investigating the effectiveness of corpus-based approaches in teaching ESP vocabulary at universities. It aims to contribute to the existing literature by providing empirical insights into how corpus-based materials can enhance ESP learning outcomes among university students. Ultimately, this review will offer core recommendations to improve the design and implementation practices of corpus linguistics courses in English education, thereby promoting improved learning outcomes in the higher education context.

The paper employs the rapid review approach to examine and subsequently synthesize findings about the application of corpus-based materials for ESP vocabulary development needs at universities. The review aims to achieve this by addressing the following research questions:

1. How can corpus-based materials enhance ESP vocabulary teaching, learning, and assessment?
2. What are the students' perceptions and opinions regarding the teaching, learning, and assessment of ESP vocabulary using corpus-based materials?
3. What are the challenges associated with corpus-based materials in terms of ESP vocabulary teaching, learning, and assessment?

Methodology

Inclusion and exclusion criteria

Research databases searched included Web of Science, Scopus, Google Scholar, ProQuest, ScienceDirect, and ERIC to ensure the credibility and scientific value of the articles under review. The search strings *"ESP and teaching", "ESP and corpus", "ESP vocabulary and corpus", and "ESP and university" were used in each of the databases* as mentioned above to search applicable articles. A total of 136 peer-reviewed articles with a publication period since 2010 were selected for the needs of this rapid literature review. Articles that are non-academic in nature, such as those published on mainstream and social media platforms, blog posts from personal or institutional websites, and review articles from technology magazines on corpora, were excluded. Articles related to corpora in other than ESP teaching and learning used as general background references were also excluded from this rapid literature review. For more details on the inclusion and exclusion criteria, refer to Table 1 below.

Table 1. *Exclusion criteria and inclusion criteria*

Inclusion criteria	Exclusion criteria
published since 2010	published before 2010
peer-reviewed journal articles	non-peer-reviewed journal articles
conference papers	non-conference papers
evidence-based articles	non-evidence-based articles
written in English	non-written in English
articles with abstract	articles without an abstract
articles on ESP teaching and learning	articles op on other than ESP teaching and learning
articles on corpus-based approaches	articles on other than corpus-based approaches
articles reporting on specific outcomes	articles reporting on general outcomes
articles relying on peer-reviewed sources	articles relying on non-peer-reviewed sources

This rapid review was conducted in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) 2020 statement elaborated by Page et al. (2021). The PRISMA chart below illustrates the number of articles available from the databases. It demonstrates the number of records identified in the databases through the screening procedure, up to the number of studies included in the rapid review process. Eventually, a simple quantitative count was conducted to support each subject matter.

The raw data were manually analyzed and subsequently interpreted using both deductive and inductive analysis approaches. During the preparatory phase of the review process, the raw data were outlined and summarized into concise formats. A coding scheme was applied to categorize data that shared similar types of information. In the description phase, the key features and characteristics of the categories derived from the coding were detailed, and links between the data and the original research questions were established. In the final stage of interpretation, the relationships, similarities, and contradictions among and between the coded categories were analyzed.

The findings and conclusions from the reviewed articles – raw data – were manually analysed following the subsequent procedure: first, the raw data were coded using a Microsoft Word document; next, the codes were organized and refined to maintain consistency throughout the process; finally, a thematic framework was developed based on the generated codes. To enhance the reliability and validity of the analysis, a second coder was engaged. The second coder was briefed on the study's purpose, methodology, research questions, and coding scheme.

Additionally, a calibration session was held to align the coding approach and ensure consistent analysis of the raw data.

Table 2 *Results of the search/selection process*

Identification						
Titles/abstracts /key words from Web of Science (n=21)	Titles/abstracts /key words from Scopus (n=29)	Titles/abstracts /key words from Google Scholar (n=34)	Titles/abstracts /key words from ProQuest (n=24)	Titles/abstracts /key words from Science Direct (n=19)	Titles/abstracts /key words from ERIC (n=9)	Total retrieved (n=136)
↓						
Screening						
Duplicates removed (n=12) Articles after duplication removed (n=124)		Articles meeting the inclusion criteria (n=95) Articles not meeting the inclusion criteria (n=29)		Total screened (n=124) Total after screening (n=95)		
↓						
Eligibility						
Full articles meeting the inclusion criteria (n=74) Full articles not meeting the inclusion criteria (n=21)			Full articles on the topic of the rapid review (n=45) Full articles not meeting the inclusion criteria (n=29)			
↓						
Included						
Articles included in the rapid review (n=45)						

Findings

RQ 1: How can corpus-based materials enhance ESP vocabulary teaching, learning, and assessment?

The subsequent part discusses how corpus-based materials can support and enhance ESP vocabulary teaching, learning, and assessment. All the researchers agree on these two points: firstly, language corpora provide access to extensive databanks that include authentic spoken and written discourse; secondly, with the assistance of relatively accessible concordancing software (Chen, & Ge, 2007; Kamasa, 2014), ESP practitioners can compile and analyse specialized corpora, examine the distribution of particular grammatical and lexical features, and even compare expert and learner texts (Belcher, 2012; Lesiak-Bielawska, 2015; Leláková & Toman, 2023).

The research confirmed that the introduction of corpus-based materials enhanced and innovated communication with students, fostering new types of interaction and providing valuable feedback (Durovič & Bauk, 2022). Similarly, it is claimed that such materials have the capacity to give a clearer insight into the vocabulary selected by students, offering teachers valuable feedback to identify gaps and other difficulties in vocabulary learning and instruction (Fraser et al., 2025).

Moreover, corpus-based materials employed promote independent work and encourage students' autonomy, which are exceptionally vital for their future careers, lifelong learning, and professional growth. Additionally, the materials address challenges arising from the varied language learning strategies preferred by individual students (Durović, 2021).

Some scholars go even further and argue that corpora and corpus-based learning are even more valuable and pertinent in the ESP context (more broadly within Languages for Specific Purposes (LSP) than in EFL, considering the distinctive features of ESP itself (Boulton et al.,

2012). The first feature describes a typical ESP learner who favors learner-centered, individualized, and autonomous learning approaches, including techniques such as hypothesis development, an inductive, problem-solving approach. The methods are more successful with adult learners, specifically ESP learners at universities. The second feature concerns the lexico-grammatical structure of ESP registers. The registers differ from general English through their selective use of particular structures, their frequent reliance on domain-specific phrasing, and the rapid pace of changes in ESP terminology to keep up with ongoing technical and professional developments. In the given context, corpora are indispensable for acquiring contextual information about the most recent ESP terminology, as compared to traditional textbooks and dictionaries (Boulton et al., 2012).

When discussing the development of individual language skills with the help of corpus-based materials, scholars do not have a unified opinion. It is claimed that these materials contribute to the enhancement of productive language knowledge, including vocabulary skills (Boulton et al., 2012), within the majority of reviewed studies (62%). Yet, offering only the core vocabulary lists as supplementary material is insufficient; emphasis should be placed on developing learners' reading skills (Hsiao-I, 2014). Alternatively, in (58%) of the investigated studies, it is asserted that they support the development of passive skills, reading and listening comprehension. Yet, such materials do not directly contribute to active skills, L1/L2 translation, speaking, and writing. Even these scholars emphasize that the utilization of authentic language in the ESP context is invaluable (Rudneva et al., 2019).

Considering the skill of writing, students are encouraged to explore various aspects of the language and to compare their own ESP writing with that of more experienced writers in their professional areas (Hundarenko, 2019). These studies demonstrate that using the corpus data can be highly effective in the classroom for increasing students' awareness (Flowerdew, 2009). However, some form of systematic guidance or mediation is often necessary to help learners interpret the data successfully.

Furthermore, the corpus approach offers a valuable opportunity to examine the concordances of specific terms within the corpus (Durović, 2021), as they include authentic examples of how the terms are used within phrases and sentences, which can be easily accessed using the same program. It is agreed that investigating concordances in context, along with information on their frequencies, stimulates the development of professional language (Rudneva et al., 2019). Moreover, for ESP learners, the process of disambiguating near-synonymous ESP words involves inferring information from specialized linguistic contexts. This approach enhances understanding, boosts engagement, and encourages motivation through self-discovery (Gavioli, 2005).

Ultimately, research indicates that language learners can gain advantages through a discovery-based learning approach, in which they actively build and utilize their own specialized corpora and vocabulary collections in the ESP context (Smith, 2020). Certain corpus-based materials can be developed in a way that promotes even learners' creativity in ESP learning (Niu & Dhamotharan, 2020).

RQ 2: What are the students´ perceptions and opinions regarding the teaching, learning, and assessment of ESP vocabulary with the help of corpus-based materials?

Much of the reviewed research has focused on learners’ receptivity to corpus-based materials for ESP needs, typically collected through interviews or questionnaires. Their results indicate that students exhibit an increased understanding of the interdisciplinary bonds between

ESP and disciplinary specialists (Lei & Liu, 2016). Simultaneously, they become aware of the increased differences in register, recognizing which lexical forms and discourse conventions are suitable in a given context.

The results of exploratory research studies indicate that students reported improvements in their writing skills, including ESP vocabulary, syntactic structure, and overall content. Additionally, their confidence in writing and their awareness of linguistic aspects were also heightened (Hundarenko, 2021; Tsai, 2021).

Students appreciate corpus-based materials as alternative methods of ESP vocabulary learning (compared to the traditional ones), together with their usefulness and efficiency (Durovič & Bauk, 2022). Students' autonomy and individual effort should be particularly emphasized in this process. The corpus approach enables students to complete at least a portion of their assignments independently, organizing their work and progressing at their own pace, while adhering to established deadlines.

On the other hand, findings of 75% of the reviewed studies indicate that many students feel uncertainty, and only under the supervision of the teacher can they learn ESP vocabulary more comprehensively and profoundly (Huang & Wang, 2023). The most negative students´ reactions are associated with more demanding and complex homework assignments that require additional time (Durovič & Bauk, 2022).

We also gathered students' opinions on the most effective way to organize future ESP vocabulary classes. The majority (66.5%) favoured a combined approach, integrating traditional textbook-based methods with the corpus method (Durovič & Bauk, 2022). In the given context, their abilities to summarize and analyse have been developed together with a strengthened learning initiative.

When discussing specific types of activities, students find motivating interactive exercises that can be designed using corpus analysis findings. For example, creating gap-fill activities helps them practice specific vocabulary areas, such as adjectives, verb tenses, and prepositions. Additionally, exercises can prompt ESP students to examine a particular discourse by identifying grammatical features and usage patterns of key vocabulary and phrases within a given corpus (Hsiao-I, 2014).

RQ 3: What are the challenges associated with corpus-based materials in terms of ESP vocabulary teaching, learning, and assessment?

The following section discusses the challenges directly related to the application of corpus-based materials in ESP vocabulary teaching.

Limited accessibility and resources are the primary problems that ESP teachers and learners face. Corpus-based materials often require special equipment, specialized software, and technical expertise, which may not be available to them. These can hinder the integration of corpus data into ESP vocabulary instruction. However, these challenges can be addressed by utilizing free online corpora and other user-friendly tools that learners can access through their mobile devices (Kern, 2013; Jaafar, 2022).

The thin line between the authenticity of corpus-based materials and their pedagogical relevance requires careful consideration as well. While corpora provide genuine language examples, selecting vocabulary that aligns with specific ESP contexts can be challenging. In the given context, teachers should ensure that corpus-derived vocabulary is both authentic and relevant to learners' professional needs (Durovič & Bauk, 2022).

The complexity of data analysis may represent another challenge, as interpreting corpus data involves understanding frequency patterns, the relationships between words revealed through collocational networks, and contextual usage. Learners and teachers may find it challenging to analyse and apply this information effectively, especially without prior training. Yet, the findings of Jaafar (2022) advocate that educators can effectively engage ESP students with the target language by utilizing corpus tools, which can help identify aspects that might be overlooked through manual analysis alone.

Integrating corpus-based methods into both teaching and testing is essential (Paker & Özcan, 2017). Yet, assessment difficulties may arise since designing assessments based on corpus data can be complex. Traditional testing methods may not effectively measure learners' ability to utilize corpus-derived vocabulary in authentic contexts, necessitating the development of innovative evaluation techniques.

Another focus of challenges in the given context has been on learner variability, which means that learners' familiarity with corpus tools and their analytical skills vary widely. This variability can impact the effectiveness of corpus-based vocabulary instruction and may require additional training and support. Granger (2009) contends that a thorough analysis of learner corpora is crucial for developing effective corpus-informed pedagogical materials tailored to ESP.

Bringing corpus-based findings into university ESP classes is understood by scholars as an essential task (Quero & Coxhead, 2018). Yet, incorporating corpus-based approaches into existing ESP courses/curriculum demands curriculum redesign and instructor training, which can be resource-intensive and time-consuming. Under these circumstances, maintaining up-to-date corpus materials that reflect current terminology and usage patterns is an ongoing challenge, as language use in specific fields evolves (Grosse & Voght, 2012; Handzilevska et al., 2023).

Using everyday words with specialized meanings can pose challenges for ESP teachers, as learners may struggle to grasp the new, context-specific definitions and concepts associated with words they already recognize in their general vocabulary (Coxhead, 2013). The key point is that this teacher's students have a well-established understanding of the everyday meaning of the word. In their university studies, these everyday words often acquire specialized meanings, transforming them into technical terms. This new, technical understanding requires students to develop knowledge of both the word's concept and its specific meaning within the scientific context. To facilitate this learning, teachers might employ various corpus-based techniques. After that, teachers might explore students' misconceptions about certain words by asking for their opinions on the meanings, identifying any misunderstandings, and then providing the scientific perspective and corpus-based context.

The results of studies based on the direct application of corpus-based materials for ESP students indicate that incomplete sentences with key words from the concordance output may pose a challenge for students. Moreover, the extensive number of concordances can seem overwhelming to non-linguistic students who are not familiar with inductive learning procedures (Rudneva et al., 2019).

Eventually, students frequently find it challenging to interpret corpus data, especially when analysing the more intricate phraseological elements of English. However, we can address these issues through careful design of teaching materials and assignments. As an example, a form of mediation is advocated within the reviewed studies when ESP students are prompted with a worksheet to notice specific linguistic features, to formulate rules, and consequently discuss in pairs or groups their findings.

Stoller and Robinson (2018) suggest the following procedure as a form of transition from general English to the language of the area they study. The first step is to recognize the key genres valued within the target disciplines; , then to analyse these genres to identify their distinctive linguistic, non-linguistic, and organizational features; after that, to select authentic texts that are relevant and accessible to students in terms of disciplinary content; and to design scaffolded instructional materials that initially help students develop observational skills, followed by more advanced tasks such as detailed analysis and writing exercises.

Another quest is connected to the acquisition of ESP vocabulary over time. We start from the premise that incorporating corpus-based data into material development and classroom activities can enhance learners' understanding of target texts in a second language (Gavioli, 2005). While the presentation of professional terminology through extensive bilingual lists may appear advantageous in the short term, it tends to be less effective over the long term due to difficulties in contextualizing the target vocabulary (Rudneva et al., 2019).

In summary, while corpus-based materials offer valuable insights into authentic vocabulary use in ESP, addressing these challenges is essential to maximize their effectiveness in teaching, learning, and assessment. As a way of example, (Hsiao-I, 2014) concludes that teachers assume two key roles in the process: firstly, as facilitators who are responsible for identifying and providing authentic, valuable resources that align with students' needs and learning objectives; and secondly, as guides who help students independently recognize language patterns and behaviors of words and phrases, fostering their development as language researchers. Additionally, teachers should focus on instructing specialized vocabulary relevant to students' subject areas. This goal can be achieved by supporting learners in developing general skills such as recognizing specialized words, interpreting definitions, connecting meanings to core concepts, and understanding word parts. By equipping students with effective vocabulary strategies, teachers can avoid directly teaching specific content areas and instead focus on guiding students to manage specialized vocabulary independently.

The integration of corpus approaches into ESP has brought about various legal and ethical challenges for educators (Bloch, 2013). A particularly contentious ethical issue concerns the use and distribution of intellectual property, as well as shifting attitudes toward plagiarism. Developing simple corpora for classroom use can also involve legal and financial considerations that ESP teachers may find challenging to navigate. The rights and restrictions related to intellectual property are typically governed by a combination of local and international laws and regulations. Nonetheless, these concerns impact teachers and students to differing degrees, especially within the context of an increasingly globalized society. Given that the definition of plagiarism varies across different settings, ESP teachers must be aware of these differences and learn how to incorporate an understanding of intellectual property rights into their teaching, ensuring ESP students are informed about the appropriate ways to use such resources.

Discussion

One of the most evident practical applications of corpus linguistic description is in language learning and teaching, particularly within the realm of English for Specific Purposes (ESP). ESP learners investigate opportunities provided by accessible language corpora; at the same time, corpora represent a challenge for ESP students. These aspects were the focus of this study, which conducted a systematic rapid review of the literature, examining peer-reviewed academic articles published between 2010 and 2025. It aimed to explore how corpus-based materials are employed

in ESP English language teaching, learning, and assessment, as well as the challenges associated with utilizing such materials in this area.

Regarding the research questions, the investigated studies supported various aspects of teaching ESP vocabulary with the application of corpus-based materials at universities.

RQ 1: How can corpus-based materials enhance ESP vocabulary teaching, learning, and assessment?

When speaking about RQ1, it was supported by majority of the studies reviewed; the authors believe that ESP vocabulary teaching, learning, and assessment can be enhanced with the utilization of corpus-based materials; this claim is in alignment with the research conducted on the effectiveness of such materials in the teaching process at universities (Aijmer, 2009; Flowerdew, 2005; Hundarenko, 2019).

The rapid review suggests that the future of corpus-based activities looks promising, offering ESP learners access to real-world contexts. As technology advances, corpus tools are expected to become more user-friendly, further enriching the ESP learning environment (Kern, 2013).

RQ 2: What are the students´ perceptions and opinions regarding the teaching, learning, and assessment of ESP vocabulary with the help of corpus-based materials?

Regarding RQ2, it must be acknowledged that the studies present a range of opinions on the application of corpus-based materials when working with ESP vocabulary.

Positively appreciated are mainly the increased perception of interdisciplinary bonds together with increased register awareness (Lei & Liu, 2016; Niu & Dhamotharan, 2020; Stoller & Robinson, 2018). Additionally, students feel that diverse linguistic competencies and writing skills are developed with the help of corpus-based materials (Curry & McEnery, 2025; Handzilevska, Tkachuk, Hundarenko & Lelakova, 2023).

Conversely, three-quarters of the studies suggest that numerous students feel uncertainty while working with corpus-based activities, especially complex homework assignments, and they admit that they can learn ESP vocabulary more effectively with a teacher (Hsiao-I, 2014; Paker & Özcan, 2017; Rudneva, Valeeva, Faizi, Guslyakova & Nigmatzyanova, 2019).

RQ 3: What are the challenges associated with corpus-based materials in terms of ESP vocabulary teaching, learning, and assessment?

The challenges identified within the rapid literature review might be summarized as follows. The primary challenge is limited access to resources, as they require specialized software, which can be overcome by utilizing free online corpora. Furthermore, a balance needs to be struck between the authenticity of these materials and their pedagogical relevance in relation to specific ESP vocabulary.

The complexity of analysing corpus data can pose difficulties for both teachers and university students; this could be alleviated by prior training. In addition to this, incorporating corpus-based materials into university ESP curricula requires curriculum redesign, which demands ongoing updates to reflect evolving language for specific purposes (Granger, 2009; Grosse & Voght, 2012; Quero & Coxhead, 2018).

Based on these results, further critical analysis is needed into the students' understanding of corpus-based materials for their ESP study needs and the challenges associated with their application. Addressing these challenges through considerate resource selection, teacher training, and ethical awareness is crucial for maximising the benefits of corpus-based activities in ESP vocabulary acquisition.

Corpus tools also assist in the evaluation of ESP course books (Chuying, 2019). It is evident from these processes that future compilation of ESP textbooks may need to pay attention to several important aspects. Firstly, it is essential to achieve a better balance among the various language skills. Listening, speaking, and related learning strategies should constitute a significant portion of the content, reflecting students' perceived needs and preferences—what they find necessary, insufficient, or desirable. It is also advisable to clearly specify and integrate sub-skills such as listening, speaking, reading, writing, vocabulary, and learning strategies within exercises and activities. Secondly, ESP textbooks should enhance their content regarding text features, recycling, organization, and difficulty levels. Improving the layout by incorporating illustrations can help capture students' interest. Additionally, it is recommended that language knowledge and skill training be revisited and recycled throughout the book, enabling students to reinforce and consolidate their learning. The content organization should follow a coherent syllabus that combines language points, skills, and subject matter, with the complexity of texts gradually increasing as ESP learners progress.

As the first comprehensive review of its kind in the field of ESP, the findings provide valuable insights for developing relevant academic policies across various ESP settings. They also offer pedagogical strategies and practical guidance for English language teachers, enabling them to leverage language corpora effectively and productively improve English language education in diverse ESP contexts.

The authors of the reviewed articles agree that learners should be provided with opportunities to acquire vocabulary implicitly through exposure to natural speech and everyday reading materials, thereby fostering their professional language. This approach helps them grasp the subtle nuances of word meanings, derivative forms, and contextual usage. In essence, it is vital to develop both learners' receptive (understanding) and productive (use) vocabulary skills. Namely, corpus-based teaching materials directly support this approach.

Research limitations

However, it is essential to acknowledge that the findings and conclusions are somewhat limited due to the small number of studies reviewed; as mentioned above, all the studies are published in English, and they were obtained solely from these major databases: Web of Science, Scopus, Google Scholar, ProQuest, Science Direct, and ERIC. Furthermore, the heterogeneity of included studies from different ESP fields, which makes the direct comparison more challenging, might be considered a limitation. Ultimately, this rapid review may be less reliable than a full systematic review due to time constraints and potential biases that can arise from the accelerated review process.

Future systematic reviews should incorporate a larger and more diverse range of studies from various ESP contexts across different regions. In this way, we could develop a more comprehensive understanding of the impact of language corpora on ESP learning, teaching, and assessment.

Conclusion

Advancements in technology have significantly broadened the array of options accessible to ESP teachers and heightened the significance of innovative teaching methods that provide ESP students with linguistic phenomena using extensive collections of machine-readable texts, known as corpora.

The weight of evidence from this rapid review of literature suggests that corpora can be beneficial for many learners in various contexts; they serve as a valuable addition to the ESP teacher's and learner's toolkit when employed thoughtfully and appropriately. To be more specific, the corpus-based approach to teaching ESP enhances linguistic competence by providing extensive data on real-life lexicogrammatical patterns, facilitating frequency-based insights into standard word choices, connotations, subtle differences in meaning, and suitable collocations.

Regarding the corpus itself, for example, when focused on university students, we can expand its scope by incorporating textbooks and potentially scientific articles. However, the latter might warrant creating a separate academic corpus specifically for analysing lexis. When focused on teachers, the corpus serves as a valuable reference resource, allowing them to verify their intuitions about the language before sharing them with their ESP students.

This study might be valuable, especially for English language teachers and ESP curriculum developers, as it provides comprehensive insights into how to employ corpus-based materials to develop students´ ESP vocabulary effectively. Moreover, this study presents insights that can inform subsequent linguistic research in this area. Future research in ESP vocabulary teaching may focus on creating core multiword academic vocabulary or core spoken academic vocabulary. Furthermore, the investigation of how cross-linguistic corpus comparisons can facilitate learners' understanding of differences and similarities between their native and target languages' terminology. And most recently, on the application of AI techniques to enhance the analysis and presentation of corpus ESP data.

Funding: This paper was published with the support of the Cultural and Educational Grant Agency of the Slovak Ministry of Education, project no. 005ŽU-4/2023.

Acknowledgments: This paper was published with the support of the Cultural and Educational Grant Agency of the Slovak Ministry of Education, project no. 005ŽU-4/2023.

Conflicts of Interest: The author declare no conflict of interest.

Authenticity: This manuscript is an original work.

Artificial Intelligence Statement: During the research process, the author utilized the Sketch Engine, a tool that employs natural language processing and computational linguistics techniques to facilitate the systematic analysis of language usage. Moreover, AI and AI-assisted technologies were used in this study for grammar and language refinement. After using these tools, the author reviewed and edited the content as needed. The author takes full responsibility for the accuracy and integrity of the content.

About the Author:

Marta Lacková works at the University of Žilina. Her research focuses on comparative and corpus linguistics. She has published two monographs on the grammatical category of determination and inter-lingual homonyms, two morphological textbooks intended for university learners, one lexicological textbook, and more than 40 scientific articles.
ORCID ID: (https://orcid.org/0000-0002-5784-7523)

References

Aijmer, K. (ed.). (2009). *Corpora and language teaching.* Amsterdam: John Benjamins.

Belcher, D.D. (2012). What ESP is and can be: An introduction. In D.D. Belcher (ed.), *English for specific purposes in theory and practice*, 1–20. Ann Arbor: University of Michigan Press.

Bloch, J. (2013). Technology and ESP. In B. Paltridge & S. Starfield (eds.), *The Handbook of English for Specific Purposes*, 407–425. John Wiley & Sons.

Boulton, A. (2016). Integrating corpus tools and techniques in ESP courses. *Concepts and frameworks in English for specific purposes,* 69, 119–137.

Boulton, A., Carter-Thomas, S., & Rowley-Jolivet, E. (2012). Issues in corpus-informed research and learning in ESP. Introduction to A. Boulton, S. Carter-Thomas & E.Rowley-Jolivet (eds), *Corpus-informed research and learning in ESP: Issues and applications* (pp. 1–14). Amsterdam: John Benjamins. DOI: 10.1075/scl.52.01bou.

Charles, M. (2013). English for academic purposes. In B. Paltridge & S. Starfield (eds.), *The Handbook of English for Specific Purposes*, 137–155. John Wiley & Sons.

Chen, Q., & Ge, G. (2007). A corpus-based lexical study on frequency and distribution of Coxhead's awl word families in medical research articles. E*nglish for specific purposes*, *26*(4), 502–514.

Chen, Y., & Liu, M. (2020). Designing corpus-informed activities for ESP learners' vocabulary development. *System*, 89, 102–105.

Chuying, O. (2019). A corpus tools-assisted evaluation of three ESP textbooks in China. *English language teaching*, *12*(6), 161–179.

Coxhead, A., Dang, T. N. Y., & Mukai, S. (2017). Single and multi-word unit vocabulary in university tutorials and laboratories: evidence from corpora and textbooks. *Journal of English for Academic Purposes*, 30, 66-78. DOI: 10.1016/j.jeap.2017.11.001.

Coxhead, A. (2025). Vocabulary and English for specific purposes. In Starfield, & C. A. Hafner, *The handbook for English for specific purposes.* https://doi.org/10.1002/9781119985068.ch6

Coxhead, A. (2013). Vocabulary and ESP. In B. Paltridge & S. Starfield (eds.), *The Handbook of English for Specific Purposes*, 115–132. John Wiley & Sons.

Curry, N., & McEnery, T. (2025). Corpus linguistics for language teaching and learning: a research agenda. *Language teaching*, *58*(2), 232–251. DOI:10.1017/S0261444824000430.

Durovič, Z., & Bauk, S. (2022). Testing corpus linguistics methods in ESP vocabulary teaching. *The journal of teaching English for specific and academic purposes. 10*(2), 319-337.

Durović, Z. (2021). Corpus linguistics methods for building ESP word lists, glossaries, and dictionaries on the example of a marine engineering word list. *Lexikos*, *31*(1), 259–282. DOI: https://doi.org/10.5788/31-1-1647.

Farhady, H., Tavassoli, K., & Irani, F.H. (2018). Selecting corpus-based grammatical structures for ESP/EAP materials. In Y. Kırkgöz & K. Dikilitaş (eds), *Key issues in English for specific purposes in higher education. English language education*, 77–92. Springer: Cham. DOI: https://doi.org/10.1007/978-3-319-70214-8_5.

Flowerdew, L. (2005). An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: countering criticisms against corpus-based methodologies. *English for specific purposes, 24* (3), 321–332.

Flowerdew, L. (2009). Applying corpus linguistics to pedagogy. *International Journal of Corpus Linguistics, 14*, p. 393–417.

Flowerdew, L. (2015). Corpus-based research and pedagogy in EAP: from lexis to genre. *Language teaching, 48*(1), 99–116. DOI: https://doi.org/10.1017/S0261444813000037.

Fraser, S., Higa, M. K., & Davies, W. (2025). Delivering an ESP pedagogic word list: integrating corpus analysis, materials design, and software development. *Languages, 10*(3), 46. DOI: https://doi.org/10.3390/languages10030046.

Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied linguistics, 35*(3), 305–327.DOI: https://doi.org/10.1093/applin/amt015.

Gavioli, L., & Benjamins, J. (2005). *Exploring corpora for ESP learning.* Amsterdam: John Benjamins.

Granger, S. (2009). The contribution of learner corpora to second language acquisition and foreign language teaching. In K. Aijmer (ed.), *Corpora and language teaching*, 13–32. Amsterdam: John Benjamins.

Grosse, C. U., & Voght, G. M. (2012). The continuing evolution of languages for specific purposes. *The Modern Language Journal, 96*, 190–202. DOI: http://www.jstor.org/stable/41478799.

Handzilevska, H., Tkachuk, O., Hundarenko, O., & Lelakova, E. (2023). Psychological features of educational bilingualism of Ukrainian and Slovak students through the prism of the resilience approach. *Inequality, informational warfare, fakes, and self-regulation in education and upbringing of youth, Youth Voice Journal, 1*, 39–48. https://www.rj4allpublications.com/product/psychological-features-of-educational-bilingualism-of-ukrainian-and-slovak-students/.

Hsiao-I, H. (2014). Teaching specialized vocabulary by integrating a corpus-based approach: implications for ESP course design at the university level. *English language teaching, 7*(5), 25–37. 10.5539/elt.v7n5p26.

Huang, L., & Wang, J. (2023). A corpus-based study on automobile English vocabulary teaching. *ICDEL '23: Proceedings of the 2023 8th International conference on distance education and learning*, 178-184. DOI: https://doi.org/10.1145/3606094.3606142.

Hundarenko, O. (2019). Challenges of teaching academic writing skills in an ESL classroom (based on international teaching experience). *Revista Romaneasca pentru educatie multidimensionala, 11(*4), 70-83.DOI: https://doi:10.18662/rrem/158.

Hundarenko, O. (2021). Students' voice on feedback and instructions in academic writing: (based on Erasmus teaching experience in Slovak and Hungarian universities). *Advanced education, 8(*19), 62–68. DOI: https://doi.org/10.20535/2410-8286.234216.

Jaafar, E.A. (2022). Collocation networks of selected words in academic writing: a corpus-based study. *Arab world English journal, 13*(1), 240–255. DOI: https://dx.doi.org/10.24093/awej/vol13no1.15.

Jitpainboon, J., & Rungswang, A. (2022). Trends of corpus linguistics used in English for specific purposes research: a case of an Asian ESP journal. *Journal of Institutional Research, South East Asia*, *20*(2), 147–160.

Kamasa V. (2014). Techniki językoznawstwa korpusowego wykorzystywane w krytycznejanalizie dyskursu. Przegląd. *Przegląd socjologii jakościowej*, *10*(2), 100-117.

Kern, N. (2013). Technology-integrated English for specific purposes lessons: real-life language, tasks, and tools for professionals. In G. Motteram (Ed.), *Innovations in learning technologies for English language teaching*, 87–115. British Council: London.

Laosrirattanachai, P. & Ruangjaroon, S. (2021). Corpus-based creation of tourism, hotel, and airline business word lists. *LEARN journal: language education and acquisition research network, 14*(1), 50–86.

Lei, L., & Liu, D. (2016). A new medical academic word list: a corpus-based study with enhanced methodology. *Journal of English for Academic Purposes, 22*, 42–53.

Lelakova, E., & Lackova, M. (2015). Application of the matrix method of presenting semantic relations in didactic practice. *Psychology research, 5*(1), 1-9.

Leláková, E., & Toman, M. (2023). A Corpus-based study of the use of commas in EFL written performance. *International journal of instruction. 16* (1) 1–20. DOI: 10.29333/iji.2023.1611a.

Lesiak-Bielawska, E. D. (2015). Key aspects of ESP materials selection and design. *English for Specific Purposes World, 46,* 1-26.

Mertler, C. A. (2019). Introduction to educational research (2nd ed.). Sage.

Martínez, A. I., Beck, C.S., & Panza, B.C. (2009). Academic vocabulary in agriculture research articles: a corpus-based study. *English for specific purposes, 28*(3), 183–198.

Muñoz, V. (2015). The Vocabulary of Semi-Popular Agriculture Articles in English: A Corpus-Based Study. *English for specific purposes, 45*(3), 26–44.

Niu, J., &. Dhamotharan, M. (2020). Developing medical learning materials to promote learners' creativity: a corpus-based case study. *Arab world English journal, 11*(4), 181–193. DOI: https://dx.doi.org/10.24093/awej/vol11no4.12.

Oleskeviciene, G.V., Mockiene L. & Stojković, N. (2021). *Corpus analysis for language studies at the university level*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021, 372. DOI: https://dx.doi.org/10.1136/bmj.n71.

Paker ,T., & Özcan, Y.E. (2017). The effectiveness of using corpus-based materials in vocabulary teaching. *International Journal of Language Academy, 5*(14), 62-81. DOI: https://doi.org/10.18033/ijla.3494.

Quero, B., & Coxhead, A. (2018). Using a corpus-based approach to select medical vocabulary for an ESP course: the case for high-frequency vocabulary. In Y. Kırkgöz, & K. Dikilitaş (eds) *Key issues in English for specific purposes in higher education. English language education,* 51-75). Springer: Cham. DOI: https://doi.org/10.1007/978-3-319-70214-8_4.

Römer, U. (2011). Corpus research applications in second language teaching. *Annual review of applied linguistics, 31*, 205-225. DPI: http://dx.doi.org/10.1017/S0267190511000055.

Rudneva, M., Valeeva, N., Faizi. R., Guslyakova, A., & Nigmatzyanova, Y. (2019). Corpus-based approach to ESP pedagogy in tertiary education. *Conference: 12th annual international conference of education, research and innovation,* 1597–1600. DOI: 10.21125/iceri.2019.0456.

Ryan, G. (2012). *Technical vocabulary acquisition through texts. A corpus and a case study in a theology classroom.* Lambert Academic Publishing.

Smith, S. (2020). DIY corpora for accounting & finance vocabulary learning. *English for Specific Purposes,* (57), 1–12. DOI: https://doi.org/10.1016/j.esp.2019.08.002.

Stoller, F. L., & Robinson, M. S. (2018). Innovative ESP teaching practices and materials development. In: Y. Kırkgöz & K. Dikilitaş (eds.), *Key issues in English for specific purposes in higher education*. Cham: Springer.

Sutarsyah, C, Nation, P., & Kennedy, G. (1994). How useful is EAP vocabulary for ESP? A corpus-based study. *RELC journal*, *25*(2), 34-50. DOI: http://dx.doi.org/10.1177/003368829402500203.

Tangpijaikul, M. (2014). Preparing Business Vocabulary for the ESP Classroom. *RELC journal*, *45*(1), 51-65. DOI: https://doi.org/10.1177/0033688214522641.

Tongpoon-Patanasorn, A. (2017). Constructing an academic word list of business English: A corpus-based approach. *Humanities and Social Sciences Journal, 34*(2), 1–31.

Tsai, YR. (2021). Exploring the effects of corpus-based business English writing instruction on EFL learners' writing proficiency and perception. *Journal of computing in higher education*, (33), 475–498. https://doi.org/10.1007/s12528-021-09272-4.

Vyatkina, N. & Boulton, A. (2017). Corpora in language learning and teaching. *Language learning and technology*, *21*(3), 1–8.

Walker, C. (2011). How a corpus-based study of the factors that influence collocation can help in the teaching of business English. *English for Specific Purposes*, *30* (2), 101–112.

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp. 289-316
DOI: https://dx.doi.org/10.24093/awej/vol16no4.17

Exploring Motivational Factors for Learning General and Medical English Among Early-Career Doctors in Ho Chi Minh City: A Mixed-Methods Study

Nghi, Tran Tin
Faculty of Foreign Languages
Ho Chi Minh City University of Industry and Trade, Ho Chi Minh city, Vietnam
Corresponding Author: nghitt@huit.edu.vn

Anh, Trang Hoang Nguyen
Institute of International Education
Hoa Sen University, Ho Chi Minh city, Vietnam
Email: dr.limederm@gmail.com

Received: 09/08/2025 Accepted: 11/08/2025 Published:12/10/2025

Abstract

English proficiency is critical for early-career medical doctors in Ho Chi Minh City to advance their professional and academic goals. Limited research exists on their motivation to learn General English and English for Medical Purposes. This study investigates the types, differences, and factors influencing motivation for learning general English and English for Medical Purposes among doctors graduating within three years. A mixed-methods approach was employed, involving a questionnaire (n=104) and semi-structured interviews (n=7) with early-career doctors in Ho Chi Minh City. Quantitative data were analyzed using SPSS for descriptive statistics and Wilcoxon Signed-Rank Tests, while qualitative data underwent thematic analysis to identify motivational factors. Instrumental motivation was the primary driver for both General English and English for Medical Purposes learning (mean=4.13, SD=0.60), followed by integrative motivation. Gender differences were significant, with females showing higher instrumental and integrative motivation for General English ($p<0.05$). Key motivational factors included career development, educational purposes, and cultural understanding, while time constraints and costs were barriers. Classroom environment and teachers had moderate influence. The findings highlight the dominance of instrumental motivation and the impact of gender and external factors on English learning. Educators should prioritize career-oriented, communicative teaching methods and authentic materials to enhance motivation. This study informs tailored English programs for medical doctors and suggests future research on time and cost influences.

Keywords: early-career doctors, English for medical purposes, general English, instrumental motivation, language learning motivation, medical education, Viet Nam|
How to cite: Nghi, T. T., & Anh, T. H. N. (2025). Exploring Motivational Factors for Learning General and Medical English Among Early-Career Doctors in Ho Chi Minh City: A Mixed-Methods Study. *Arab World English Journal, 16*(4):289-316.
https://dx.doi.org/10.24093/awej/vol16no4.17

Introduction

Motivation is essential in learning because it can impact actions and make them happen (Schunk & Usher, 2012). General English (GE) helps medical learners learn more effectively (Khansir & Dehkordi, 2017), while English for Specific Purposes (ESP) and English for Medical Purposes (EMP) give them attention and urge them to achieve their learning goals (Tomak & Šendula-Pavelić, 2017). Several studies have focused on the motivation to learn General English and ESP among medical students (Hosseini & Shokrpour, 2019; Mathis et al., 2021; Tomak & Šendula-Pavelić, 2017). Nonetheless, Martín-González and Chaves-Yuste (2024) claimed that the literature on motivation in English for Specific Purposes is insufficient. In particular, no studies have examined the motivation of medical students and doctors in learning both GE and EMP. Thus, it is important to investigate the type of motivation, levels of motivation, and motivational factors for learning English among doctors graduating three years or less in Ho Chi Minh City.

The study has three aims. The first is to investigate the level of motivation to learn English (GE and ESP) among medical doctors working in Ho Chi Minh City and graduating three years or less. Secondly, the author wants to find how different the kinds of motivation in learning GE and ESP of Ho Chi Minh City medical doctors graduating three years or less are. Finally, the study aims to find factors that impact the GE and ESP learning motivation of Ho Chi Minh City medical doctors graduating three years or less.

To achieve these objectives, this study aims to investigate the following questions:

1. *To what extent are Ho Chi Minh City medical doctors graduating three years or less motivated to learn GE and ESP?*
2. *How different are the kinds of motivations between learning GE and learning ESP among Ho Chi Minh City medical doctors graduating three years or less?*
3. *What factors influence the motivation of Ho Chi Minh City medical doctors who have graduated three years or less to learn GE and ESP?*

This study investigates medical doctors who have graduated from a medical university for three years or less and are now working in Ho Chi Minh City. The findings are expected to reveal the level and types of motivation and form a specific understanding of motivational factors in learning both GE and EMP among doctors in Ho Chi Minh City. Medical students or other medical staff, such as nurses, were not included because of the large differences in their working and learning environments.

Literature Review

English in the Contemporary World and in Vietnam

English is essential not only in public schools but also in private language institutions. There are many types of English, such as GE and ESP (Canese & Spezzini, 2023). GE is the most popular form, and the students who learn it believe that it will be useful for their future. ESP includes English for Medical Purposes, which looks into the academic and professional needs of medical learners and offers relevant terms and language skills (Alqurashi, 2016).

English is important because of its impact on career development in science and technology (H. T. M. Nguyen, 2012). It is the medium of communication, a requirement for a decent job, because English people do not suffer from language barriers (Andayani, 2022; Daif-Allah & Aljumah, 2020). Moreover, English plays a crucial role in education, scholarship, and university graduation (Andayani, 2022; N. Nguyen, 2012). Džuganová (2019) suggested that proficiency in English helped them broaden their horizons and offer opportunities for application. Misdiagnosis

leading to mortality could occur because of doctors' lack of English proficiency (Bharathi & Pushpanathan, 2022). Wahyuni (2021) indicated that all four English skills, especially listening and speaking, are necessary for medical doctors.

There are some problems when students learn both GE and ESP in universities, which are related to instructors, facilities, and learning techniques (Albright, 2018; Nguyen & Pham, 2016). In particular, ESP course books focused only on vocabulary and lacked English skills teaching. Maruyama (1996) noted that having poor vocabulary led to the belief that ESP vocabulary was not common and unnecessary in daily activities and that English teachers in ESP programs were not familiar with scientific terms. Furthermore, a lack of theoretical frameworks to support ESP education has been identified (Chen, 2011; Nguyen & Pham, 2016).

In Vietnam, medical students must learn 12 credits for foreign languages, and most of them choose English. Both the GE and EMP were taught. Students were at different levels, and the classes were mostly organized in the afternoon. The main teaching method is the grammar-translation technique, in which teachers translate everything from English into Vietnamese and pay attention to analyzing sentence structures for students to understand. Classroom activities solely involve learning vocabulary, grammar, reading comprehension, and writing.

Motivation for Learning a Foreign Language

Motivation is a complex aspect that makes it impossible to define it clearly (Gardner, 2006). Motivation is essential in learning English among medical students because it is a psychological factor that determines the learning process of learning and teaching (Khansir & Dehkordi, 2017). In the field of English for specific purposes, motivation, whether intrinsic or extrinsic, could have significant, positive, or moderate correlations with academic achievement (Bekai & Harkouss, 2018).

Motivation can be divided into four types: instrumental, integrative, extrinsic, and intrinsic. According to Brown (2014), there are two types of motivation: instrumental and integrative. In particular, instrumental motivation shows the desire to attain academic or occupational goals, such as passing tests and having good jobs. Meanwhile, integrative motivation reflects learners' liking of foreign cultures and their wishes for integration (Ulfaika & Arsiani, 2023). In addition, self-determination theory (Ryan & Deci, 2017) claims that learners could be influenced by both internal and external factors. Therefore, motivation can be divided into intrinsic and extrinsic motivations. Intrinsic motivation refers to the internal and self-determined stimuli related to human factors. By contrast, extrinsic motivation is linked to reasons other than inner satisfaction.

Previous Studies

Othman et al. (2018) conducted research on postgraduate Afghanistan students, and they concluded that students were extrinsically motivated, and that teachers, pedagogic styles and classroom activities were the main motivation factors in learning English.

Regarding general English for medical groups, Indriani (2020) showed that the factors that influence motivation are learners' personalities, educators, and teaching methods. In addition, Mathis et al. (2021) indicated that medical students' motivation was instrumental, showing that students wanted to use English in their future profession.

In terms of EMP, Tomak and Šendula-Pavelić (2017) indicated that healthcare students are more integratively motivated. Moreover, Hosseini and Shokrpour (2019) concluded that some

factors influence learning motivation in ESP classes, such as instructors' characteristics, learning materials, and personal aspirations.

Taking a closer look at the motivation for learning English in Vietnam, Nguyen (2019) found that non-English majored participants were highly motivated to learn English, and instrumental motivation was the dominant motivation. The author also suggested a link between the school year and parental role in English learning motivation. In Vietnam, Le et al. (2024) research on medical students related to learning both GE and EMP indicated that demotivational factors on students' motivation were linked to lack of instructors' responses, lack of communication in class, outdated ESP textbooks, and teachers' frequent inability to explain complex medical terms.

Although a wide range of studies have been conducted in the field of motivation in learning English, most have focused mainly general English Martín-González and Chaves-Yuste (2024) indicated that studies related to learning motivation mainly focused on GE but lacked literature in ESP. In Vietnam, no studies have investigated the motivation of medical doctors to learn English, including General English, and especially English for Specific Purposes.

Methodology

Research Design

This study used a mixed-methods approach. The quantitative phase used questionnaires, whereas the qualitative component of the study included in-depth interviews with doctors to explore their motivational factors.

Participants

The participants of this study were medical doctors currently working in Ho Chi Minh City who had graduated from university within three years. The participants had graduated from high school before 2018, so most of them learned in exam-oriented classes that lacked communication. They have to attend postgraduate studies or training programs to obtain medical licences. During this period, GE and EMP are needed for them to acquire knowledge, and they want to enhance their language skills so that they are eligible for their future jobs.

Data Collection

Quantitative phase

A Vietnamese version of the questionnaire was administered to 104 medical doctors. It took approximately 20 min to complete the survey.

Qualitative Phase

Individual, semi-structured interviews were conducted to verify the quantitative results. This phase included seven participants via Facebook Messenger. The information was then stored on the researcher's laptop and transcribed in English. Each interview lasted from to 15-25 minutes.

Data Analysis

Quantitative Analysis

Quantitative data were analyzed using SPSS version 20 (Statistical, Armonk, NY, USA). The participants' answers were examined for abnormalities and missing data. The results showed that all 104 questionnaires met the standards, with sufficient and complete answers.

The Cronbach's alpha test was performed to assess reliability. Then, a descriptive statistics test was conducted, including the calculation of the mean scores. Illustrations of descriptive

information are presented in tables and charts. The researcher divided the statements of the questionnaire into eight smaller themes to investigate the motivational factors.

Qualitative Analysis

Qualitative data were analyzed using thematic analysis. In particular, information from the interviews was transcribed and read to familiarize them with the content. The data were then grouped into themes and reviewed.

Results

Quantitative Data Analysis

Level of Motivation

Table 1. *Levels of motivation for learning general English (GE) and English for medical purposes (EMP) among early-career doctors in Ho Chi Minh City*

n=104	Maximum	Minimum	Mean	SD	Level of motivation
GE-INT	5,00	1,57	3,88	0,62	High
GE-INS	5,00	1,75	4,13	0,60	High
GE-EXT	5,00	1,57	3,25	0,67	Moderate
GE-INR	5,00	1,60	3,38	0,78	Moderate
EMP-INT	5,00	1,83	3,58	0,67	High
EMP-INS	5,00	2,25	4,13	0,60	High
EMP-EXT	5,00	1,43	3,31	0,67	Moderate
EMP-INR	5,00	1,60	3,37	0,74	Moderate

We first report descriptive statistics for all motivational subscales. Table 1 summarizes participant characteristics and the descriptive results for general English (GE) and English for medical purposes (EMP), indicating stronger instrumental motives relative to intrinsic and integrative motives, alongside time and workload as the most cited barriers.

The differences between kinds of motivation

Table 2. *Pairwise comparisons of motivation types for learning general English (GE) and English for medical purposes (EMP)*

	GE-INT	**GE-EXT**	**GE-INR**
GE-INS	$p^{(1)} < 0,0083$	$p^{(1)} < 0,0083$	$p^{(1)} < 0,0083$
GE-INT	-	$p^{(1)} < 0,0083$	$p^{(1)} < 0,0083$
GE-EXT	-	-	$p^{(1)} = 0,15$
	EMP-INT	EMP-EXT	EMP-INR
EMP-INS	$p^{(1)} < 0,0083$	$p^{(1)} < 0,0083$	$p^{(1)} < 0,0083$
EMP-INT	-	$p^{(1)} < 0,0083$	$p^{(1)} < 0,0083$
EMP-EXT		-	$p^{(1)} = 0,51$

As shown in Table 2, statistically significant differences were observed between GE-INS and GE-INT ($p< 0,0083$), GE-INS and GE-EXT ($p< 0,0083$), GE-INS and GE-INR ($p< 0,0083$), GE-INT and GE-EXT ($p< 0,0083$), GE-INT and GE-INR ($p< 0,0083$), EMP-INS and EMP-INT ($p< 0,0083$), EMP-INS and EMP-EXT ($p< 0,0083$), EMP-INS and EMP-INR ($p< 0,0083$), EMP-INT and EMP-INR ($p< 0,0083$), and EMP-INT and EMP-INR ($p< 0,0083$). However, no statistically significant difference was found between GE-EXT and GE-INR ($p=0,15$), and between EMP-EXT and EMP-INR ($p = 0,51$) according to the Wilcoxon Signed-Rank Test.

Table 3. *Differences in motivation types between general English (GE) and English purposes (EMP)*

GE	EMP	p value
GE-INS	EMP-INS	p[(1)] = 0,86
GE-INT	EMP-INT	p[(1)] < 0,05
GE-EXT	EMP-EXT	p[(1)] = 0,14
GE-INR	EMP-INR	p[(1)] = 0,96

Regarding the comparison between GE and EMP, the Wilcoxon Signed-Rank Test did not show statistically significant differences between GE-INS and EMP-INS (p = 0,86), GE-EXT and EMP-EXT (p= 0,14), or GE-INR and EMP-INR (p = 0,96) (Table 3). However, there was a significant difference between the GE-INT and EMP-INT groups (p < 0,05).

Factors that Influence Learning Motivation

Demographic Factors

Table 4a. *Influence of demographic factors on motivation for learning general English (GE)*

		GE-INS p		**GE-INT p**		**GE-EXT p**		**GE-INR p**	
		Mean ± SD	p-value	Mean ± SD	p-value	Mean ± SD	p-value	Mean ± SD	p-value
Gender	Male	3,98 ± 0,1	0,04 (a)	3,67 ± 0,1	0,002 (a)	3,21± 0,1	0,35 (a)	3,3 ± 0,12	0,38 (c)
	Female	4,24 ± 0,06		4,04 ± 0,07		3,29 ± 0,09		3,44 ±0,1	
Years after graduation	1	4,22 ± 0,1	0,65 (b)	3,85 ± 0,11	0,07 (b)	3,3 ± 0,14	0,41 (b)	3,33 ± 0,17	0,19 (b)
	2	4,22 ± 0,12		4,13 ± 0,11		3,38 ± 0,13		3,65 ± 0,17	
	3	4,06 ± 0,08		3,79 ± 0,09		3,18 ± 0,09		3,3 ± 0,1	
Postgradu-ate study	Yes	4,11 ± 0,07	0,51 (a)	3,85 ± 0,67	0,37 (a)	3,27 ± 0,08	0,96 (a)	3,47 ± 0,1	0,04 **(a)**
	No	4,16 ± 0,12		3,94± 0,13		3,21 ± 0,13		3,17 ± 0,11	
Years of studying English	<5	3,98 ± 0,11	0,32 (b)	3,75± 0,15	0,45 (b)	3,29 ± 0,14	0,74 (b)	3,18 ± 0,15	0,46 (b)
	5- <10	4,1 ± 0,11		3,81 ± 0,09		3,11 ± 0,11		3,38 ± 0,14	
	10-<15	4,05 ± 0,16		3,85± 0,17		3,24 ± 0,14		3,32 ± 0,15	
	>15	4,27 ± 0,09		4,00 ± 0,09		3,33 ± 0,13		3,51 ± 0,14	

Table 4b. *Influence of demographic factors on motivation for learning English for medical purposes (EMP)*

		EMP-INS		**EMP-INT**		**EMP-EXT**		**EMP-INR**	
		Mean ± SD	p-value	Mean ± SD	p-value	Mean ± SD	p-value	Mean ± SD	p-value
Gender	Male	4,00 ± 0,1	0,049 (a)	3,51 ± 0,1	0,31 (c)	3,24 ± 0,1	0,39 (a)	3,33 ± 0,11	0,56(a)
	Fem-ale	4,24 ± 0,07		3,64 ± 0,08		3,37± 0,09		3,4 ± 0,09	

ISSN: 2229-9327

Years after graduation	1	4,2 ± 0,1	0,98 (b)	3,61 ± 0,15	0,45 (d)	3,33 ± 0,14	0,90 (d)	3,52 ± 0,15	0,25(b)
	2	4,11 ± 0,14		3,73 ± 0,14		3,36 ± 0,14		3,52 ± 0,15	
	3	4,12 ± 0,08		3,52 ± 0,83		3,29 ± 0,09		3,25 ± 0,1	
Postgraudate study	Yes	4,17 ± 0,06	0,57 (a)	3,59± 0,08	0,83 (c)	3,36 ± 0,07	0,24 (c)	3,4 ± 0,08	0,59(a)
	No	4,05 ± 0,13		3,56 ± 0,13		3,19 ± 0,14		3,29 ± 0,15	
Years of studying English	<5	4,06 ± 0,14	0,68 (b)	3,42 ± 0,13	0,61 (d)	3,34 ± 0,15	0,73 (b)	3,21 ± 0,16	0,47(b)
	5- <10	4,15 ± 0,13		3,54 ± 0,14		3,21 ± 0,14		3,51 ± 0,14	
	10- <15	4,02 ± 0,15		3,6 ± 0,17		3,27 ± 0,13		3,23 ± 0,15	
	>15	4,23 ± 0,08		3,67 ± 0,1		3,39 ± 0,11		3,43 ± 0,13	

Table 5. *Influence of demographic factors on motivation for learning general English (GE) and English for medical purposes (EMP) (data sepeated b/w mean + SD)*

Group	GE-INS Mean	GE-INS SD	GE-INS p	GE-INT Mean	GE-INT SD	GE-INT p	GE-EXT Mean	GE-EXT SD	GE-EXT p	GE-INR Mean	GE-INR SD	GE-INR p	EMP-INS Mean	EMP-INS SD	EMP-INS p	EMP-INT Mean	EMP-INT SD	EMP-INT p	EMP-EXT Mean	EMP-EXT SD	EMP-EXT p
Gender Male	3.98	0.1	0.04	3.67	0.1	2	3.21	0.1	0.35	3.3	0.12	0.38	4.0	0.1	49	3.51	0.1	0.31	3.24	0.1	0.39
Gender Female	4.24	0.06		4.04	0.07		3.29	0.09		3.44	0.1		4.24	0.07		3.64	0.08		3.37	0.09	
Years after graduation 1	4.22	0.1	0.65	3.85	0.11	0.07	3.3	0.14	0.41	3.33	0.17	0.19	4.2	0.1	0.98	3.61	0.15	0.45	3.33	0.14	0.9
Years after graduation 2	4.22	0.12		4.13	0.11		3.38	0.13		3.65	0.17		4.11	0.14		3.73	0.14		3.36	0.14	
Years after graduation 3	4.06	0.08		3.79	0.09		3.18	0.09		3.3	0.1		4.12	0.08		3.52	0.83		3.29	0.09	
Postgraduate study Yes	4.11	0.07	0.51	3.85	0.67	0.37	3.27	0.08	0.96	3.47	0.1	0.04	4.17	0.06	0.57	3.59	0.08	0.83	3.36	0.07	0.24
Postgraduate study No	4.16	0.12		3.94	0.13		3.21	0.13		3.17	0.11		4.05	0.13		3.56	0.13		3.19	0.14	
Years of studying English <5	3.98	0.11	0.32	3.75	0.15	0.45	3.29	0.14	0.74	3.18	0.15	0.46	4.06	0.14	0.68	3.42	0.13	0.61	3.34	0.15	0.73
Years of studying English 5-<10	4.1	0.11		3.81	0.09		3.11	0.11		3.38	0.14		4.15	0.13		3.54	0.14		3.21	0.14	
Years of studying English 10-<15	4.05	0.16		3.85	0.17		3.24	0.14		3.32	0.15		4.02	0.15		3.6	0.17		3.27	0.13	
Years of studying English >15	4.27	0.09		4.0	0.09		3.33	0.13		3.51	0.14		4.23	0.08		3.67	0.1		3.39	0.11	

Tables 4 and 5 summarize the influence of demographic factors on GE and EMP motivation. Overall, no statistically significant differences were observed when the study compared the four motivation types of EMP or GE with demographic factors, such as gender, years after graduation, postgraduate study, and years of studying English. Nonetheless, the study found statistically significant differences between male and female in EMP-INS (p= 0,049), GE-INS (p= 0,04) and GE-INT (p=0,002) with female having higher mean values. In addition, the study reported a statistically significant difference in GE-INR between the two groups based on postgraduate studies (p=0,04), with the former having a higher mean value.

Thematic Factors

Table 6. *Thematic factors influencing motivation for learning general English (GE) and English for medical purposes (EMP)*

Sub-themes	Mean ± SD (level)	GE	Mean	SD	Level
Career and Professional Development	3,93 ± 0,54 (High)	INS1: Learning GE helps me to progress in my future career.	4,27	0,67	High
		INS2: Learning GE facilitates my work (such as talking to patients, diagnosing and giving treatment).	4,3	0,68	High
		INS3: Learning GE allows me to get a higher salary.	3,88	0,90	High
		INS8: Learning GE helps me conduct medical research and write medical papers.	4,27	0,84	High
		EXT3: I learn GE because my job requires GE.	3,69	0,89	High
		EXT5: I learn GE because my patients and clients use English.	3,18	0,95	Moderate
Social recognition and pressure	3,22 ± 0,82 (Moderate)	EXT4: Being able to use GE proficiently allows me to earn respect from other people.	3,69	0,89	High
		EXT7: I learn GE because my friends and coworkers learn English	2,79	1,00	Moderate
		EXT2: My parents want me to learn GE	3,23	1,20	Moderate
Educational purposes	3,88 ± 0,61 (High)	INS4: Learning GE is essential for my higher education.	4,13	0,88	High
		INS5: Learning GE helps me participate in medical conferences.	4,04	0,88	High
		INS6: Learning GE helps me pass exams.	3,80	0,87	High
		INS7: Learning GE helps me to grasp up-to-date knowledge.	4,35	0,66	High
		EXT1: I learn GE because it is a compulsory subject in university	3,12	1,1	Moderate
Classroom Environment and teachers	3,12 ± 1,03 (Moderate)	EXT6: The effort I put into GE learning largely depends on the quality of English classes.	3,12	1,03	Moderate
Cultural Understanding and Social Interaction	3,76 ± 0,63 (High)	INT1: Learning GE helps me understand foreign movies, music, books, etc.	4,30	0,72	High
		INT2: Learning GE helps me understand foreign cultures and ways of living.	3,74	0,90	High
		INT3: Learning GE gives me chances to make friends and stay connected with foreigners.	4,14	0,85	High
		INT5: Learning GE helps me get married to English-speaking people.	2,88	0,96	Moderate

ISSN: 2229-9327

Immigration and Living	4,02 ± 0,71 (High)	INT4: Learning GE allows me to travel to different places around the world.	4,14	0,85	High
		INT6: Learning GE helps me live in foreign countries	3,70	1,02	High
		INT7: Learning GE helps me integrate more easily into English-speaking communities.	4,23	0,70	High
Personal Interest and Enjoyment	3,23 ± 0,88 (Moderate)	INR1: I learn GE because I like it without any reasons.	3,34	0,98	Moderate
		INR2: Learning GE is a hobby for me.	3,17	1,00	Moderate
		INR4: I learn GE because I enjoy speaking English to other people.	3,19	1,04	Moderate
Self-challenge and Self-fulfillment	3,6 ± 0,81 (High)	INR3: I challenge myself by learning GE.	3,16	1,04	Moderate
		INR5: I learn GE because I feel satisfied when I can communicate in English without translation.	4,04	0,84	High
Sub-themes	Mean ± SD (Level)	EMP	Mean	SD	Level
Career and Professional Development	3,96 ± 0,61 (High)	INS9: Learning EMP helps me to progress in my future career.	4,13	0,75	High
		INS10: Learning EMP facilitates my work in hospitals (such as diagnosing and giving treatment).	4,12	0,79	High
		INS11: Learning EMP allows me to get a higher salary.	3,88	0,89	High
		INS16: Learning EMP helps me conduct medical research and write medical papers.	4,4	0,68	High
		EXT10: I learn EMP because my requires EMP.	3,84	1,02	High
		EXT12: I learn EMP because my patients and clients use English.	3,41	1,00	High
Social recognition and pressure	3,36 ± 0,86 (Moderate)	EXT11: Being able to use EMP proficiently allows me to earn respect from other people.	3,59	1,03	High
		EXT9: My parents want me to learn EMP	2,85	1,18	Moderate
		EXT14: I learn EMP because my friends and coworkers learn English.	2,95	0,94	Moderate
Educational purposes	3,98 ± 0,63 (High)	INS12: Learning EMP is essential for my higher education.	4,13	0,80	High
		INS13: Learning EMP helps me to get more opportunities to participate in medical conferences.	4,15	0,82	High
		INS14: Learning EMP helps me pass exams.	3,86	0,88	High

		INS15: Learning EMP helps me to grasp up-to-date medical knowledge and information.	4,41	0,71	High
		EXT8: I learn EMP because it is a compulsory subject in university.	3,33	1,07	Moderate
Classroom Environment and teachers	3,23 ± 0,98 (Moderate)	EXT13: The effort I put into English learning largely depends on the quality of English classes.	3,23	0,98	Moderate
Cultural Understanding and Social Interaction	3,67 ± 0,62 (High)	INT8: Learning EMP allows me to become a part of an international doctors' communities.	3,93	0,83	High
		INT9: Learning EMP helps me make friends with international doctors.	4,02	0,80	High
		INT10: Learning EMP helps me talk with English-speaking people.	4,06	0,81	High
		INT11: Learning EMP helps me get married to English-speaking people.	2,65	0,93	Moderate
Immigration and Living	3,42 ± 0,95 (High)	INT12: Learning EMP helps me live in foreign countries.	3,25	1,05	Moderate
		INT13: Learning EMP helps me integrate more easily into English-speaking communities.	3,59	1,03	High
Personal Interest and Enjoyment	3,21 ± 0,84 (Moderate)	INR6: I learn EMP because I like it without any reasons.	3,30	0,92	Moderate
		INR7: Learning EMP is a hobby for me.	3,15	0,94	Moderate
		INR9: I learn EMP because I enjoy speaking English to other people.	3,16	1,03	Moderate
Self-challenge and Self-fulfillment	3,61 ± 0,75 (High)	INR8: I challenge myself by learning EMP.	3,26	0,90	Moderate
		INR10: I learn EMP because I feel satisfied when I can communicate in English without translation.	3,95	0,93	High

As shown in Table 6, the highest mean for GE was for Immigration and Living (4,024 ± 0,71), followed closely by Career and Professional Development (3,93 ± 0,54), underscoring the aims, and educational purposes demonstrated strong motivational influence (3,88 ± 0,614).,a and Cultural Understanding and Social Interaction also played a significant role in learning motivation (3,764 ± 0,627). Moreover, Self-challenge and Self-fulfillment were effective factors in GE learning (3,6 ± 0,805). Social Recognition and Pressure and Personal Interest and Enjoyment had moderate motivational impact (3,22 ± 0,816 and 3,23 ± 0,88 respectively). Finally, the classroom environment and teachers affected motivation at the lowest level (3,12 ± 1,027).

In terms of EMP, Educational Purposes (3,98 ± 0,63) was the highest, followed by Career and Professional Development (3,96 ± 0,61). Cultural Understanding and Social Interaction and Self-challenge and Self-fulfillment were strong motivators (3,67 ± 0,62 and 3,61 ± 0,75 respectively). Immigration and living could motivate EMP learning, and social recognition and pressure aspects were reported to have a moderate influence on EMP learning motivation (3,36 ±

0,86). The classroom Environment and Teachers had only moderate motivational influence (3,23 ± 0,978), while Personal Interest and Enjoyment had the lowest motivational mean (3,21 ± 0,84),

Qualitative Data Analysis

General English

Career and Professional Development

Some participants (n=3) were aware that GE could facilitate their job and career.

"*I think English is a tool that will help everyone in the future, help me graduate and have a better job."* (Male 2)

Social Recognition and Pressure

The doctors in our study (n=2) indicated the influence of pressure. While one participant (female 3) considered pressure a deterrent to motivation, the other showed his love for a competitive class.

"I don't feel like learning English because I feel shy and ashamed if I don't read correctly, don't read well, and don't speak well especially, when I have to do [speak] so in such a crowded class." (Female 3).

"The classmates' level of English should be equal to yours so that you can feel a bit **z**

Educational Purposes

More than half of the participants (n=4) claimed that they had learned because of their education. The main reasons were to gain easier access to international information, IELTS knowledge, and overseas study.

"*The first reason is that I have easier access to more information from abroad. Because America or Europe will always have more updated sources of information, articles or scientific research, learning English will help you access those sources of information faster*." (Male 2)

"I studied for English exams, at that time I was studying for IELTS." (Male 2)

English will be a great opportunity, especially for those who want to study abroad." (Male 1)

Classroom Environment and Teachers

All participants (n=7) in our study indicated that the classroom environment and teachers could influence GE learning motivation. Both positive and negative factors were related to this aspect, as reported in our study. On the one hand, positive factors related to this aspect were comfortable learning environments, applicable teaching content, suitable materials, interaction during class, competitive games, positive teachers' attitudes, and teachers who were able to tailor to students'levels. On the other hand, based on their GE learning experience, dotors in the qualitative phase claimed that factors such as large class size, borining topics, traditional teaching styles, lack of communication, and negative teachers' attitudes negatively influence their motivation.

Cultural Understanding and Social Interaction

All the participants in the interview (n=7) indicated that cultural understanding and social interaction could influence their GE learning motivation. This is because GE allows them to communicate with other people, understand foreign ways of living, and express their feelings.

"It helps me discover many new cultures. I also love to communicate with people from many countries. The core of learning English is really a tool for me to exchange information." (Male 1)

"It helps me convey my information not only in Vietnamese but also in English. It helps me convey ideas that I feel are not suitable to say in Vietnamese. "When I am too angry or I want to express my feelings but find it a bit cheesy, I use English to make it more suitable for my personality." (Female 1)

Immigration and Living

Some participants (n=2) claimed that living abroad urged them to learn about GE. GE allowed doctors to travel freely and to settle in other countries.

"I see that wherever I travel, there is English. In general, it's easier for me to read anything. Then when I travel abroad, I feel a little more confident." (Female 3)

"The most important thing about learning English is my relationship with my intimate friends. It means a lot to my life because one day I have to stay in America." (Female 3)

Personal Interest and Enjoyment

Two participants (n=2) experienced factors related to personal interest and enjoyment. Being able to speak or listen to English directly gave them a sense of happiness.

"When I started to speak English, I felt... very satisfied." (Male 2)

"English can also help me entertain myself. I think everyone wants to watch movies or listen to music without reading subtitles. Hearing the character's voice makes me happy. I do not really like watching voiceover films. (Male 1)

Time and Cost

Time and cost are newly explored factors that influence GE learning motivation. Lack of time to study could prevent doctors from studying English, while tuition could have both positive and negative effects.

"The most important issue for everyone, I think, is usually time. Doctors like us do not have enough time. We have to work far away from home and study other subjects related to our profession, so we cannot allocate time to study English. Second, the cost issue can be a problem because many people are also afraid of the high cost of English classes." (Male 3).

"I like the fact that if I pay a big lump sum of money for an English course, I will have the motivation to get out of the house and go to school." (Male 1)

English for Medical Purposes

Career and Professional Development

Five participants showed that career and professional development influenced their motivation to learn about EMP. This is because they were able to treat foreign patients, increase their income, and apply for work. The need for EMP learning is essential when working in private medical settings.

"I think Medical English can be applied in my future work such as discussing medical issues with foreign doctors". (Male 3)

"*It can affect your income because working and knowing English will help you have more opportunities to examine and treat foreigners, which will also increase your income in your future work*." (Male 2)

"*Medical English can be used in my work such as examining patients and other related tasks, especially in private medical companies*." (Female 4)

Educational Purposes

Almost all the doctors (n=5) agreed that learning the EMP was for education. Their motivation to learn EMP stems from updating knowledge from books and articles, accessing new information, and obtaining scholarships.

"If your reading ability is perfect, you will shorten the time you spend on reading documents, which means faster and more effective knowledge absorbption." (Female 1)

"For medicine, updating the latest medical news is important. Therefore, accessing those sources of information through English is what I think is the core and necessary for those who are studying medicine." (Male 2)

"Learning English helps me find scholarships." (Female 2)

Classroom Environment and Teachers

All interviewees indicated the influence of classroom environment and teachers on EMP learning motivation.On the one hand, positive factors were young and friendly educators, adaptive teaching methods, applicable content and practical activities. However, based on their past learning experience, EMP learning motivation decreases when there are inapplicable and non-communicative teaching methods, large class size, lack of thorough explanation, poor English pronunciation, inauthentic materials, poor understanding of content knowledge, and mixed English accents.

Cultural Understanding and Social Interaction

Two doctors indicated that cultural understanding and social interaction could impact their learning motivation because they helped them communicate with colleagues and widen their social circle.

"Medical English helps me to broaden my social circle." (Female 4)

"Learning English helps me talk to my colleagues such as doctors and professors." (Female 2)

Immigration and Living

According to one participant, immigration and living helped her travel around the world.

Medical English helps me to travel around the world." (Female 4)

Personal Interest and Enjoyment

One participant in the study found an interest in learning the EMP. He felt a sense of contentment when being able to read and learn his major without Vietnamese translation.

Learning Medical English allows me to read content knowledge and learn my major directly from the source material without relying on translations. I am so happy for that. (Male 3)

Self-challenge and Self-fulfillment

One participant in our study expressed a sense of fulfillment when he realized his progress in learning English.

"*When I saw the progress in learning English I made over 30 days, I felt motivated to keep going*." (male 3)

Time and Cost

Similar to GE, EMP learning motivation was driven by time and cost.

"We do not have enough time to learn, although we really want to learn. We are too tired because of pressure at work. Finance is also an issue. English courses are quite expensive…" (Male 3)

Discussion

Types of Motivation

Doctors in Ho Chi Minh City were mainly instrumentally motivated for both GE and EMP learning. This result is in line with that of Y. Wang (2023), indicating that most medical students were instrumentally motivated. As for explanations, English was compulsory for medical postgraduates to participate in medical research activities and conduct high-quality studies (C. Wang, 2023).

Integrative motivation received a high mean score, which means that doctors wished to integrate into other nations and get to know the styles of international colleagues at work to facilitate collaboration. This was aligned with Yelahina and Fedchyshyn (2019) results, which stated that medical participants were more integratively motivated because of their aspiration to blend in with the target language community.

In terms of intrinsic and extrinsic motivation for learning GE and EMP, both types of motivation showed a moderate degree, with intrinsic motivation being higher than extrinsic. The results indicated that learning assisted their careers and education rather than interest and joy. Moreover, the level of extrinsic motivation for learning GE and EMP was also the lowest among the four, which means that medical doctors were minimally influenced by external factors, such as tests and teachers. This result is similar to that reported by Pavel (2020). According to Pavel (2020), general medicine students are intrinsically motivated rather than extrinsically motivated, and they are more inclined to learn English. One possible explanation is that medical students and doctors are aware of the importance of GE and EMP learning in allowing them to participate effectively in medical lectures and work efficiently with international patients and colleagues.

The similarity between the ranks of the four types of English learning motivation suggested that medical doctors were influenced by all four kinds of motivation, regardless of the English type. Medical doctors placed a high priority on learning both GE and EMP to achieve career progress and academic success or converse with foreigners. They had pragmatic motivation rather than being purely affected by interests or external factors such as exams and social expectations. Therefore, identical patterns between GE and EMP motivations suggest that both GE and EMP programs should focus on incorporating communication and career-oriented approaches.

The Differences between Kinds of Motivation

Statistically significant differences between the four types of motivation in both GE and EMP learning showed that they had different influences on doctors' GE and EMP learning processes. The result was similar to that of Nguyen (2019), who stated that non-Engish-majored university students had higher instrumental motivation than integrative motivation. The major distinctions between integrative and instrumental motivations compared to intrinsic and extrinsic motivations showed that doctors paid attention to the pragmatic purposes of GE and EMP, including career prospects and fluent communication with international colleagues and foreigners, rather than learning English because of love or external factors. However, the study reported no statistically significant difference between the extrinsic and intrinsic motivation of GE, and the

same pattern was reported for EMP learning. This result indicates that the two kinds of motivation affected GE and EMP equally to a moderate degree.

When comparing GE and EMP, our study witnessed that between GE and EMP learning, instrumental, extrinsic and intrinsic motivations were not no statistic and significant difference between GE and EMP 's instrumental, extrinsic and intrinsic motivatić. This indicates the same pattern of motivation between the two types of English. This means that regardless of the content of English learning, doctors focused on their careers and were minimally influenced by personal interests and external pressures. However, there was a statistically significant difference in integrative motivation between GE and EMP. The results indicate that doctors in the study were inclined to learn GE because of its help in learning about English culture and communicating with people through a wide range of topics. This is because GE covers a wide range of topics (Canese & Spezzini, 2023), whereas EMP deals with medical terminology and language (Alqurashi, 2016). Consequently, GE allows young medical doctors to integrate more easily into English-speaking communities.

Factors that Influence Motivation

Demographic Factors

Based on our study, gender, years after graduation, postgraduate study, and years of graduation did not fully or significantly impact both GE and EMP learning motivation. These similarities between the different demographic factors showed that doctors learned English because of factors other than demographic characteristics.

Regarding the role of gender in learning motivation, our results were in line with those of Alghamdi et al. (2023) and Tomak and Šendula-Pavelić (2017), indicating that no significant differences between the two sexes related to EMP learning motivation were present. However, that females had significantly higher GE-INS, GE-INT, and EMP-INS. These results were similar to the conclusion of Daif-Allah and Aljumah (2020), who showed that females were statistically and significantly motivated to a higher degree than their male counterparts. As for the explanation, according to Saidi and Al-Mahrooqi (2012), it was reasonable to see that males were less motivated to learn English than their female counterparts because females were more enthusiastic and had higher self-efficacy.

In terms of years after graduation, no statistically significant differences were reported between years after graduation and learning motivation. This suggests that leaving school and being exposed to the working environment did not decrease or increase doctors' English learning motivation. This also means that medical participants can maintain a high level of learning motivation for use in their jobs, education, and integration.

The results were similar to those of Tomak and Šendula-Pavelić (2017), who claimed that there were no significant differences in intrinsic motivation between groups below 27 and those aged 27 or over. However, there were major differences in integrative and instrumental motivation between the two groups, with younger participants being more motivated. Thus, future studies should include a larger range of years after graduation, and explore contexts that maintain or diminish the drive to learn English**.**

Looking at postgraduate study and its effect on motivation, there were no significant differences between postgraduate study and GE or EMP learning motivations. This showed that doctors were motivated to study English, regardless of their academic level. Esanova and Buriyev

(2023) reported that medical articles and conferences in English will change rapidly; therefore, the demand for learning GE and EMP will influence doctors at any stage of their profession.

Regarding the influence of years studying English, there were no statistically significant differences between years of studying English and GE or EMP learning motivation. The results were consistent with the conclusion of Nguyen (2019) study, which indicated that non-English major students were not motivated regardless of exposure time to English. One explanation is that non-English major students, including medical doctors, experienced uninteresting language learning, which did not positively influence GE and EMP learning motivation.

Thematic Factors

Career and Professional Development

In our study, doctors were highly motivated to study GE and EMP for career purposes. In particular, GE not only allowed medical doctors to use it at work for treatment, but it also helped them to progress in their careers and receive pay raise. Mathis et al. (2021) supported this finding because the author concluded that being able to work with medical staff urged medical participants to study GE. This is because the medical environment requires a high level of accuracy, based on teamwork and updated knowledge. Another reason is that, according to Nazarieh et al. (2023), proficiency in English allows medical doctors to choose suitable medical techniques and increases treatment success. The study further suggested that English is a requirement to work in the medical field, which was also presented in(Y. Wang, 2023)'s study, indicating that medical students with high English proficiency can find jobs easily.

Regarding EMP learning motivation, Career and Professional Development ranked second. The doctors in our study indicated that they learned the EMP to help them with their work. This was supported by the results of Katsara (2008) indicated that the motivation for learning ESP was due to its advantages for future jobs. To be more specific, EMP could facilitate their hospital tasks, offer higher salaries, and allow career advancement. This is in line with the results ofTomak and Šendula-Pavelić (2017). The author indicated that medical students learn EMP because of its benefits in professional progress, better work, and high-paid occupations.

Social Recognition and Pressure

Social Recognition and Pressure in GE learning motivation had a moderate motivational influence. This suggests that doctors could earn respect from other people if their GE is good. English plays a crucial role in acquiring new knowledge and fostering collaboration in the medical field. Thus, English proficiency can also improve personal image in the eyes of others.

In our study, social recognition and pressure aspects moderately influenced EMP learning motivation. More specifically, earning respect from other people could motivate doctors to learn the EMP. This means that doctors knew that their reputation would increase if they could use their EMP proficiently. Tomak and Šendula-Pavelić (2017) also concluded that gaining greater respect from others could affect EMP learning motivation.

Educational Purposes

Educational purposes demonstrated a strong impact on motivation. In particular, medical doctors learned GE because it was highly needed for their higher education, passing exams, and acquiring the latest knowledge. Y. Wang (2023) agreed that improving English skills allowed medical students to participate in international education and comprehend medical articles and textbooks written in English. In fact, the United States Medical Licensing Examination (USMLE)

or Australian Medical Council (AMC) examinations are conducted in English; therefore, proficiency in GE helps them understand the common structures used in the exams.

For EMP learning motivation, the highest mean was for Educational Purposes. Doctors in our study learned EMP because it was important for their education. Katsara (2008) also highlighted that almost all ESP students were motivated to learn because they considered ESP useful for their study. In particular, from our study, EMP helped doctors take part in medical conferences, grasp up-to-date knowledge, and pass exams. There are several reasons for this result. First, medicine requires life-long learning (Hachoumi et al., 2023). Second, English is the most commonly used language to update medical information. Finally, the USMLE and AMC are English tests that require test takers to understand not only basic English but also medical terms to solve clinical problems. Therefore, doctors have no choice but to learn EMP intensively to gain new knowledge and pass standardized medical tests.

Classroom Environment and Teachers

The classroom environment and teachers had the lowest motivational means for GE and EMP learning motivation. The quantitative and qualitative phases indicated some factors related to motivation, such as teachers, teaching methods and styles, materials, activities, and poor understanding of content knowledge.

First, teachers' attitudes can influence motivation and positive attitudes can promote learning motivation. Huang (2022) concluded that among a variety of educators, positive characteristics such as mindfulness and compassion could reduce learners' negative feelings and increase motivation.

Second, teaching methods and styles play a crucial role in determining student motivation. Our study indicated that it was necessary for English teachers to understand their students and realize students' problems on time to prevent demotivation. Several studies have reported results similar to those of our study. Arroyyani et al. (2024) indicated that demotivation occurred when teachers delivered lessons in a disorganized manner when they failed to recognize students' problems. Le et al. (2024) claimed that student motivation could decrease if teachers were unable to respond to students' questions. In addition, our study reported a high demand for classroom communication. Katsara (2008) suggested that the communicative approach increases ESP learners' motivation and interest in broadening their horizons. Le et al. (2024) and Meng (2022) also identified lack of communication in class as a demotivating factor, showing that when teachers focused on introducing medical terminologies too much, they could not let students participate in classroom activities and develop four English skills, especially speaking and listening.

Third, the materials and content of the class are important. It was indicated from our interview that doctors in our study needed practical situations, time to share personal experiences of using medical terms in real life, how to apply them, and medicine-related materials. This is similar to the results of Katsara (2008) highlighted the importance of using authentic materials so that students can become familiar with English culture. Moreover, Le et al. (2024) and Meng (2022) showed that outdated ESP textbooks could adversely affect learning motivation.

In addition, activities in the class and learning environment could influence learning motivation. Our study showed that doctors were involved in a project, discussion with friends, or playing games, rather than doing grammar exercises or repeating words after teachers. Indriani (2020) claimed that, for higher learning motivation in EMP classes, the use of appropriate activities is necessary. Katsara (2008)and Y. Wang (2023) suggested that ESP learners should be exposed to group work and games to foster group cohesion and provide students with interactive and

communication environments. Finally, interviews revealed that a large class would pose a hindrance to learning motivation. This finding was supported by Erlina et al. (2020), who claimed that overcrowding could interfere with attention and concentration.

In addition, for low EMP learning motivation, teachers' poor understanding of content knowledge, which led to a lack of thorough explanation, was the culprit. Le et al. (2024) indicated that demotivation could arise from teachers' frequent inability to explain complex medical terms due to their insufficient understanding of analogus references. Our study also indicated that mixed English accents, including both American and British, could hinder EMP learning motivation among doctors.

Cultural Understanding and Social Interaction

Cultural Understanding and Social Interaction also play a significant role in GE learning motivation. Our study showed that doctors were eager to understand foreign cultures from books, music, and movies, as well as by making international friends. Our results were in contrast to those of Mathis et al. (2021), who showed that medical students did not learn English for sole entertainment or enjoy English-language music better. This is because medical students in Mathis et al. (2021) study had yet to realize the importance of cultural understanding and social interaction as doctors in our study.

Regarding EMP learning motivation, Cultural Understanding and Social Interaction were strong motivators**.** EMP helps participants in our study establish relationships with international doctors and allows them to become a part of the global medical community. Tomak and Šendula-Pavelić (2017) proposed a similar point in that doctors learned EMP because they were interested in interacting with members of international medical communities.

Immigration and Living

Among the eight sub-themes related to learning GE, Immigration and Living had the highest means. Doctors in our study prioritized learning English because it could help them travel, adapt, and integrate into a new country. The same pattern has been reported in other studies. The highest-scoring motivational item for learning English is that which helps medical students enjoy traveling abroad (Mathis et al., 2021).

Immigration and living reasons can also motivate doctors to learn about EMP. Our research showed that learning EMP could largely assist doctors in integrating more easily into English-speaking communities. The results were similar to those of Yelahina and Fedchyshyn (2019) andTomak and Šendula-Pavelić (2017), showing that medical students learn EMP so that they can travel, get to know other nations and cultures, and integrate into the target language society.

Personal Interest and Enjoyment

Personal Interest and Enjoyment had the lowest motivational mean in EMP and nearly the lowest in GE, indicating that doctors are less likely to learn EMP or GE for pleasure or entertainment. This is true because EMP are designed for professional use and education rather than entertainment, and they mainly focus on their careers.

Self-challenge and Self-fulfillment

Self-challenge and Self-fulfillment were high in GE's learning motivation. This indicates that mastering GE is linked to personal achievement. In particular, the lack of translation services poses a hindrance to healthcare access (Gil-Salmerón et al., 2021). Therefore, the ability to communicate without translation allows doctors to provide immediate and timely treatment.

Similar to GE learning, motivation, Self-challenge and Self-fulfillment were rated highly for EMP learning. In our study, doctors felt a great sense of personal satisfaction when they communicated in English without translation. This indicates that misleading translations can lead to adverse medical events (Karwacka, 2014). Misunderstanding of medical terms could decrease patients' trust and negatively impact the treatment plan (Anazawa et al., 2012). Therefore, if doctors could become both practitioners and interpreters, they could reduce fatality and create fluent doctor-doctor and doctor-patient communication.

Time and Cost

Time and cost are newly explored factors that influence doctors' learning motivation. While the time constraint to study was negative for learning motivation for both GE and EMP, high tuition fees could be negative for both types of motivation and exclusively positive for GE. Regarding the time limit to study, there have been no studies investigating the impact of a lack of time to study learning motivation. Wilcox et al. (2016) suggested that businesses can reduce the time spent on finishing tasks. Ahmad et al. (2024) concluded that high tuition fees could lead to low academic performance, motivation, and confidence. However, Beneito et al. (2018) showed that students who pay fees in full could pass one module with a minimal number of registered times, pass courses in the first time more easily, and improve their academic grades compared with those who were exempted from paying fees. Therefore, we suggest that further studies focus on investigating how time constraints and tuition fees influence learning motivation.

Future Research Directions

Further research should focus on newly explored themes, such as the influence of time and cost of learning on doctors' English learning motivation. They should look into the changes in motivation over time as doctors progress in their professions and the impacts of different teaching methods, geographical regions, and medical specialties on motivation.

Conclusion

The study was carried out on medical doctors who had graduated within three years in Ho Chi Minh City, and a mixed method approach was applied. The findings revealed that for doctors in the study, regardless of whether they learned GE or EMP, instrumental motivation was the primary driving force, followed by integrative motivation. Doctors were intrinsically and extrinsically on a moderate level. The study also found that gender differences existed in some kinds of motivation for learning GE and EMP, and postgraduation could influence intrinsic motivation when doctors learn GE. Most of the results of the interviews were identical to those of the quantitative phase. However, some issues related to poor understanding of content knowledge, negative attitudes toward teachers, and dull teaching methods that lack communication were identified as negatively influencing English learning motivation. Overall, the study suggests that future English programs should focus on career development, academic learning, and communication when teaching GE and EMP to medical doctors. Moreover, real-world applications, authentic and updated materials, and interactive activities are required to enhance English learning motivation. The findings of our study lay the foundation for not only future research, but also for choosing suitable teaching methods and creating effective English programs for medical doctors.

Funding Statement: This research received no specific grant or financial support from any funding agency in the public, commercial, or not-for-profit sectors.
Ethics Statement: This study was approved by the Ethics Committee of Ho Chi Minh City University of Industry and Trade and Hoa Sen University. All participants provided informed consent, and the research adhered to ethical guidelines for studies involving human subjects, ensuring confidentiality and voluntary participation.
Conflict of Interest Statement: The authors declare no affiliations with or involvement in any organization or entity with a financial interest in the subject matter or materials discussed in this manuscript.
Author Contributions Statement: TTN and THNA designed and implemented the research, conducted the analysis of results, and wrote the manuscript. TTN conceptualized the original idea and supervised the project.

About the Authors:
Dr. Tran Tin Nghi is an experienced educator and researcher interested in second language acquisition and pedagogy. His work focuses on developing and implementing effective teaching methodologies, using corpus-based techniques in language teaching, AI in Education, and the role of literacy in second language learning. He has a strong record of scholarly publications in prestigious Scopus-indexed journals. His commitment to ongoing professional development is evident through his active involvement in leading organizations like VietTESOL, STESOL, and VietCALL. Nghi Tran can be contacted at nghitt@huit.edu.vn. https://orcid.org/0000-0001-6549-3895

Trang Hoang Nguyen Anh, Dermatologist, MSc in English Language Teaching. Trang Hoang Nguyen Anh obtained his Master's degree from Hoa Sen University and completed his Dermatology Residency at Pham Ngoc Thach University of Medicine. His research interests include IELTS preparation and English for Medical Purposes. https://orcid.org/0009-0002-4341-2243

References
Ahmad, A., Hasan, M., & Ghorbanpour, M. (2024). Education cost as a new fickle in higher education for students learning via quantitively multinomial logistic regression. *Scientific Reports*, *14*(1), 1–15. https://doi.org/10.1038/s41598-024-81074-x

Albright, J. (2018). *English tertiary education in Vietnam*. Routledge.

Alghamdi, A., Albawardi, A., Alzuabi, N., & Alshaiji, L. (2023). Does Gender Matter? Motivation and Learning EFL: A Saudi Case Study. *International Journal of Arabic-English Studies*, *23*(1).

Alqurashi, F. (2016). English for medical purposes for Saudi medical and health professionals. *Advances in Language Literary Studies, 7*(6), 243–252. https://doi.org/10.7575/aiac.alls.v.7n.6p.243

Anazawa, R., Ishikawa, H., & Kiuchi, T. J. T. (2012). The accuracy of medical interpretations: A pilot study of errors in Japanese-English interpreters during a simulated medical scenario. *Translation & Interpreting: The International Journal of Translation and Interpreting Research*, *4*(1), 1–20.

Andayani, E. S. (2022). The importance of learning and knowing English in higher education in Indonesia. *Research Development Journal Of Education*, *8*(1), 372–379. https://doi.org/10.30998/rdje.v8i1.13315

Arroyyani, R., Setianingsih, I. S., & Sari, D. M. (2024). ESP Students' Demotivation In Learning English: A Study Of An Islamic Education Study Program. *Journal on Education*, *7*(01), 1056–1065.

Bekai, W., & Harkouss, S. (2018). The role of motivation in teaching speaking in esp: a case study of a Lebanese Private University. *Journal of Teaching English for Specific and Academic Purposes*, *6*(1), 141–150. https://doi.org/10.22190/JTESAP1801141B

Beneito, P., Boscá, J. E., & Ferri, J. (2018). Tuition fees and student effort at university. *Economics of Education Review*, *64*, 114–128.

Bharathi, K. N., & Pushpanathan, T. (2022). English for medical purposes and its status in India. In *Proceedings of the International Conference on Best Innovative Teaching Strategies (ICON-BITS 2021)* (pp. 221-224). ICON-BITS. https://doi.org/10.2139/ssrn.4027451,

Brown, H. D. (2014). *Principles of language learning and teaching: A course in second language acquisition*. Pearson.

Canese, V., & Spezzini, S. (2023). *Teaching english in global contexts: Language, Learners and Learning*. ISL-FIL-UNA.

Chen, Y. (2011). The institutional turn and the crisis of ESP pedagogy in Taiwan. *Taiwan International ESP Journal*, *3*(1), 17–30. https://doi.org/10.6706/TIESPJ.2011.3.1.2

Daif-Allah, A. S., & Aljumah, F. H. (2020). Differences in Motivation to Learning English among Saudi University Students. *English Language Teaching*, *13*(2), 63–74.

Džuganová, B. (2019). Medical language-A unique linguistic phenomenon. *Jahr–European Journal of Bioethics*, *10*(1), 129–145. https://doi.org/10.21860/j.10.1.7

Erlina, D., Desvitasari, D., Marzulina, L., & Risfina, A. (2020). Students' Demotivating Factors in English Language Learning at One State Madrasah Tsanawiyah in Palembang, South Sumatera, Indonesia. *Indonesian Research Journal in Education*, 120–136. https://doi.org/10.22437/irje.v4i1.9024

Esanova, M., & Buriyev, D. (2023). The challenges in teaching english to medical students. *Евразийский журнал академических исследований*, *3*(10), 255–260.

Gardner, R. C. (2006). *The socio-educational mole of second language learning*. Rowley, Massachusetts: Newbury House.

Gil-Salmerón, A., Katsas, K., Riza, E., Karnaki, P., & Linos, A. (2021). Access to Healthcare for Migrant Patients in Europe: Healthcare Discrimination and Translation Services. *18*(15), 7901. https://doi.org/10.3390/ijerph18157901

Hachoumi, N., Eddabbah, M., & El Adib, A. R. (2023). Health sciences lifelong learning and professional development in the era of artificial intelligence. *International Journal of Medical Informatics*, *178*, 105171. https://doi.org/https://doi.org/10.1016/j.ijmedinf.2023.105171

Hosseini, A., & Shokrpour, N. (2019). Exploring motivating factors among Iranian medical and nursing ESP language learners. *Cogent Arts & Humanities*, *6*(1), 1634324. https://doi.org/10.1080/23311983.2019.1634324

Huang, J. (2022). The role of English as a foreign language teachers' mindfulness and compassion in fostering students' foreign language enjoyment. *Frontiers in Psychology*, *13*, 899298. https://doi.org/10.3389/fpsyg.2022.899298

Indriani, K. S. (2020). Factors influencing motivation in learning English of Mangusada Badung General Hospital staff. *OKARA: Jurnal Bahasa dan Sastra*, *14*(1), 13–22. https://doi.org/10.19105/ojbs.v14i1.2972

Karwacka, W. (2014). Quality assurance in medical translation. *The Journal of Specialised Translation*, *21*, 19–34.

Katsara, O. (2008). Aspects of motivation within the context of an ESP course. *ESP World Journal*, *7*(3), 1–26.

Khansir, A. A., & Dehkordi, F. G. (2017). The role of motivation in the teaching and learning English. *Modern Journal of language teaching methods*, *7*(9), 072–085.

Le, H. V. L., Ngo, T. T., & Dang, G. N. (2024). Students' Perceptions of Demotivating Factors in Learning English for Specific Purposes at a Military Medical University in Vietnam. *Journal of Language Teaching and Research*, *15*(5). https://doi.org/10.17507/jltr.1505.08

Martín-González, D., & Chaves-Yuste, B. (2024). From English for general purposes to English for specific purposes: The role of motivation in higher education in Spain. *ESP Today*, *12*(1), 26–48.

Maruyama, H. (1996). Difficulties in teaching technical English in Japan. *LFE: Revista de lenguas para fines específicos*, *3*(1), 215–232.

Mathis, B. J., Mayers, T., & Miyamasu, F. (2021). English as a vocational passport: Japanese medical students and second language learning motivation. *Education Sciences*, *12*(1), 8. https://doi.org/10.3390/educsci12010008

Meng, Y. J. T. (2022). English learning demotivating factors among non-english majors in medical university. *Theory and Practice in Language Studies*, *12*(5), 931–936. https://doi.org/10.17507/tpls.1205.13

Nazarieh, M., Delzendeh, M., & Beigzadeh, A. (2023). Motivation as an integral factor in English language learning for medical students. *Res Dev Med Educ*, *12*(1), 7–7. https://doi.org/10.34172/rdme.2023.33133

Nguyen, H. C. (2019). Motivation in learning English language: a case study at Vietnam national university, Hanoi. *European Journal of Educational Sciences*, *6*(1), 49–65. https://doi.org/10.19044/ejes.v6no1a4

Nguyen, H. T. M. (2012). Primary English language education policy in Vietnam: Insights from implementation. In R. Baldauf, R. Kaplan, K. Kamwangamalu, & P. Bryant (Eds.), *Language planning in primary schools in Asia* (pp. 121–145). Routledge.

Nguyen, H. T. T., & Pham, M. T. T. (2016). Difficulties in teaching English for specific purposes: Empirical study at Vietnam universities. *Higher Education Studies*, *6*(2), 154–161. https://doi.org/10.5539/hes.v6n2p154

Nguyen, N. (2012). How English has displaced Russian and other foreign languages in Vietnam since "Doi Moi". *International Journal of Humanities Social Science*, *2*(23), 259–266.

Othman, N. A., Manap, M. R., Ramli, N. F. M., & Kassim, A. A. M. (2018). Motivation in English language Learning: A Case Study of Afghanistan Students in Malaysia. In *Proceedings of the 4th International Conference ILANNS 2018* (pp. 377-393). ILANNS.,

Pavel, E. (2020). Language learner motivation and strategies in English for medical purposes. *Philobiblon*, *25*(1), 125–138.

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford publications.

Saidi, A. A., & Al-Mahrooqi, R. (2012). The influence of gender on Omani college students' English language learning strategies, comprehension and motivation. *International Journal of Applied Linguistics and English Literature*, *1*(4), 230–244.

Schunk, D. H., & Usher, E. L. (2012). Social cognitive theory and motivation. In R. M. Ryan (Ed.), *The Oxford handbook of human motivation* (Vol. 2, pp. 13–27). Oxford University Press. https://doi.org/10.1016/j.cedpsych.2019.101832

Tomak, T., & Šendula-Pavelić, M. (2017). Motivation towards studying English for specific purposes among students of medical and healthcare studies. *Jahr–European Journal of Bioethics*, *8*(2), 151–170. https://doi.org/10.21860/j.8.2.1

Ulfaika, R., & Arsiani, D. (2023). The Instrumental and Integrative Motivation of Students in Blended Learning Class at Eight Graders of Junior High School. *Conference on English Language Teaching*, 653–672.

Wahyuni, S. (2021). English language needs for medical students: A link and match of academic and professional career. *ENGLISH FRANCA: Academic Journal of English Language Education for Health*, *5*(1), 169. https://doi.org/10.29240/ef.v5i1.2146

Wang, C. (2023). On improving the English language ability of Science Citation Index papers for medical postgraduates through flipped classroom teaching. *Korean journal of medical education*, *35*(1), 85. https://doi.org/doi: 10.3946/kjme.2023.251

Wang, Y. (2023). Motivation of Chinese Medical Students in Learning English as a Second Language: A Quantitative Study. *Creative Education*, *14*(7), 1358–1369.

Wilcox, K., Laran, J., Stephen, A. T., & Zubcsek, P. P. (2016). How being busy can increase motivation and reduce task completion time. *Journal of personality and social psychology*, *110*(3), 371. https://doi.org/10.1037/pspa0000045

Yelahina, N., & Fedchyshyn, N. (2019). Motivation as a key component of teaching English for medical students. *Медична освіта*(4), 107–112.

Appendices

Appendix A. Doctors in Ho Chi Minh City were surveyed to assess their level and motivation to learn English.

Questionnaires

BACKGROUND INFORMATION OF PARTICIPANTS
1.Gender 1. Male 2. Female 3. Not want to mention
2.Year of birth
3.Years of graduation from medical university 1 year 2 years 3 years > 3 years
4.Currently studying/ Finish postgraduate programs in medicine Yes No

5.Time of learning English
1. < 5 Years
2. 5 -< 10 years
3. 10 - < 15 years
4. ≥ 15 years

MOTIVATION
1 = strongly disagree
2 = partly disagree
3.= neutral
4 = partly agree
5 = strongly agree
INTEGRATIVE MOTIVATION

Order	Category	CODING	Statement	1	2	3	4	5	Vietnamese translation
1	GE	INT1	Learning GE helps me understand foreign movies, music, books, etc.						Học tiếng Anh tổng quát giúp tôi hiểu phim, nhạc, sách nước ngoài, v.v.
2	GE	INT2	Learning GE helps me understand foreign cultures and ways of living						Học tiếng Anh tổng quát giúp tôi hiểu văn hóa và lối sống nước ngoài
3	GE	INT3	Learning GE gives me chances to make friends and stay connected with foreigners						Học tiếng Anh tổng quát cho tôi cơ hội kết bạn và kết nối với người nước ngoài
4	GE	INT4	Learning GE allows me to travel to different places around the world						Học tiếng Anh tổng quát cho phép tôi đi du lịch đến nhiều nơi khác nhau trên thế giới
5	GE	INT5	Learning GE helps me get married to English-speaking people						Học tiếng Anh tổng quát giúp tôi kết hôn với người nói tiếng Anh
6	GE	INT6	Learning GE helps me live in foreign countries						Học tiếng Anh tổng quát giúp tôi sống ở nước ngoài
7	GE	INT7	Learning GE helps me integrate more easily into English-speaking communities						Học tiếng Anh tổng quát giúp tôi hòa nhập dễ dàng hơn với cộng đồng nói tiếng Anh
8	ESP	INT8	Learning EMP allows me to become a part of an international doctors' communities						Học tiếng Anh Y khoa cho phép tôi trở thành một phần của cộng đồng bác sĩ quốc tế
9	ESP	INT9	Learning EMP helps me make friends with international doctors						Học tiếng Anh Y khoa giúp tôi kết bạn với các bác sĩ quốc tế
10	ESP	INT10	Learning EMP helps me talk with English-speaking people						Học tiếng Anh Y khoa giúp tôi nói chuyện với người nói tiếng Anh
11	ESP	INT11	Learning EMP helps me get married to English-speaking people						Học tiếng Anh Y khoa giúp tôi kết hôn với người nói tiếng Anh
12	ESP	INT12	Learning EMP helps me live in foreign countries						Học tiếng Anh Y khoa giúp tôi sống ở nước ngoài

13	ESP	INT13	Learning EMP helps me integrate more easily into English-speaking communities						Học tiếng Anh Y khoa giúp tôi hòa nhập dễ dàng hơn với cộng đồng nói tiếng Anh
INSTRUMENTAL MOTIVATION									
Number	Category	CODING	Statement	1	2	3	4	5	Vietnamese translation
1	GE	INS1	Learning GE helps me to progress in my future career						Học tiếng Anh tổng quát giúp tôi thăng tiến trong sự nghiệp tương lai
2	GE	INS2	Learning GE facilitates my work (such as talking to patients, diagnosing and giving treatment)						Học tiếng Anh tổng quát tạo điều kiện thuận lợi cho công việc của tôi (chẳng hạn như nói chuyện với bệnh nhân, chẩn đoán và điều trị)
3	GE	INS3	Learning GE allows me to get a higher salary						Học tiếng Anh tổng quát giúp tôi có được mức lương cao hơn
4	GE	INS4	Learning GE is essential for my higher education						Học tiếng Anh tổng quát là điều cần thiết cho giáo dục đại học của tôi
5	GE	INS5	Learning GE helps me participate in medical conferences						Học tiếng Anh tổng quát giúp tôi tham gia các hội thảo y khoa
6	GE	INS6	Learning GE helps me pass exams						Học tiếng Anh tổng quát giúp tôi vượt qua các kỳ thi
7	GE	INS7	Learning GE helps me to grasp up-to-date knowledge						Học tiếng Anh tổng quát giúp tôi nắm bắt được những kiến thức cập nhật
8	GE	INS8	Learning GE helps me conduct medical research and write medical papers.						Học tiếng Anh tổng quát giúp tôi tiến hành nghiên cứu y học và viết các bài báo y khoa.
9	ESP	INS9	Learning EMP helps me to progress in my future career						Học tiếng Anh Y khoa giúp tôi thăng tiến trong sự nghiệp tương lai
10	ESP	INS10	Learning EMP facilitates my work in hospitals (such as talking to patients, diagnosing and giving treatment)						Học tiếng Anh Y khoa tạo điều kiện thuận lợi cho công việc của tôi tại bệnh viện (chẳng hạn như nói chuyện với bệnh nhân, chẩn đoán và điều trị)
11	ESP	INS11	Learning EMP allows me to get a higher salary						Học tiếng Anh Y khoa giúp tôi có được mức lương cao hơn
12	ESP	INS12	Learning EMP is essential for my higher education						Học tiếng Anh Y khoa là điều cần thiết cho giáo dục đại học của tôi
13	ESP	INS13	Learning EMP helps me to get more opportunities to participate in medical conferences						Học tiếng Anh Y khoa giúp tôi có thêm cơ hội tham gia các hội thảo y tế
14	ESP	INS14	Learning EMP helps me pass exams						Học tiếng Anh Y khoa giúp tôi vượt qua kỳ thi
15	ESP	INS15	Learning EMP helps me to grasp up-to-date medical knowledge and information.						Học tiếng Anh Y khoa giúp tôi nắm bắt được những kiến thức, thông tin y tế cập nhật nhất.
16	ESP	INS16	Learning EMP helps me conduct medical research and write medical papers.						Học tiếng Anh Y khoa giúp tôi tiến hành nghiên cứu y học và viết các bài báo y học.

ISSN: 2229-9327

EXTRINSIC MOTIVATION									
Order	Category	CODING	Statement	1	2	3	4	5	Vietnamese translation
1	GE	EXT1	I learn GE because it is a compulsory subject in university						Tôi học tiếng Anh tổng quát vì đây là môn học bắt buộc ở trường đại học
2	GE	EXT2	My parents want me to learn GE						Bố mẹ tôi muốn tôi học tiếng Anh tổng quát
3	GE	EXT3	I learn GE because my job requires GE						Tôi học tiếng Anh tổng quát vì công việc của tôi yêu cầu tiếng Anh tổng quát
4	GE	EXT4	Being able to use GE proficiently allows me to earn respect from other people						Việc có thể sử dụng tiếng Anh tổng quát thành thạo cho phép tôi nhận được sự tôn trọng từ người khác
5	GE	EXT5	I learn GE because my patients and clients use English						Tôi học tiếng Anh tổng quát vì bệnh nhân và khách hàng của tôi sử dụng tiếng Anh
6	GE	EXT6	The effort I put into GE learning largely depends on the quality of English classes.						Công sức tôi bỏ ra cho việc học tiếng Anh tổng quát phần lớn phụ thuộc vào chất lượng các lớp học tiếng Anh.
7	GE	EXT7	I learn GE because my friends and coworkers learn English						Tôi học tiếng Anh tổng quát vì bạn bè và đồng nghiệp của tôi học tiếng Anh
8	ESP	EXT8	I learn EMP because it is a compulsory subject in university						Tôi học tiếng Anh Y khoa vì đây là môn học bắt buộc ở trường đại học
9	ESP	EXT9	My parents want me to learn EMP						Bố mẹ tôi muốn tôi học tiếng Anh Y khoa
10	ESP	EXT10	I learn EMP because my job requires EMP						Tôi học tiếng Anh Y khoa vì công việc của tôi yêu cầu tiếng Anh Y khoa
11	ESP	EXT11	Being able to use EMP proficiently allows me to earn respect from other people						Việc có thể sử dụng tiếng Anh Y khoa thành thạo giúp tôi nhận được sự tôn trọng từ người khác
12	ESP	EXT12	I learn EMP because my patients and clients use English						Tôi học tiếng Anh Y khoa vì bệnh nhân và khách hàng của tôi sử dụng tiếng Anh
13	ESP	EXT13	The effort I put into English learning largely depends on the quality of English classes.						Công sức tôi bỏ ra vào việc học tiếng Anh phần lớn phụ thuộc vào chất lượng các lớp học tiếng Anh.
14	ESP	EXT14	I learn EMP because my friends and coworkers learn English						Tôi học tiếng Anh Y khoa vì bạn bè và đồng nghiệp của tôi học tiếng Anh
INTRINSIC MOTIVATION									
Order	Category	CODING	Statement	1	2	3	4	5	Vietnamese translation
1	GE	INR1	I learn GE because I like it without any reasons						Tôi học tiếng Anh tổng quát vì tôi thích nó mà không cần lý do
2	GE	INR2	Learning GE is a hobby for me						Học tiếng Anh tổng quát là một sở thích của tôi

3	GE	INR3	I challenge myself by learning GE						Tôi thử thách bản thân bằng cách học tiếng Anh tổng quát
4	GE	INR4	I learn GE because I enjoy speaking English to other people						Tôi học tiếng Anh tổng quát vì tôi thích nói tiếng Anh với người khác
5	GE	INR5	I learn GE because I feel satisfied when I can communicate in English without interpretation/ translation .						Tôi học tiếng Anh tổng quát vì tôi cảm thấy hài lòng khi có thể giao tiếp bằng tiếng Anh mà không cần phiên dịch/biên dịch.
6	ESP	INR6	I learn EMP because I like it without any reasons						Tôi học tiếng Anh Y khoa vì tôi thích nó mà không cần lý do
7	ESP	INR7	Learning EMP is a hobby for me						Học tiếng Anh Y khoa là một sở thích của tôi
8	ESP	INR8	I challenge myself by learning EMP						Tôi thử thách bản thân bằng việc học tiếng Anh Y khoa
9	ESP	INR9	I learn EMP because I enjoy speaking English to other people						Tôi học tiếng Anh Y khoa vì tôi thích nói tiếng Anh với người khác
10	ESP	INR10	I learn EMP because I feel satisfied when I can communicate in English without interpretation/ translation .						Tôi học tiếng Anh Y khoa vì tôi cảm thấy hài lòng khi có thể giao tiếp bằng tiếng Anh mà không cần phiên dịch/biên dịch.

INFLUENCING FACTORS

	It influenced me to learn English
Exciting learning materials	
Being able to gain access to materials	
Teacher’s personality	
Teacher’s teachings styles	
Practical lessons	
Working environments that include English	

Appendix B. Interview questions

English	Vietnamese
In the first part we will talk about "GENERAL ENGLISH"	Phần đầu chúng ta sẽ nói về "TIẾNG ANH TỔNG QUÁT"
1. Based on your general English learning experience, what makes you excited about learning English? Can you please tell me about it?	1. Dựa trên trải nghiệm học tiếng Anh tổng quát của bạn, điều gì khiến bạn thấy thích thú với việc học tiếng Anh? Bạn có thể vui lòng kể về nó giúp mình.
2. Based on your general English learning experience, what makes you feel unsatisfied with learning English? Can you please tell me about it?	2. Dựa trên trải nghiệm học tiếng Anh tổng quát của bạn, điều gì khiến bạn thấy chưa hài lòng với việc học tiếng Anh? Bạn có thể vui lòng kể về nó giúp mình.

3. How do you think the dissatisfaction with learning GENERAL English can be changed so that you enjoy learning GENERAL English more?	3. Bạn nghĩ rằng những điều chưa hài lòng với việc học tiếng Anh TỔNG QUÁT có thể được thay đổi như thế nào để bạn thích học tiếng Anh TỔNG QUÁT hơn?
4. What elements do you hope a GENERAL ENGLISH class in the future will have so you can enjoy learning?	4. Bạn mong muốn một lớp TIẾNG ANH TỔNG QUÁT trong tương lai sẽ có yếu tố gì để bạn có thể thích thú học ?
In the following part we will talk about "MEDICAL PROFESSIONAL ENGLISH"	Phần sau chúng ta sẽ nói về "TIẾNG ANH CHUYÊN NGÀNH Y KHOA"
1. Based on your experience learning MEDICAL English, what makes you interested in learning English? Can you please tell me about it?	1. Dựa trên trải nghiệm học tiếng Anh Y KHOA của bạn, điều gì khiến bạn thấy thích thú với việc học tiếng Anh? Bạn có thể vui lòng kể về nó giúp mình.
2. Based on your experience learning MEDICAL English, what makes you feel unsatisfied with learning English? Can you please tell me about it?	2. Dựa trên trải nghiệm học tiếng Anh Y KHOA của bạn, điều gì khiến bạn thấy chưa hài lòng với việc học tiếng Anh? Bạn có thể vui lòng kể về nó giúp mình.
3. How do you think the things you are not satisfied with learning MEDICAL English can be changed so that you enjoy learning MEDICAL English more?	3. Bạn nghĩ rằng những điều chưa hài lòng với việc học tiếng Anh Y KHOA có thể được thay đổi như thế nào để bạn thích học tiếng Anh Y KHOA hơn?
4. What elements do you hope a future MEDICAL ENGLISH class will have that you can enjoy learning?	4. Bạn mong muốn một lớp TIẾNG ANH Y KHOA trong tương lai sẽ có yếu tố gì để bạn có thể thích thú học ?

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp.317-329
DOI: https://dx.doi.org/10.24093/awej/vol16no4.18

The Synergistic Effect of Cognitive and Emotional Intelligence in Foreign Language Learning: Concepts and Pedagogical Perspectives

M. Pınar BABANOĞLU
Department of Foreign Languages, Mersin University, Turkey
Corresponding Author: mpinar@mersin.edu.tr

Reyhan AĞÇAM
Department of Foreign Languages, Kahramanmaraş Sütçü İmam University, Turkey
Email: reyhanagcam@gmail.com

M. Fatih ADIGÜZEL
Department of Foreign Languages, Mersin University, Turkey
Email: fatihadiguzel@mersin.edu.tr

Tuçe ÖZTÜRK KARATAŞ
Department of Foreign Languages, Mersin University, Turkey
Enail: ecut14@gmail.com

Esin DÜNDAR
Department of Foreign Languages, Mersin University, Turkey
Enail: dundar.esin@gmail.com

Received:08/27/2025 Accepted: 11/25/2025 Published:12/10/2025

Abstract
This conceptual review examines the synergistic relationship between cognitive intelligence and emotional intelligence in the context of second/ foreign language learning. This relationship is considered vital in terms of revealing the effects of these intelligences on language education from a pedagogical perspective. Cognitive intelligence involves analytical reasoning, memory, and linguistic aptitude which are all significant factors for understanding grammar, vocabulary and language structure. Contrarily, emotional intelligence comprises self-awareness, empathy, motivation and social competence that play a pivotal role in effective communication, building confidence, and demonstrating cultural sensitivity. The dynamic interplay between cognitive and emotional intelligences considerably influences learners' overall success. Emotional intelligence components such as motivation and emotional regulation enhance cognitive performance, including attention and memory which particularly important for maintaining focus, managing anxiety, and fostering resilience when faced with challenging language tasks. Empathy, another core emotional intelligence element, facilitates intercultural and pragmatic competence, thereby enriching communicative effectiveness. This review highlights the importance of a holistic approach to language instruction that recognizes the dual effect of both intelligences in fostering linguistic proficiency, communicative competence, and learner engagement. It also invites our attention to the inquiry of foreign language learning environments that simultaneously involve activities designed with cognitive and emotional intelligence components for various age groups.

Keywords: affective factors, cognitive intelligence, emotional intelligence, foreign language learning, individual differences

How to cite: BABANOĞLU, M. P., AĞÇAM, R., ADIGÜZEL, M.F., KARATAŞ,T. O., & DÜNDAR, E.(2025). The Synergistic Effect of Cognitive and Emotional Intelligence in Foreign Language Learning: Concepts and Pedagogical Perspectives. *Arab World English Journal, 16*(4): 317-329. https://dx.doi.org/10.24093/awej/vol16no4.18

Introduction

Language is a basic tool used in everyday communication that seems effortless. Beneath its apparent simplicity lies a complex and dynamic system encompassing phonology, grammar, semantics, pragmatics and discourse (Dronkers et al., 2019; Fromkin et al., 2017). It is deeply rooted in social and cultural contexts, facilitating interaction among living beings and shapes and reflects individual identities, relationships, and collective worldviews (Kramsch, 1998; Gee, 2014). Reflecting upon this paradox of complexity and simplicity, foreign language learning is the natural extension or dimension of language itself. From this point of view, learning a second or foreign language (L2 hereafter) is a multifaceted process which involves cognitive and affective dimensions such as mental and emotional operations under-taken by learners in their pursuit of proficiency. Unlike first language (L1 hereafter) acquisition, which mostly occurs subconsciously during early childhood, L2 learning typically requires deliberate effort, conscious attention to grammar rules, and formal instruction (Ellis, 1994; Krashen, 1982). It is well evidenced that L2 learning is influenced by a range of cognitive, social and affective variables including age, motivation, working memory, metalinguistic awareness and exposure (Lenneberg, 1967; Lightbown & Spada, 2006; Miyake & Friedman, 1998; Robinson, 2002). Given this complexity, it becomes essential to explore the cognitive and affective factors that significantly influence L2 learning. Cognitive intelligence (CI) related-cognitive skills such as memory, attention, executive functions as cognitive flexibility, and overall CI determine how effectively learners can process, store, and retrieve linguistic input (Skehan, 1998). Affective factors, including motivation, anxiety, self-confidence, and attitudes toward L2 are equally important, since they shape learners' emotional responses and engagement with the learning process (Dörnyei, 2005; Gardner, 1985). In this context, Emotional intelligence (EI) defined as the ability to recognize, understand, and manage one's own emotions and those of others (Salovey & Mayer, 1990), emerges as a critical component to regulate affective factors such as foreign language anxiety which may inhibit language learning process (Krashen, 1982). Learners with higher EI often have better chances to cope with the challenges of language learning, manage anxiety, maintain motivation, and build positive relationships in social and communicative contexts (Dewaele et al., 2008; Pisghadam, 2009). Likewise, it was shown that general intelligence accounts only for 25% of life and academic success (Yurdakul et al., 2008), while non-cognitive-traits such as motivation, grit and EI skills like self-awareness play a vital role in the process as well. All in all, CI and EI contribute to learners' language achievement together (Taheri et al., 2019) with a more holistic understanding of what makes language learners successful, emphasizing that language learning is not only a mental activity but also a deeply human and affectively driven experience. Hence, this study evaluates the foundational roles of CI and EI in L2 learning, examining their inner mechanisms to demonstrate that both domains are interrelated and equally significant in contributing to L2 learner success.

The Present Study

With a holistic perspective, this study which is a conceptual review aims to present the synergistic effect of cognitive intelligence (CI) and emotional intelligence (EI) in the context of second/ foreign language learning drawing on various concepts and pedagogical perspectives. Based on this general purpose, this conceptual literature review which is organized around four sections specifically focuses on

1. Various definitions and foundations of the conceptual term of *Intelligence*,

2. The main cognitive processes and functions of second/foreign language learning considering the cognitive aspects of intelligence and language aptitude
3. Emotional intelligence to understand the affective aspect of L2 learning
4. Lastly, the synergistic relationship between cognitive and emotional intelligence in Foreign Language Learning shedding light on the cognitive and affective language areas in the brain

Literature Review

Understanding Intelligence: Definitions and Foundations

The adjective 'intelligent' is commonly used to describe people who have a set of abilities to understand complex ideas, think clearly, learn quickly, solve problems and adapt to the environment effectively. The term is derived from everyday observations of human behaviour and reflects our tendency to value mental sharpness, adaptability and insight –qualities that form the basis of the broader concept of intelligence. Indeed, the Encyclopaedia Britannica (2010) defines it as the ability to learn, adapt, and use knowledge effectively in various contexts and is often measured using IQ (from "intelligence quotient"), a number used to express the relative intelligence of a person. In education, it is viewed as the ability to learn, understand, and respond to new challenges, while in psychology, it refers more specifically to application of knowledge, abstract thinking, and adaption to one's environment. Intelligence involves a complex interplay of mental functions used to understand, adapt to, shape, and select different environments, though its definition is still under debate (Neisser et al., 1996; Sternberg, 2012).

Intelligence primarily manifests a set of cognitive abilities, including reasoning, problem-solving, and memory, which individuals use to learn from experience and interact effectively with their surroundings. The concept of a general cognitive ability, mostly associated with the term 'g factor', introduced by Spearman (1927), is widely accepted as a core component underlying overall intellectual performance (Jacobs et al., 2007). It is broadly acknowledged that intelligence arises from a combination of inherited traits and environmental influences (Plomin & Spinath, 2004), yet human cognitive ability is one of the most inherited dimensions (Jacobs et al., 2007). Our cognitive abilities allow us not only to acquire and build knowledge, solve problems and adapt to new situations, but also to process information, identify patterns and apply abstract thinking. Accordingly, general intelligence is the broad capacity, while CI is to use existing knowledge and cognitive processes such as attention, learning, memory, and reasoning. However, the ability of processing emotional information is also seen as a form of intelligence (Mayer et al., 2008) that addresses perceiving, understanding and managing emotions (Bru-Luna et al., 2021).

In language learning, cognitive abilities shape the mental framework which plays a central role by enabling learners to decode linguistic structures, process new language input, and build a distinct L2 mental lexicon alongside their L1 (Carroll, 1990; Paradis, 2004; Skehan, 1998). It is a fact that they are accompanied by affective factors in language learning, which hold key positions by maintaining motivation for L2 competency, boosting self-confidence to engage effectively in social interactions, and regulating and reducing negative emotions like foreign language anxiety (Dörnyei, 2005; Pishghadam, 2009; Thao et al., 2023).

Cognitive Side of Intelligence and Language Aptitude

Based on the aforementioned definitions, intelligence is mainly defined as the ability to learn, reason, solve problems, and adapt to new situations, encompassing cognitive abilities such as

memory, processing speed, and logical reasoning that enable effective functioning across contexts. Carroll (1962) highlights its educational relevance, noting that intelligence influences how efficiently students comprehend and respond to teachers' instructions. Research indicates that general intelligence strongly predicts academic performance, with notable correlations between intelligence scores and school grades (Roth et al., 2015). At the same time, education contributes to the development of cognitive abilities, revealing a reciprocal relationship: formal schooling enhances academic skills and yields measurable gains in general intelligence (Ritchie & Tucker-Drob, 2018).

Human cognitive functioning involves processes such as memory, attention, perception, spatial and processing speed, and executive functions like cognitive flexibility—collectively represented by CI, which measures an individual's capacity to reason, solve problems, and process information (Neisser et al., 1996; Schubert et al., 2024). Language, as a distinctive cognitive function, draws on these mechanisms. Core cognitive abilities—perception, attention, memory, processing speed, and executive control—are fundamental to L2 learning. Figure 1 illustrates cognitive processes contributing to L2 learning, which largely overlap with those active in L1 acquisition (Cook, 1985; 1997). While L1 acquisition is typically implicit, L2 learning often relies on explicit and conscious processing (Ellis, 1994; Krashen, 1982).

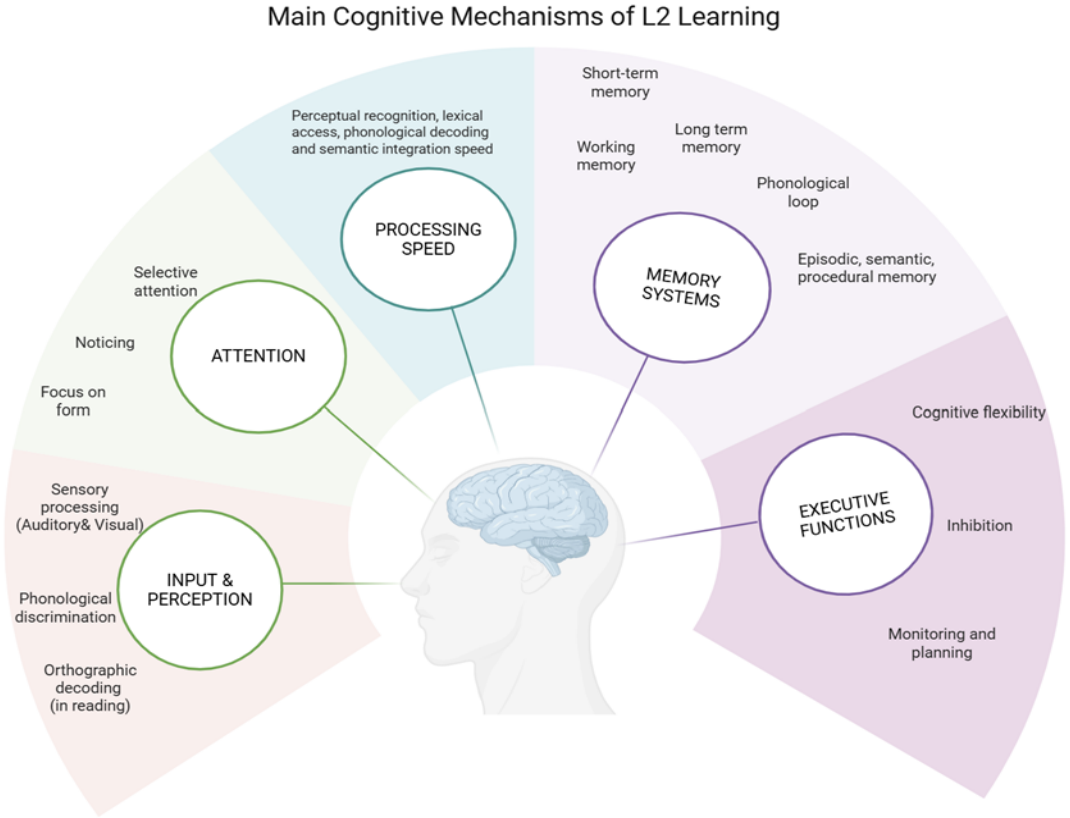


Figure 1. Main cognitive processes and functions of second/foreign language learning

Despite being not limited to the functions presented in Figure 1, the primary cognitive mechanisms involved in L2 learning are generally the ones that are already active in L1 acquisition (Cook, 1985; 1997). Input perception and processing engage auditory and visual systems, phonological discrimination, and orthographic decoding (Cook, 1977; Field, 2004). Learners must

accurately perceive and decode input before comprehension or storage. Attention enables selective focus on linguistic features and supports noticing—awareness of grammar or vocabulary (Schmidt, 1990). Instructional techniques such as focus on form help direct attention to language structures (Doughty & Varela, 1998). Memory systems are central: working memory supports temporary storage during sentence construction, while short- and long-term memory underlie vocabulary and rule learning (Baddeley, 2003). The phonological loop aids verbal rehearsal using procedural memory, while declarative memory supports semantic and episodic knowledge (Ullman, 2005). Executive functions—including cognitive flexibility, inhibition of L1 interference, and monitoring language output—govern higher-order control processes (Bialystok, 2009; Miyake et al., 2000). These rely on regions such as the anterior cingulate cortex and parietal lobes (Bialystok, 2009), facilitating language switching and suppression of interference essential for tasks like conversation and translation. Thus, the neural and cognitive architecture of the brain critically underpins L2 learning (Abutalebi & Green, 2016; Bialystok, 2009; Li et al., 2014).

Language aptitude, a cognitive trait related to L2 learning, refers to an individual's ability to acquire languages other than their mother tongue (Kristiansen, 1990). Gardner and MacIntyre (1992) conceptualized aptitude as a set of cognitive abilities facilitating L2 acquisition, distinct from CI, as it specifically relates to the efficiency of language learning rather than overall academic performance. They identified components such as phonemic coding ability, grammatical sensitivity, and memory capacity as crucial for internalizing linguistic material. Skehan (1998) highlighted aptitude's alignment with formal language instruction. Gardner and MacIntyre (1992) further situated aptitude within a socio-educational model, emphasizing its interaction with affective factors like motivation and anxiety. They concluded that successful L2 learning requires both cognitive skills and emotional readiness, framing aptitude as a dynamic, rather than fixed, trait. Although aptitude represents a cognitive advantage, effective learning still depends on motivation and regulation of negative emotions such as anxiety.

Affective Side of L2 Learning and Emotional Intelligence

L2 learning involves not only cognitive work but also emotional progress (Babanoğlu, 2025), and emotional factors significantly influence this process. Throughout the demanding and sometimes monotonous process of L2 learning, learners often experience a wide range of emotions and feelings (Oz et al., 2015). In order to learn effectively, they must feel confident, less anxious and capable of expressing themselves in the target language. In this regard, Arnold (2021) highlights the crucial role of emotional factors in learning and teaching, particularly in language education.

In L2 learning, affect is mostly associated with emotions, attitudes, beliefs, personality traits, anxiety, tolerance of ambiguity and motivation that learners experience during the process, all of which influence their behaviours and learning (Hurd, 2008). Krashen (1982) formalised this idea with his well-known "Affective Filter Hypothesis". Accordingly, when learners are motivated and self-confident with relatively low anxiety levels, the 'filter' is lowered, allowing more comprehensible input to reach the language acquisition device in mind. High anxiety, low self-esteem or low motivation, on the other hand, raise the filter, preventing input from being internalised. The following variables sit at the centre of Krashen's model:

- Motivation: the drive to learn,
- Self-confidence/self-esteem: the belief that one can succeed,

- Anxiety: the tension that can inhibit language performance.

Oxford (1990; 2011) extended this view to emphasize that learners can strategically regulate their affect. She highlights the significance of affective strategies such as lowering anxiety, encouraging oneself, and taking one's emotional temperature. While Krashen (Ibid) emphasizes how affect filters input, Oxford (Ibid) highlights how learners can act on their affect to keep the filter low. This adds an agentive and skill-building dimension: Affect is not just a state that "happens" to learners — it's something they can learn to handle.

The rising emphasis on emotions in language learning has led to a greater focus on the concept of EI, which was first introduced by Salovey and Mayer (1990) and popularised by Daniel Goleman (1995) in the discipline of psychology. According to Mayer et al. (2008), EI is “the ability to carry out accurate reasoning about emotions and the ability to use emotions and emotional knowledge to enhance thought” (p. 511). Accordingly, while affective factors explain how emotions influence language acquisition, EI provides a framework for managing and leveraging those emotions effectively. EI involves recognizing, understanding and regulating emotions to enhance learning and relationships. In language classrooms, those with higher EI can manage frustration, communication breakdowns and build positive relationships, reducing the affective filter and facilitating language improvement.

Despite controversies (e.g. overhyped claims, uncertain development and limited impact evidence) (Matthews et al., 2002), extensive research in the field reveals that EI is still an attractive research topic in education due to its relevance to student motivation, self-regulation and interpersonal competence. Zins et al. (2004), for instance, stated that EI not only supports broader personal development but also academic success. Similarly, Oz et al. (2015) noted that improving emotional intelligence can boost motivation and have a positive impact on attitudes towards learning foreign languages, leading to better language proficiency.

Synergy of Cognitive and Emotional Intelligence in L2 Learnin

Emotions and cognition are commonly believed to function in different areas. They were long viewed separate from each other due to methodological limitations in early research on brain and mind. However, research findings obtained in later periods highlight that understanding a complex behaviour requires recognition of the essential interaction between the two systems (Pessoa et al., 2006). Likewise, although emotions and cognition have been regarded as two distinct processes in L2 learning, and cognitive skills and language aptitude have been favoured as being considered more important form a traditional point of view, researchers now recognise that the two are dynamically linked and influence language performance and proficiency (MacIntyre, 2002; MacIntyre et al., 2024; Qin et al., 2022; Swain, 2013). Lett and O’Mara (1990) point out that although cognitive abilities consistently predict success in foreign language learning, non-cognitive features like motivation and attitude have significant potential in predicting and enhancing success in acquiring foreign language skills.

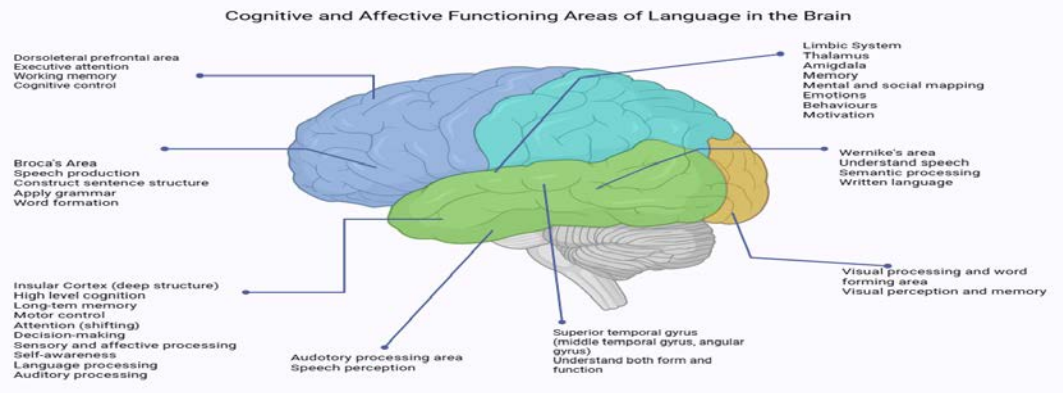


Figure 2. Cognitive and affective language areas in the brain

Figure 2 illustrates the distributed yet integrated nature of language processing in the human brain. It highlights the cognitive and affective areas that are essential for language learning. Core cognitive regions such as Broca's area (involved in syntactic construction and speech production) and Wernicke's area (involved in semantic comprehension) are supported by broader networks, including the dorsolateral prefrontal cortex (involved in executive attention and working memory (Miller & Cohen, 2001), and the posterior superior temporal gyrus (involved in both semantic memory and form-function mapping (Binder & Desai, 2011). Figure 2 also emphasises the affective components of language acquisition. The limbic system contributes to emotional salience and motivation (Pessoa, 2008), and the insular cortex, a hub in the salience network, facilitates attentional switching, interoceptive awareness and affective-social processing, which are critical for communicative competence (Uddin, 2015). This functional integration is consistent with contemporary models of L2 learning, which view it not merely as symbolic or grammatical decoding, but as a cognitive process that is socially and emotionally grounded and which recruits memory, motivation, attention and emotional evaluation across a wide network of cortical and subcortical regions.

Figure 2 subsequently shows that the way in which cognitive and affective brain systems work together in language processing and acquisition, particularly in the process of L2 learning. Traditional language centres, such as Broca's and Wernicke's areas, coordinate grammatical structure and semantic comprehension (Friederici, 2011), while broader cognitive control regions such as the dorsolateral prefrontal cortex and posterior parietal cortex support memory systems, attentional focus and strategic learning (Miyake & Friedman, 2012). Importantly, emotional and social brain systems, including the limbic system and insula, become active when language learning involves meaningful, emotionally resonant or socially contextualised input (Immordino-Yang & Damasio, 2007). Together, these systems form the neurological basis for EI, the capacity to perceive, regulate and respond to emotional cues. EI complements CI by enhancing motivation, empathy and resilience during language learning (MacIntyre et al., 2002). These networks reflect an embodied, socially grounded model of L2 learning, in which cognitive effort and emotional engagement are deeply interconnected processes. Oz et al. (2015) point out that “integration of human’s reasoning of emotions and cognition, i.e. emotional intelligence and cognitive intelligence, might provide positive outcomes in intellectual and emotional development of language learners.” (p. 417). The following section provides a closure on the relevant literature reviewed so far with a focus on the integration of CI and EI components in L2 learning.

Conclusion

Though traditionally viewed as separate domains, cognitive and emotional intelligences are deeply interconnected in human brain. While cognitive processes such as memory, attention and reasoning are primarily associated with the prefrontal cortex, emotional regulation and motivation –key components of EI –are integrated through the limbic system and constantly interact with executive control functions of CI. This dynamic coordination enables individuals to process language effectively, stay motivated, manage anxiety and engage meaningfully in communication. Therefore, understanding how these systems operate in tandem is crucial for gaining insight into the mental architecture that supports L2 learning.

Since CI has long been regarded as the dominant factor in learning success, Arnold (2011) argues that, considering current understanding of education, it remains difficult to effectively address emotional aspects alongside cognitive ones. Nonetheless, there is a growing body of research with the aim of enhancing educational programs by incorporating elements such as character development, thinking abilities, learning preferences and EI. Research continues to support the significance of both cognitive and non-cognitive traits in successful L2 learning. As emphasized in Lett and O'Mara (1990), while cognitive abilities reliably predict foreign language achievement, non-cognitive features like motivation and attitude towards foreign language learning also play a powerful role in shaping fruitful learning outcomes. These emotional traits influence how learners approach challenges, persist through difficulties, and engage with and process L2 linguistic input effectively. Together, CI and EI form a complementary and mutually reinforcing foundation for L2 learning.

As for the pedagogical implications generated in light of these insights, Lustiyantie et al. (2025) concluded that a holistic approach blending cognitive strategy development with emotional support is essential in language education. Teachers should design learning environments that strengthen cognitive functions, such as memory training and cognitive flexibility, and that foster emotional well-being through encouragement, empathy and stress reduction strategies. Particularly, EI can be cultivated through activities like motivation tasks, role plays, debates, writing tasks and group work (Tevdovska, 2017) in language classrooms that not only encourage learners to engage with others and reflect on emotions but also stimulate cognitive processes like critical thinking, problem-solving, and perspective-taking –essential for deeper language comprehension and use. Integration of CI and EI into classroom practices has the potential to enhance learner engagement and motivation, ultimately leading to more effective and enjoyable language learning experiences.

Funding: No funding was received to carry out the study
Acknowledgments: The authors declare that this work is entirely our own, completed without external assistance, and requires no acknowledgements.
Conflicts of Interest: The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.
Authenticity: Authors declare that the work submitted is our own and was completed collaboratively. All sources, references, and assistance used have been properly acknowledged. We confirm that no part of this work has been plagiarized or produced through unauthorized means, and each author accepts responsibility for the integrity and authenticity of the content.
Artificial Intelligence Statement: The authors, confirm that artificial intelligence tools (Grammarly, DeepL, and ChatGPT) were used only for proofreading and language refinement.

About the Author:
M. Pınar Babanoğlu currently works at Dept. of Foreign Languages, Mersin University. Her research interests are English as a Second/Foreing Language Learning and Teaching, Second Language Acquisition, Learner Corpus Research.ORCID ID: https://orcid.org/0000-0001-8166-974X

Reyhan Ağçam works at Dept. of Foreign Languages, Kahramanmaraş Sütçü imam University. Her research interests are English as a Second/Foreign Language Learning and Teaching, Second Language Acquisition, Learner Corpus Research. ORCID ID: https://orcid.org/0000-0002-5445-9031

Muhammet Fatih Adıgüzel currently works at Dept. of Foreign Languages, Mersin University. His research interests include Linguistics, Second Language Acquisition, Corpus Research. ORCID ID: https://orcid.org/0000-0002-6962-0078

Tuçe Öztürk Karataş works at Dept. of Foreign Languages, Mersin University. Her research interests are English as a Second/Foreign Language Learning and Teaching and language testing. ORCID ID: https://orcid.org/0000-0002-3844-3574

Esin Dündar works at Dept. of Foreign Languages, Mersin University. Her research interests are English as a Second/Foreign Language Learning and Teaching, Curriculum Studies, Sustainability in ELT and EFL coursebook Evaluation.
ORCID ID: https://orcid.org/0000-0002-8160-0176

References

Abutalebi, J., & Green, D. W. (2016). Neuroimaging of language control in bilinguals: Neural adaptation and reserve. *Bilingualism: Language and Cognition, 19*(4), 689–698. https://doi.org/10.1017/S1366728916000225

Arnold, J. (2011). Attention to affect in language learning. *Anglistik: International Journal of English Studies, 22*, 11–22. https://files.eric.ed.gov/fulltext/ED532410.pdf

Arnold, J. (2021). Affective factors in language learning. In M. Simons & T. Smits (Eds.), *Language Education and Emotions: Research into Emotions and Language Learners, Language Teachers and Educational Processes*, (pp.1-15). London: Routledge.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*(3), 189–208. https://doi.org/10.1016/S0021-9924(03)00019-4

Babanoğlu, M. P. (2025). Decoding the relationship between emotional intelligence and foreign language anxiety: A systematic review and meta-analysis. *Language Teaching Research, 0*(0), 1-31. https://doi.org/10.1177/13621688251345726

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*(1), 3–11. https://doi.org/10.1017/S1366728908003477

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences, 15*(11), 527–536. https://doi.org/10.1016/j.tics.2011.10.001

Bru-Luna L., Martí-Vilar M., Merino-Soto C., Cervera-Santiago J. (2021). Emotional intelligence measures: A systematic review. *Healthcare, 9*(1696). https://doi.org/10.3390/healthcare9121696

Carroll, J. B. (1962). The prediction of success in intensive foreign language training. Reprint from *Training Research and Education*, pp. 87-136. University of Pittsburgh Press. https://files.eric.ed.gov/fulltext/ED038051.pdf

Carroll, J. B. 1990. Cognitive abilities in foreign language aptitude: Then and now. In T. S. Parry and C. W. Stansfield (Eds.), *Language Aptitude Reconsidered*, (pp. 11–29). Englewood Cliffs, NJ: Prentice Hall.

Cook, V. (1985). Chomsky's Universal Grammar and second language learning. *Applied Linguistics*, *6*(1), 2–18. https://doi.org/10.1093/applin/6.1.2

Cook, V. J. (1977). Cognitive processes in second language learning. *IRAL*, *15*, 1, 1–20.

Cook, V. (1997). The consequences of bilingualism for cognitive processing. In A. de Groot & J. Kroll (Eds.), *Tutorials in Bilingualism*, (pp. 279-300). NY: Psychology Press.

Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty, & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114-138). Cambridge: Cambridge University Press.

Dronkers, N. F., Ivanova, M. V., & Baldo, J. V. (2017). What do language disorders reveal about brain–language relationships? From classic models to network approaches. *Journal of the International Neuropsychological Society*, *23*(9–10), 741–754. https://doi.org/10.1017/S1355617717001121

Dörnyei, Z. (2005). Motivation in second-language learning: The L2 motivational self system, In Z. Dörnyei (Ed.), *Motivation and second language acquisition* (pp. 9–42). Honolulu, HI: University of Hawaii Press).

Dewaele, J.-M., Petrides, K. V., & Furnham, A. (2008). Effects of trait emotional intelligence and sociobiographical variables on communicative anxiety and foreign language anxiety among adult multilinguals: A review and empirical investigation. *Language Learning*, *58*(4), 911–960. https://doi.org/10.1111/j.1467-9922.2008.00482.x

Ellis, R. (1994). *The study of second language acquisition.* Oxford: Oxford University Press.

Field, J. (2004). *Psycholinguistics: The key concepts*. Routledge.

Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews*, *91*(4), 1357–1392. https://doi.org/10.1152/physrev.00006.2011

Fromkin, V., Rodman, R. & Hyams, N. (2017). *An Introduction to Language*, 11th ed. Boston, MA: Cengage Learning.

Gee, J. P. (2014). *An introduction to discourse analysis: Theory and Method*, 4th ed. London and New York: Routledge.

Gardner, R. C. (1985). *Social psychology and second language learning: the role of attitudes and motivation*. London: Edward Arnold.

Gardner, R. C., & MacIntyre, P. D. (1992). A student's contributions to second language learning. Part 1: Cognitive variables. *Language Teaching*, *22*, 211-220. http://dx.doi.org/10.1017/S026144480000700X

Goleman, D. (1995). *Emotional intelligence: Why it can matter more than IQ*. Bantam Books.

Hurd, S. (2008). Affect and strategy use in independent language learning. In S. Hurd & T. Lewis (Eds.), *Language Learning Strategies in Independent Settings*, pp. 218-236. Multilingual Matters.

Immordino-Yang, M. H., & Damasio, A. (2007). We feel, therefore we learn: The relevance of affective and social neuroscience to education. *Mind, Brain, and Education*, *1*(1), 3–10. https://doi.org/10.1111/j.1751-228X.2007.00004.x

Jacobs, N., Van Os, J., Derom, C., & Thiery, E. (2007). Heritability of intelligence. *Twin Research and Human Genetics*, *10*(S1), 11-14. https://doi.org/10.1375/twin.10.supp.11

Kramsch, C. (1998). *Language and Culture*. Oxford: Oxford University Press.

Krashen, S. D. (1982). *Principles and practice in second language acquisition.* Pergamon.

Kristiansen, I. (1990). Nonverbal intelligence and foreign language learning. *Research Bulletin*, 73. Helsinki: Department of Education, University of Helsinki.

Lenneberg, E.H. (1967). Biological foundations of language. New York: John Wiley and Sons. https://doi.org/10.1080/21548331.1967.11707799

Lett, J. & O'Mara, F. E. (1990). Predictors of success in an intensive language learning context: correlates of language learning at the Defense Language Institute Foreign Language Center. In Parry, T. S. & Stansfield, C. W. (Eds.), *Language Aptitude Reconsidered*, (pp. 222-260). Englewood Cliffs, N. J.: Prentice Hall Regents.

Li, P., Legault, J., & Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: Anatomical changes in the brain. *NeuroReport*, *25*(13), 135–141. https://doi.org/10.1097/WNR.0000000000000277

Lightbown, P. M., & Spada, N. (2006). *How languages are learned* (3rd ed.). Oxford University Press.

Lustiyantie, N., Chaeruman, U. A., & Rajaguk-guk, S. B. (2025). Cognitive and affective roles in language and literature learning at the secondary school level (a literature study). *Eduvest-Journal of Universal Studies*, *5*(5), 5121-5131. http://dx.doi.org/10.59188/eduvest.v5i5.49940

MacIntyre, P. D. (2002). Motivation, anxiety and emotion in second language acquisition. In P. Robinson (Ed.), *Individual differences and instructed language learning*, (pp. 45-68). Amsterdam: John Benjamins. https://doi.org/10.1075/lllt.2.05mac

MacIntyre, P. D., Dörnyei, Z., Clément, R., & Noels, K. A. (2002). Motivation, emotion, and language learning. *Studies in Second Language Acquisition*, *24*(1), 45–68. https://doi.org/10.1017/S027226310200102X

MacIntyre, P. D., Gregersen, T., & Mercer, S. (2024). *Positive psychology in SLA: Emotions, motivation, and wellbeing*. Bristol, UK: Multilingual Matters.

Matthews, G., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., ... & Warm, J. S. (2002). Fundamental dimensions of subjective state in performance settings: task engagement, distress, and worry. *Emotion*, *2*(4), 315-340. https://doi.org/10.1037/1528-3542.2.4.315

Mayer, J. D., Salovey, P., & Caruso, D. R. (2004). TARGET ARTICLES: Emotional intelligence: Theory, findings, and implications. *Psychological Inquiry*, *15*(3), 197-215. https://doi.org/10.1207/s15327965pli1503_02

Mayer, J. D., Roberts, R. D., & Barsade, S. G. (2008). Human abilities: Emotional intelligence. *Annu. Rev. Psychol.*, *59*(1), 507-536. https://doi.org/110.1146/annurev.psych.59.103006.093646

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167

Miyake, A., & Friedman, N. P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. Miyake & N. P. Friedman (Eds.), *Foreign Language Learning Psycholinguistic Studies on Training and Retention*. Psychology Press, pp. 339-364.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*(2), 77–101. https://doi.org/10.1037/0003-066X.51.2.77

Oxford, R. L. (2011). Strategies for learning a second or foreign language. *Language Teaching, 44*(2), 167-180. https://doi.org/10.1017/S0261444810000492

Oxford, R. L. (1990). Styles, strategies, and aptitude: Connections for language learning. *Document Resume*, 73-125.

Oz, H., Demirezen, M., & Pourfeiz, J. (2015). Emotional intelligence and attitudes towards foreign language learning: Pursuit of relevance and implications. *Procedia – Social and Behavioral Sciences, 186*, 416–423. https://doi.org/10.1016/j.sbspro.2015.04.118

Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nat Rev Neurosci, 9*, 148–158. https://doi.org/10.1038/nrn2317

Pessoa, L., Japee, S., Sturman, D., & Ungerleider, L. G. (2006). Target visibility and visual awareness modulate amygdala responses to fearful faces. *Cerebral Cortex, 16*(3), 366-375. https://doi.org/10.1093/cercor/bhi115

Pishghadam, R. (2009). A quantitative analysis of the relationship between emotional intelligence and foreign language learning. *Electronic Journal of Foreign Language Teaching*, *6*(1), 31-41.

Plomin, R., & Spinath, F. M. (2004). Intelligence: genetics, genes, and genomics. *Journal of Personality and Social Psychology, 86*(1), 112. https://doi.org/10.1037/0022-3514.86.1.112

Qin, T., Zhang, L. J., & Wang, W. (2022). Understanding learners' emotion and cognition in a flipped English classroom: A person-centered approach. *Frontiers in Psychology, 13*, 1005237. https://doi.org/10.3389/fpsyg.2022.1005237

Ritchie S. J., & Tucker-Drob, E. M. (2018). How much does education improve intelligence? A Meta-analysis. *Psychol Sci*., *29*(8), 1358-1369. https://doi.org/10.1177/0956797618774253

Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence, 53*, 118–137. https://doi.org/10.1016/j.intell.2015.09.002

Robinson, P. (2002). *Individual differences and instructed language learning.* John Benjamins.

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality*, *9*(3), 185-211. https://doi.org/10.2190/DUGG-P24E-52WK-6CDG

Schmidt, R. W. (1990). The role of consciousness in second language learning1. *Applied linguistics, 11*(2), 129-158. https://doi.org/10.1093/applin/11.2.129

Schubert, A.-L., Löffler, C., Wiebel, C., Kaulhausen, F., & Baudson, T. B. (2024). Don't waste your time measuring intelligence: Further evidence for the validity of a three-minute speeded reasoning test. *Intelligence, 102*, 101804. https://doi.org/10.1016/j.intell.2023.101804

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Spearman, C. E. (1927). *The Abilities of Man: Their Nature and Measurement*. London: Macmillan & Co.

Sternberg, R. J. (2012). Ethical drift. *Liberal Education, 98*(3), 60.

Swain, M. (2013). The inseparability of cognition and emotion in second language learning. *Language Teaching, 46*(2), 195–207. https://doi.org/10.1017/S0261444811000486

Taheri, F., Geshani, A., Fatahi, J., Jalaie, S., & Tavakoli, M. (2019). Acceptable noise levels in Arabic-Persian bilinguals. *Auditory and Vestibular Research, 28*(1), 14-21. https://doi.org/10.18502/avr.v28i1.411

Tevdovska, E. S. (2017). The impact of emotional intelligence in the context of language learning and teaching. *SEEU Review, 12*(1), 125-134. https://doi.org/10.1515/seeur-2017-0009

Thao, L. T., Thuy, P. T., Thi, N. A., Yen, P. H., Thu, H. T. A., & Tra, N. H. (2023). Impacts of emotional intelligence on second language acquisition: English-major students' perspectives. *SAGE Open*, *13*(4). https://doi.org/10.1177/21582440231212065

The Editors of Encyclopaedia Britannica (2010, November 9). *Intelligence summary*. *Encyclopedia Britannica*. https://www.britannica.com/summary/human-intelligence-psychology

Uddin, L. Q. (2015). Salience processing and insular cortical function and dysfunction. *Nature Reviews Neuroscience*, *16*(1), 55–61. https://doi.org/10.1038/nrn3857

Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and Context in Adult Second Language Acquisition* (pp. 141–178). Georgetown University Press.

Yurdakul, N., Ker-Dinçer, M., Vural, Z. B. A., & Beril, Z. (2008). Searching for excellence in educational communication: The role of IQ, EQ and SQ. *Ahmet Yesevi University Board of Trustees, 45*, 147-164.

Zins, J. E., Weissberg, R. P., Wang, M. C., & Walberg, H. J. (Eds.). (2004). *Building academic success on social and emotional learning: What does the research say?* New York: Teachers College Press.

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp.330-342
DOI: https://dx.doi.org/10.24093/awej/vol16no4.19

Cross-cultural Pragmatics and Code-switching in Multilingual EFL Classrooms

Valeriia Smaglii
Odesa National Maritime University, Odesa, Ukraine
Corresponding Author: valeriya.smagliy@gmail.com

Svitlana Yukhymets
Odesa National Maritime University, Odesa, Ukraine
E-mail: yukhymets.svetlana@gmail.com

Yevheniia Kornielaieva
Odesa National Maritime University, Odesa, Ukraine
E-mail: kornelaeva2017@gmail.com

Inna Kivenko
Odesa National Maritime University, Odesa, Ukraine
E-mail: innakivenko@gmail.com

Received: 08/14/2025 Accepted:10/29/2025 Published:12/10/2025

This study explores how cross-cultural pragmatic choices and code-switching jointly shape social action and interactional competence in multilingual classrooms where English is a foreign language. It addresses a key gap in applied linguistics: although code-switching has been widely studied, little is known about how alternation operates within conversational structure to support learning and participation. Understanding these mechanisms is pedagogically important for using bilingual resources more effectively in English-language education. The research bases on a mixed-method corpus of twenty-six hours of naturally occurring classroom interaction, including teacher-led lessons, pair and group work, and assessment talk. Conversation analysis and quantitative observation identify how different types of code-switching – between and within sentences and in short tags – serve pragmatic and organizational functions such as managing turn entry, maintaining precision, and facilitating closure. Findings show that English predominates in main communicative acts. At the same time, learners' first languages appear at sequence boundaries to clarify meaning, express politeness, and reinforce inclusion. Alternation patterns typically follow a brief first-language introduction, an English head action, and a short return to the first language, enhancing comprehension and task completion without reducing authenticity. The study redefines code-switching as a strategic pedagogical resource that promotes understanding, collaboration, and pragmatic development.

Keywords: Bilingual education, code-switching, conversation analysis, interactional competence, multilingual classroom, pragmatics

How to cite: Smaglii, V., Yukhymets, S., Kornielaieva, Y., &Kivenko, I. (2025). Cross-cultural Pragmatics and Code-switching in Multilingual EFL Classrooms. *Arab World English Journal, 16*(4): 330-342. https://dx.doi.org/10.24093/awej/vol16no4.19

Introduction

Multilingual English-as-a-foreign-language classrooms are dynamic contact zones where distinct politeness systems, participation frameworks, and turn-taking conventions converge. Communicative success in such environments depends on how learners transfer and interpret illocutionary force and facework across languages (the pragmalinguistic dimension) and how these align with the locally recognized rights and obligations of interaction (the sociopragmatic dimension). Within this ecology, code-switching is not a symptom of linguistic deficiency but a strategic interactional resource and indexical signal. Speakers alternate languages to execute conversational sequence and participation, to express stance, to signal phase transitions in activity, and to access lexical or grammatical resources that may be unavailable or socially marked in the target language.

This study integrates insights from cross-cultural pragmatics and conversation analysis to explain how bilingual resources shape classroom interaction. Cross-cultural pragmatics provides the framework for understanding the variation in politeness, indirectness, and mitigation strategies across cultural contexts, while conversation analysis supplies the methodological precision to trace these pragmatic choices within the turn-by-turn unfolding of classroom talk. Together, they allow for a fine-grained mapping of how language alternation operates simultaneously as a practical strategy and as a structural resource for organizing social action.

Following the principles of interactional sociolinguistics and CA, the study treats code-switching as a contextualization device closely tied to the architecture of conversation: pre-, base-, and post-expansions, preference organization (agreement or disagreement formats), and repair sequences (self- and other-initiated or completed). Structurally, the analysis differentiates between inter- and intra-sentential switching and tag-switching, as well as between insertion, alternation, and congruent lexicalization, distinguishing the roles of the matrix and embedded languages. These patterns systematically align with pragmatic functions such as request mitigation through first-language particles, disagreement softening through stance markers, and precision moves in technical or institutional discourse.

Despite extensive pedagogical debate – ranging from “target-language-only” policies to translanguaging approaches – there remains a lack of empirically grounded, micro-analytic accounts of how code-switching shapes the pragmalinguistic packaging and sociopragmatic alignment of classroom interaction. Previous studies (Alzamil, 2019; Criado & González-Romero, 2023) have primarily quantified first-language use or reported learners’ and teachers’ attitudes, without examining where in the turn or sequence switches occur, what conversational actions they accomplish (for example, upgrading, downgrading, or rekeying), and how they affect participation and uptake.

The present study fills this gap by connecting micro-level linguistic alternation with macro-level interactional competence. Using conversation analysis as its methodological foundation, it systematically links pragmatic functions to sequential position, turn construction, and repair organization. The aim is to demonstrate how cross-cultural pragmatic alternatives and code-switching jointly accomplish social action and build interactional proficiency in multilingual English classrooms. Specifically, the study seeks to: (1) describe the distribution and typology of switching across activity types (teacher-fronted talk, group work, and assessment discourse) and relate switch points to floor management; (2) analyze how code-switching shapes the pragmalinguistic realization of speech acts (requests, refusals, advice, assessments, disagreement) and their sociopragmatic fit; (3) explore how alternation aligns with sequence expansion,

preference formats, and repair trajectories; (4) model stance work and face mitigation achieved through alternation; and (5) validate correlations between code-switching patterns and indicators of learning opportunity, such as comprehension checks, uptake, and task completion, alongside learners' perceptions of authenticity and cognitive effort.

Literature review

Code-switching in English-as-a-foreign-language classrooms has been repeatedly demonstrated to be a structured, purposeful practice rather than a random fallback strategy. Across a range of instructional contexts, switching functions as a pedagogical and interactional tool for achieving clarity, managing participation, and aligning with local discourse norms. Studies in Turkish secondary settings show that both teachers and students employ code-switching to manage tasks, clarify meaning, and adapt participation frameworks to the local classroom ecology, with switches systematically embedded in the unfolding sequence of talk (Ataş & Sağın-Şimşek, 2021). Teachers' perceptions strongly shape this practice: when classrooms are conceptualized as bilingual spaces, code-switching is reframed as context-sensitive scaffolding rather than linguistic deficiency, though tolerance levels vary between shared-L1 and multilingual groups (Gallagher, 2020). Interview- and questionnaire-based research supports this view, identifying pedagogical, managerial, and socio-emotional functions for switching while warning against mechanical overuse (Hazaymeh, 2022; Temesgen & Hailu, 2022). Recent observational work extends these findings by showing how code-switching indexes participation, uptake, and alignment in real-time task interaction (Amiri, 2024). This body of research redefines code-switching as an interactionally structured mechanism that mediates access to content, procedure, and affiliation within classroom discourse, foregrounding its pragmatic and social organization.

A second strand of research reconceptualizes this practice through the lens of translanguaging pedagogy, where multilingual resources are mobilized as deliberate tools for meaning-making and identity construction. Systematic reviews suggest that pedagogical translanguaging can enhance language outcomes when embedded in task design and assessment, though empirical evidence remains fragmented and often short-term (Huang & Chalmers, 2023). Micro-analytic studies in English-medium instruction further demonstrate how teachers dynamically "rekey" student responses into academic registers through multimodal and multilingual cues, thereby socializing disciplinary discourse in real time (Tai, 2024). Intervention research documents how translanguaging training reshapes teachers' professional identities and classroom discourse, fostering greater openness to linguistic hybridity (Yasar Yüzlü & Dikilitaş, 2022; Yüzlü, 2022). Taken together, these studies shift attention from the frequency of first-language use toward the interactional and identity-building processes through which multilingual repertoires support stance-taking, footing, and register development.

A third body of research situates code-switching and translanguaging within interlanguage and cross-cultural pragmatics, emphasizing their role in developing pragmatic competence. Empirical studies show that multilingual learners negotiate politeness, indirectness, and mitigation strategies through dynamic alternation between languages. For instance, Chinese English speakers coordinate refusal strategies through layered mitigation to balance facework and preference organization (Fang, 2024). Instructional studies demonstrate that targeted pragmatic interventions improve comprehension of indirect speech acts (Nguyen & Pham, 2022), while production-oriented tasks reveal that politeness and contextual appropriateness weigh more heavily on learner performance than syntactic complexity alone (Kim & Oh, 2025). Both online and classroom-based

programs confirm measurable gains in pragmatic accuracy, with explicit teaching showing a slight advantage (Rafiq & Yavuz, 2024), and dynamic-assessment models highlighting the role of mediated feedback in fostering pragmatic awareness (Rezai et al., 2023). A recent meta-synthesis of EFL and ESL contexts reveals that pragmatic competence is consistently lower in EFL settings due to limited interactional authenticity. Yet, code-switching and translanguaging practices can partially bridge this gap by extending participation and contextual grounding (Wang et al., 2024). Together, these findings position multilingual alternation not merely as a surface phenomenon but as a key mediator of pragmatic development, linking linguistic form to interpersonal appropriateness and communicative success.

Despite these converging insights, essential limitations remain. Most studies rely on self-report data or global performance ratings rather than fine-grained, interaction-sensitive measures. Few systematically connect switch loci, such as pre-expansions, repair sequences, or stance marking, to observable indicators of pragmatic comprehension or speech act realization. Longitudinal research is scarce, and operationalizations of translanguaging remain inconsistent across contexts. This article addresses these gaps by integrating cross-cultural pragmatics and code-switching within a conversation-analytic framework, linking sequential placement of switches, discourse markers, and repair organization to measures of interactional competence and pragmalinguistic achievement.

Methodology

A mixed-methods design combines conversation analysis and interactional sociolinguistics with light quantitative modeling. Naturally occurring interaction is recorded in multilingual Odesa Maritime National University EFL classes across teacher-fronted instruction, pair/group work, and assessment talk. The sequence is the basic analytic unit; turns are segmented into turn-constructional units and transition relevance places. The choice of Odesa Maritime National University is grounded in its linguistic and institutional diversity: it enrolls students from Ukraine, Turkey, India, and several African and Asian countries, creating a naturally multilingual environment where English functions as both a target language and a lingua franca for professional training. While the findings emerge from a specific institutional context, the analytical focus on interactional mechanisms supports broader transferability to other multilingual EFL and English-medium-instruction environments where comparable participation structures and pedagogical aims exist.

Participants

Data come from intact classes in a multilingual EFL program (5 instructors; 100 students across diverse L1S). Participation required enrollment and consent; learner proficiency is documented via CEFR-aligned records for descriptive comparisons.

Material

Across one term, 20 lessons (80 minutes) yield about 26 hours of audio/video. Episodes rich in negotiation and evaluation are sampled for close analysis. Recordings use ceiling microphones and a fixed camera; audio is 44.1 kHz/16-bit. Transcription follows CA conventions in ELAN/EXMARaLDA with aligned tiers for speaker, language, code-switch type, action label, stance, repair, and discourse markers. Ten to fifteen percent of transcripts are second-coded for quality.

Analysis proceeds in three layers: (1) CA explication of how alternation configures action and sequence, illustrated with extracts; (2) corpus-style summaries of distributions and co-occurrences; (3) mixed-effects models linking alternation/pragmatic variables to uptake and appropriateness (random intercepts for class and speaker; fixed effects for activity type, role, proficiency band), plus lag-sequential checks of whether switches increase the probability of successful uptake in the next turn(s).

Ethics and data management

Institutional approval and informed consent have been obtained. Transcripts are anonymized; media are stored securely. De-identified transcripts, the codebook, and analysis scripts will be shared in an open repository; raw media remain access-controlled.

Results

Across the multilingual English-as-a-foreign-language corpus, code-switching displays systematic distribution rather than randomness, differing markedly by activity type and participation structure. Two parameters guide this analysis: locus in sequence and turn design (pre-expansions, within-turn insertions, post-expansions, and transition relevance places) and structural-functional typologies (inter- vs. intra-sentential switching, tag-switching, insertion, alternation, and congruent lexicalization, with matrix versus embedded language roles).

In teacher-led instruction, code-switching is predominantly intersentential and occurs at structural boundaries such as topic openings, task briefings, or lesson recapitulations. Teachers frequently alternate matrix languages: the first language frames management talk or metapragmatic explanation, while English predominates in disciplinary content. Brief tag-switches and stance markers (e.g., evaluative particles, vocatives) typically follow assessment turns or serve to rekey interactional alignment. Insertion functions to clarify procedures and highlight key terms, while alternation signals shifts in footing—from explanation to directive mode. These patterned switches maintain teacher control, ensure clarity, and facilitate collective uptake with minimal cognitive load.

During peer collaboration, intrasentential switching and alternation are most typical. English usually serves as the matrix language for task-related discourse, while brief L1 insertions assist in lexical retrieval, stance marking, or problem-solving. Conversely, when learners plan or negotiate off-record, the matrix reverts to the L1 with English embeddings for task-specific terminology. Switches frequently align with transition relevance places, enabling learners to hold or yield the floor, self-select, or seek confirmation. Insertion repairs lexical gaps and saves processing time, alternation reallocates task roles, and tag-switching reinforces solidarity and agreement. Overall, these patterns create a more symmetrical participation structure: weaker speakers rely on L1 scaffolding to sustain interaction without abandoning the English task frame.

In evaluative and testing sequences, English remains the matrix language to preserve institutional formality, while short L1 insertions soften judgments or clarify feedback. Tag-switches occur in post-expansions as stance tokens (“добре, well done”), whereas alternations appear at pre-expansions during reformulation of grading criteria or feedback frames. Code-switching here performs a face-management function, balancing authority with affiliation and ensuring that evaluation is received as constructive rather than punitive.

Synthesizing across genres, intersentential switching dominates boundary work and institutional control; it supports precision and local repair; tag-switching serves affective and

relational purposes across all contexts. Insertion facilitates accuracy and efficiency; alternation marks rekeying of participation roles; congruent lexicalization occurs mainly in high-proficiency dyads with shared bilingual grammar, promoting rapid co-construction. Matrix–embedded distribution shows English as the matrix for on-record institutional talk, with the L1 governing metacommunicative and managerial frames. Sequentially, switches cluster at pre-expansions to initiate participation, within-turn positions to sustain fluency, and post-expansions to close sequences or secure uptake.

In sum, code-switching functions as a contextualization device that manages floor rights, stance, and participation symmetry. Its patterned use across instructional, collaborative, and evaluative activity types demonstrates that alternation is not a random fallback but a methodical interactional resource for balancing clarity, engagement, and institutional order in multilingual classrooms.

A quantitative overview of the 26-hour multilingual classroom corpus shows that intersentential switching makes up about **40%** of all instances, mainly in teacher-led talk for boundary marking and management. Intrasentential switching accounts for roughly **35%**, dominating peer interaction where learners negotiate meaning or repair. Tag-switching constitutes about **15%**, functioning as stance or alignment markers across contexts. Insertions and alternations together form around **10%**, used for lexical precision, mitigation, and task role adjustment. English serves as the matrix language in about **70%** of task-related sequences, while the first language frames approximately **30%** of metacommunicative segments. Overall, these proportions confirm that code-switching follows patterned, context-dependent principles rather than random alternation.

Students overwhelmingly use English modal and conventional-indirect forms (*would/could you…?, would it be possible to…?, do you mind if…?)* with preparators and grounders *(I was wondering…, maybe…, because…*). Downtoners (*a bit, just*), stance-evidentials (*I think, it seems*), and tag questions *(…, right?*) reduce imposition even further. Vocatives and address terms (*Sir/Miss/guys/team*) set up participation footing, plural vocatives indexing inclusive alignment in peer work.

With interlocutors, students favor high-indirectness English head acts but frequently supply L1 grounders or post-requests for uptake (e.g., L1 reason giver or time limit), reserving the on-record move in English and shifting facework to L1. In peer work, English remains the Matrix for task talk, but brief L1 insertions offer precision (terms, times) or solidarity particles that legitimate self-selection at TRPs. In teacher-fronted requests for management, teachers alternate to L1 intersententially at pre-expansions for the sake of gaining silence/attention, and revert to English for the directive proper. Refusals have a typical dispreference structure: delay (micropauses), hesitations(*uh/um*), stance softeners (*maybe, I'm not sure*), and accounts/alternative offers (*I can't today, but tomorrow*…). Formulae such as *I'm afraid… and I don't think I can*… convey negative preference while sustaining rapport.

To others, students tend to code-switch to carry out accounts and apology formulae (L1 apology tokens/sorry with culture-specific particles). However, the head refusal is still in English (*I can't because*…). Among friends, head acts of refusal might be in English or L1, depending on face-threat: English for recordable task consequences; L1 for off-record stance (teasing, softening). In assessment speech, teachers employ English summaries (you met the requirement on X) but insert L1 padding before a negative outcome (short praise or empathy particle) to reduce face-threat before the English-grade statement. Admonitions are stated in the form of deontic modals

(you should/need to/have to), mitigated imperatives (*let's + bare verb; maybe try + -ing*), and suggestive frames (*how about…, it might be better to*…).

Evidentials (*by…, by the rubric*…) underpin epistemic warrant; term address is consonant with roles (buddy/guys vs first names). In peer planning, English Matrix head acts (you have to know the prices first) are mixed with L1 insertions for procedural explanation or justification (L1 because-clause), especially at pre- or post-expansions. Teachers tend to open with an L1 stance marker to downshift authority, followed by the actionable advice in English to leave it institutionally "on record." Positive assessment clusters around English evaluatives (*good/clear/interesting*), intensifiers (*really/significantly*), and upgrade formulae (much better than last time). Negative or mixed assessment use hedged formats (*it's kind of unclear…, maybe consider*…) combined with repairables (you could rephrase X as Y).

Teachers employ English as a Matrix for outcomes and criteria; brief L1 tag-switches provide affective stance (praise, encouragement) or "cushion" before a downgrade. English for evaluative features (coherence, grammar), L1 for affiliative alignment (compliments, humor), especially in post-expansions to complete the sequence without threatening solidarity. Disagreement is packaged through concessive prefaces (*I see, but*…), turn-initial discourse markers (*well, actually*), stance downtoners (*maybe, I'm not sure*), and evidence displays (*because…, according to*…). Vocatives and address terms help localize the face-threat (*look, team.*…). Preference organization shows delay and accounts before the counter-claim. Students in peer work will initiate the counter-claim in English but use L1 stance particles or laughter tokens within-TCU to down-grade, and then return to English to build the proposed repair. With teachers, students employ L1 clarification questions (initiators of open-class, candidate understandings) rather than overt disagreement; teachers may rekey to L1 briefly to signal trouble, then repeat the institutional version in English. Cross-role and cross-linguistic resource mapping shows that modals such as would, could, should, and need to be mainly used as English head-act resources for all activities, and the first language only rarely contributes modal morphology but provides stance particles to scale force. Particles and stance markers like *well, so, actually, maybe, and I think* in English mark epistemics and turn structure. L1 particles occur in sensitive positions like refusals, disagreement, or downgrades to indicate solidarity or to soften face-threats. Evidentials are realized in English lexical signals such as *I think, it seems, or according to* that have public warrant, whereas L1 reportative particles or hearsay occasionally emerge in peer talk as part of establishing mutual ground. English address terms and vocatives, specifically their plural forms such as guys or team, serve to manage the inclusive floor. In contrast, L1 solidarity or kinship terms occur in post-expansions to achieve compliance or to close sequences.

English conventional formulae, particularly conventional-indirect request frames and apology or concern formulas, anchor on-record action, while L1 formulae are more typical of accounts and mitigations in relation to the head act. In the case of matrix vs. embedded functions, English is the matrix for on-record assessable actions such as request heads, advice directives, and assessment findings, while L1 is embedded within grounders, stance shifts, and mitigation, especially in pre- and post-expansion positions. Finally, in terms of position within a sequence, switches cluster in pre-expansions, where they function as preparators or attention shifts, within-TCU units to provide lexical specificity or stance particles, and in post-expansions, where they provide backup accounts and ratifications, hence coordinating language choice with the on-ramp or off-ramp of sensitive actions and with floor transfer.

Sequential positioning in conversation reveals that code-switching tends to cluster at pre-, base-, and post-expansions, where participants negotiate entry into the sequence, commitment to an action, and closure. In pre-expansions, brief alternations into the first language help draw attention, preview delicate business, or install contingencies before an English head act. A teacher, for instance, may begin with *[L1] одну хвилинку, слухаємо* and then continue in English *okay let's start with your outline*, prompting the student's response *yeah (.) I'll present the intro first*. At the base of the sequence, which carries the on-record request, advice, or assessment, English remains the matrix language to ensure accountability in the EMI context. Post-expansions, however, frequently contain short L1 tag-switches that soften or ratify outcomes, as when a teacher concludes an assessment with *good (.) [L1] так тримати 'keep it up'*.

Preference organization is evident in dispreferred actions such as refusals or disagreements, which follow the well-documented pattern of delay and mitigation. Code-switching appears most often in the slots before and after the dispreferred base. Learners usually phrase the head act in English, for example, *I can't join the rehearsal today*, but then shift into L1 for the accompanying account or apology, as in *[L1] вибач, сімейні справи 'sorry, family matters'*, which recalibrates the face-threat before the peer's acceptance *no problem (.) we'll record it*.

Repair sequences also reveal systematic coordination of codes. When comprehension troubles arise, repair is often launched with English open-class initiators such as *sorry?* to maintain institutional transparency. Still, candidate understandings may be voiced in L1 if a technical term or culturally grounded label is involved, thereby speeding resolution. For instance, a suggestion *we could take the shuttle* may be followed by *sorry*, eliciting clarification through *[L1] маленький автобус 'small bus'*, which is then ratified with *ah (.) minibus*.

Self- and other-repair trajectories demonstrate further patterned use. In self-initiated self-repair, learners insert L1 fragments mid-turn to hold the floor while searching for lexis, before resuming English for the replacement, as in *we should con— [L1] як це? 'what's the word?' confirm the booking*. In other-initiated self-repair, the prompt typically comes in English, either as an open-class initiator or a repetition with interrogative intonation, and the learner supplies the fix, often adding a short L1 rationale afterward. Thus, a teacher's prompt *the thesis is…?* may be answered with *that social media helps (.) sorry [L1] забув слово 'I forgot the word'*, combining repair with an account that explains the slip.

In multilingual interaction, intra-sentential insertions into English TCUs complete precise terms or stance particles without relinquishing the floor, serving as turn-holding devices that compensate for retrieval delay caused by cognitive load or technological mediation. Code-switching at transition-relevance points regulates entry and exit: L1 prefaces offer entry to the floor in the form of pre-requests or calls to attention, and L1 tags in post-expansions offer support for alignment or incidental praise, facilitating smooth turn transfer. Learners self-select after L1 prefaces, whereas teachers use L1 strategically at TRPs to allocate turns, then switch back to English for the head act to sustain the public record.

Stance work is also finely calibrated through alternation. Epistemic stance in English, framed by expressions such as *I think* or *it seems*, provides public warrant for claims and assessments, while L1 insertions lower the cost of uncertainty by invoking shared background and proximity, as in *I think the leading cause is economic (.) [L1] мені здається 'it seems to me'*. Deontic stance relies on English modals such as *should*, *need to*, or *have to* to convey obligation, yet L1 softeners prepare the ground for directives and redistribute rights in advance of face-threatening acts, as in *[L1] можна маленьку пораду? 'may I give a small advice?' → you might reorganize the second*

paragraph. Affective stance is often split across codes: the assessable content remains in English for accountability, while positive affect and solidarity appear in L1, as in *your pronunciation improved (.) [L1] молодець 'well done'*.

Alternation thus recalibrates face-threat at three systematic pressure points. In pre-expansions, L1 is a gentle lead-in to potentially sensitive acts; in TCUs, L1 particles lubricate friction in disagreement or rejection since the counter-claim is still in English; and in post-expansions, L1 affiliative tags authorize downgrades or seal acceptance. In all these positions, English holds institutional accountability while L1 does mitigation, accounts, and affiliative work.

Participants orient to stable indexical mappings: English indexes public, recordable, evaluatable action, while L1 indexes proximity, precision, and solidarity. In peer work with high shared L1, congruent lexicalization – mixed frames within the same TCU – indexes collaborative problem-solving and accelerates proposal building. At the same time, in heterogeneous groups, brief L1 tags still do facework, but English is retained as the matrix code for mutual ground.

The overall outcome of pegging alternation to sequence slots and TRPs is more symmetrical participation and easier progressivity. Code-switching at structurally delicate junctures of points regulates entry rights, keeps long turns under lexical search, and closes vulnerable sequences without generating spirals of repair. Stance-indexical division of labor, with English doing on-record anchoring and L1 assuming mitigation and affiliation, lowers interpersonal risk while safeguarding the visibility of assessable outcomes.

Analysis across activity types and proficiency levels shows that code-switching systematically supports comprehension, uptake, and inclusion when balanced within turn design. Brief L1 pre-expansions before English head acts reduce ambiguity and clarify task entry, increasing participation though slightly lowering authenticity when overused. Within-turn L1 insertions enhance precision and efficiency, shortening repair and aiding lower-proficiency learners in maintaining the floor, provided L1 segments remain brief. Alternation for meta-planning clarifies role division and accelerates task convergence but may exclude non-shared-L1 peers unless followed by English summaries. Post-expansion L1 cushioning softens evaluations and improves uptake by reducing face-threat, while repair sequences combining English initiators with L1 candidate interpretations resolve problems faster and reinforce correct reuse.

The dominant pattern (English for core acts, L1 for mitigation) optimizes understanding and accountability. Excessive or extended L1 use, however, delays completion and increases cognitive load. Stance alignment further refines these effects: English evidentials with L1 tokens enhance epistemic uptake; English modals with L1 softeners balance deontic control; and L1 post-expansions of affective stance mitigate threat and sustain rapport. The most effective configuration is a three-part alternation – L1 preface, English head act, L1 post-cushion – maximizing comprehension, inclusion, and authenticity while maintaining manageable processing demands.

Discussion

Our results show that in multilingual English as a Foreign Language classrooms, code-switching is not a random fallback but a planned interactional resource. Alternation is synchronized with turn and sequence organization to balance interpersonal risk and institutional accountability. The recurrent pattern L1 as soft entry and affiliative exit framing an English head act helps learners anticipate the next move, legitimizes their participation, and accelerates task convergence.

These findings refine previous functional accounts of classroom switching. As in Turkish classrooms (Ataş & Sağın-Şimşek, 2021), switches help with task management and alignment, but our study adds sequential precision by showing how they cluster at pre-expansions, within-turn insertions, and post-expansion tags. The focus on where switching occurs clarifies how learners coordinate uptake in real time. Teacher stance also shapes these patterns. Gallagher (2020) shows that "bilingual space" orientations make switching legitimate; our data specify how this legitimacy is enacted, balancing English-medium instruction accountability with inclusion – English for on-record acts and L1 at transition points for entry, mitigation, or closure.

Research on translanguaging (Huang and Chalmers, 2023) reports benefits when multilingual practice links directly to tasks and assessment. Our conversation-analytic approach complements these findings by identifying the micro-mechanisms through which alternation accomplishes stance and sequence design. This view aligns with Tai (2024), who describes how teachers rekey student talk into academic registers using multilingual cues. We extend this by demonstrating the fine-grained turn-by-turn construction of such moments. At the micro-pragmatic level, our results echo Martín-Laguna and Alcón-Soler (2018), who relate discourse markers to turn organization: in our data, English markers anchor assessable head acts, while L1 particles surface in delicate sequences—before disagreement, during refusal accounts, or after downgrades—to recalibrate face-threat and maintain progressivity.

Pedagogically, both "target-language only" policies and unrestricted L1 use are inefficient. The evidence supports structured alternation at key sequence points: an L1 preface to ease entry, an English head act for clarity and accountability, and an L1 post-cushion to sustain rapport. In multilingual classes, this pattern should be reinforced by brief English recaps to maintain shared understanding. Teachers can apply these findings by planning when and why to alternate codes—for example, to preview a new task, model an academic move, or close interactional trouble—rather than avoiding or improvising switches. Training in recognizing pre-/base-/post-expansion functions could help teachers manage switching strategically, improving comprehension and inclusion without undermining authenticity.

This study is based on a 26-hour corpus from one institution within a single-term window. Results may therefore reflect local norms and teacher practices rather than generalizable patterns across contexts. The limited sample size and disciplinary setting (maritime English) restrict broader applicability. Future research should include longitudinal and cross-institutional data to track how code-switching shapes long-term pragmatic development and whether similar mechanisms appear in other English-medium instruction disciplines. Quantitative modeling of stance and uptake could further test how alternation predicts learning outcomes.

Conclusion

This study shows that code-switching in multilingual English-as-a-foreign-language classrooms is a structured, purposeful strategy aligned with conversational organization. Language alternation supports both communicative precision and social alignment: English anchors accountable classroom actions, while learners' first languages provide clarification, mitigation, and rapport. Across classroom types, switching patterns follow a stable configuration (L1 preface, English head act, and L1 post-cushion) which promotes comprehension, inclusion, and smooth task progression.

By linking cross-cultural pragmatics and conversation analysis, the study explains how alternation coordinates sequence structure, stance, and face management. It reframes multilingual

classroom discourse as an interactional resource rather than a deviation from target-language norms. Pedagogically, the findings support planned alternation over rigid monolingual policies: teachers can time brief L1 use to preview, clarify, and close tasks while maintaining English accountability. Further longitudinal and cross-institutional research should explore how such "sandwich architecture" fosters durable pragmatic and interactional competence.

Funding: This research is not funded.
Acknowledgments: Not applicable.
Conflicts of Interest: The authors declare no conflict of interest.
Authenticity: This manuscript is an original work
Artificial Intelligence Statement: The authors confirm that all AI-generated content was carefully reviewed, edited, and integrated to ensure academic integrity, originality, and compliance with scholarly standards. During the preparation of this work, the authors used AI to check the grammar and translate several paragraphs. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

About the Authors

Valeriia Smaglii, doctor of philological sciences, professor, head of the Department of Foreign Languages and Translation, Odesa National Maritime University, Odesa, Ukraine. Her research interests are Language Teaching and Learning, English for Specific Purposes, Cognitive Linguistics. ORCID ID: https://orcid.org/0000-0002-6222-7652

Svitlana Yukhymets, Candidate of Pedagogical Sciences, Associate Professor, Associate Professor of Department of Foreign Languages and Translation, Odesa National Maritime University, Odessa, Ukraine. Her research interests cover a range of linguistic problems, many of which fall within the scope of translation as well as foreign language teaching methods.
ORCID ID: https://orcid.org/0000-0003-3350-7310

Yevheniia Kornielaieva, candidate of philological sciences, associate professor, associate professor of the Department of Foreign Languages and Translation, Odesa National Maritime University, Odesa, Ukraine. Her research interests are Language Teaching and Learning, English for Specific Purposes, Lexicology, Lexical semantics.
ORCID ID: https://orcid.org/0000-0001-9516-1963

Inna Kivenko, Candidate of Philological Sciences, Associate Professor, Associate Professor of the Department of Foreign Languages and Translation, Odesa National Maritime University, Odesa, Ukraine. Research interests are Language Teaching and Learning, English for Specific Purposes, Pragmalinguistics, Cognitive Linguistics, Communicative Linguistics, Applied Phonetics, Cross-Cultural Communication
ORCID ID: https://orcid.org/0000-0003-2044-5456

References:

Alzamil, A. (2019). The Effects of the Use of First Language on Learning English as a Second Language: Attitudes of Arabic EFL Learners. *Arab World English Journal, 10* (3) 192- 201. https://dx.doi.org/10.24093/awej/vol10no3.13

Amiri, M. (2024). Code switching in EFL classroom discourse. *Journal of Second Language Teaching & Learning, 13*(2), 21–44. https://doi.org/10.22051/jsltl.2024.XXXXXX

Ataş, U., & Sağın-Şimşek, Ç. (2021). Discourse and educational functions of students' and teachers' code-switching in EFL classrooms in Turkey. *Linguistics and Education, 65*, 100981. https://doi.org/10.1016/j.linged.2021.100981

Criado, R., & González-Romero, P. (2023). Exploring First Language Use in Non-Formal Foreign Language Education: A Mixed-Method Approach. *Arab World English Journal, 14* (4) 19-45. https://dx.doi.org/10.24093/awej/vol14no4.2

Fang, X. (2024). Making refusals via English as a lingua franca: Chinese English speakers' strategies and sequences. *Journal of Pragmatics, 230*, 1–14. https://doi.org/10.1016/j.pragma.2024.103511

Gallagher, F. (2020). Considered in context: EFL teachers' views on the classroom as a bilingual space and codeswitching in shared-L1 and in multilingual contexts. *System, 91*, 102262. https://doi.org/10.1016/j.system.2020.102262

Hazaymeh, W. A. (2022). Teachers' perceptions of code-switching functions and effects in the EFL classroom. *European Journal of Educational Research, 11*(3), 1839–1849. https://doi.org/10.12973/eu-jer.11.3.1839

Huang, X., & Chalmers, H. (2023). Implementation and effects of pedagogical translanguaging in EFL classrooms: A systematic review. *Languages, 8*(3), 194. https://doi.org/10.3390/languages8030194

Kim, M. J., & (co-authors). (2025). Examining EFL learners' oral advice production: The role of pragmatic demands. *System. Advance online publication*. https://doi.org/10.1016/j.system.2025.103234

Martín-Laguna, S., & Alcón-Soler, E. (2018). Development of discourse-pragmatic markers in a multilingual classroom: A mixed-method research approach. *System, 75*, 68–80. https://doi.org/10.1016/j.system.2018.03.009

Nguyen, M. T. T., & Pham, T. T. T. (2022). Instructional effects on L2 pragmatic comprehension: The case of indirect refusals and indirect opinions. *The Language Learning Journal, 50*(4), 427–443. https://doi.org/10.1080/09571736.2022.2088443

Rafiq, R., Wang, X., Hui, P., & Wen, S. (2024). The effectiveness of online pragmatic training on essential pragmatic competence: A randomized experimental study. *Heliyon, 10(11),* eXXXXX. https://doi.org/10.1016/j.heliyon.2024.XXXXX

Rezai, A., Saeidi, M., & Dabbagh, A. (2023). Cultivating interlanguage pragmatic comprehension through graduated dynamic assessment. *Language Testing in Asia, 13,* 18. https://doi.org/10.1186/s40862-022-00179-w

Tai, K. W. H. (2024). Shaping student responses into academic expressions: Analysing an English-medium instruction history classroom from a translanguaging perspective. *International Journal of Bilingual Education and Bilingualism, 27*(4), 550–580. https://doi.org/10.1080/13670050.2023.2232089

Temesgen, A., & Hailu, E. (2022). Teachers' codeswitching in EFL classrooms: Functions and motivations. *Cogent Education, 9*(1), 2124039. https://doi.org/10.1080/2331186X.2022.2124039

Wang, Y., Ghayth, K., & co-authors. (2024). Pragmatic competence of EFL and ESL learners and its influencing factors: A review. *Eurasian Journal of Applied Linguistics, 10*(3), 122–140. https://doi.org/10.32601/ejal.10312

Yasar Yüzlü, M., & Dikilitaş, K. (2022). Translanguaging in the development of EFL learners' foreign language skills in Turkish context. *Innovation in Language Learning and Teaching, 16*(2), 176–190. https://doi.org/10.1080/17501229.2021.1892698

Yüzlü, M. Y. (2022). The impact of translanguaging-driven in-service training on EFL teachers' professional identity reconstruction. *Linguistics and Education, 71*, 101079. https://doi.org/10.1016/j.linged.2022.101079

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp. 343-377
DOI: https://dx.doi.org/10.24093/awej/vol16no4.20

Mobile-Assisted Language Learning in Tertiary EFL Education: A Decade of Systematic Evidence (2015–2025)

Li Chengli
Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Malaysia, 43600
Corresponding author**:** p131127@siswa.ukm.edu.my

Harwati Hashim
Centre of Learning and Teaching Innovation, Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Malaysia, 43600;
Email**:** harwati@ukm.edu.my

Supyan Hussin
Institute of Ethnic Studies, Universiti Kebangsaan Malaysia, Bangi, Malaysia, 43600;
Email**:** supyan@ukm.edu.my

Received:08/14/2025 Accepted: 11/11/2025 Published:12/10/2025

Abstract
The rapid advancement of mobile technologies has reshaped global education, particularly language learning in higher education. The systematic literature review of mobile-assisted language learning in tertiary EFL contexts from 2015 to 2025 aims to synthesise empirical evidence on how mobile technologies have been implemented and examined in English learning, focusing on key trends, pedagogical strategies, and theoretical frameworks. The significance of this study lies in consolidating fragmented research to provide evidence-based insights for educators, researchers, and policymakers seeking to integrate mobile tools effectively into tertiary English education. Following the PRISMA guidelines, 42 empirical studies were analysed from major academic databases. Results show a clear shift from technology-focused studies to pedagogically and theoretically integrated approaches, indicating that MALL has become a learner-centred and context-sensitive practice that supports both language acquisition and instructional design in tertiary EFL education. Overall, the findings highlight the growing pedagogical potential of MALL in fostering active and autonomous learning, underscoring its impact on modernising tertiary EFL instruction and shaping future research and policy directions.
Keywords: EFL learners, higher education, mobile-assisted language learning, systematic literature review

How to cite: Chengli, L., Hashim, H., & Hussin, S. (2025). Mobile-Assisted Language Learning in Tertiary EFL Education: A Decade of Systematic Evidence (2015–2025). *Arab World English Journal, 16*(4): 343-377. https://dx.doi.org/10.24093/awej/vol16no4.20

Introduction

The ubiquity of mobile technologies has transformed global education, with language learning being among the most profoundly affected fields (Kukulska-Hulme & Shield, 2008; Pegrum, 2014). Mobile-Assisted Language Learning (MALL) marks a significant departure from conventional classroom-based practices, introducing flexible, learner-centered approaches that overcome the limits of time and place (Sharples et al., 2005; Stockwell, 2013). This transformation is especially visible in higher education, where it has become an especially promising context for adopting MALL, shaped by several intersecting conditions. University learners generally display strong digital skills and have reliable access to mobile devices (Dou et al., 2025; Jeong, 2023). In addition, the self-directed nature of tertiary studies aligns closely with the independent learning opportunities that mobile technologies provide(Shen et al., 2023; Wu, 2015). Over time, research on MALL has progressed from initial studies assessing technical viability to more advanced inquiries addressing pedagogical effectiveness, learner perceptions, and theoretical integration with established models of language acquisition (Hwang & Fu, 2019; Kacetl & Klímová, 2019). The field has expanded considerably, with investigations examining a wide range of applications including vocabulary growth, reading and writing skills, and oral communication. This broader focus mirrors both technological progress and enhanced insight into how mobile tools can support multiple dimensions of language learning (An et al., 2021; Lai et al., 2022).

Although numerous empirical studies have been published, the field still lacks comprehensive syntheses that could offer solid, evidence-informed recommendations to teachers, researchers, and policymakers (Hema & Gandhimathi, 2025). This absence is particularly problematic in light of the considerable financial commitments made to educational technologies and the increasing emphasis on evidence-based decision-making in higher education (Guo et al., 2024; Lazaro & Duart, 2023). Accordingly, the main aim of this systematic review is to critically examine and synthesise empirical research on MALL in tertiary English as a foreign language (EFL) education from 2015 to 2025, with the goal of providing a comprehensive and evidence-based understanding to inform future research, instructional practice, and policy development in higher education.

Several pressing issues underscore the significance of this review. To begin with, the COVID-19 pandemic accelerated the integration of mobile technologies in higher education and reshaped research priorities. Rather than focusing solely on adoption, studies began examining learner autonomy, digital equity, and the sustainability of technology-mediated pedagogy. This shift reflects a move toward viewing MALL not just as a technological tool but as a pedagogical response that supports continuity, interaction, and engagement during educational disruptions (Jeong, 2023; Lee, 2021). This rapid transition underscored both the opportunities and constraints of mobile technologies in facilitating language learning, pointing to the need for a systematic assessment of existing research (Šramová, 2024; Zou et al., 2021).

In addition, the diversity of existing MALL studies poses difficulties for practitioners who wish to apply evidence-driven practices. Research differs widely in terms of theoretical grounding,

research design, learner populations, and outcome measures, which makes deriving clear conclusions about effectiveness or best practices challenging(Li et al., 2022; Mihaylova et al., 2022). A systematic review is therefore essential to map patterns, consolidate insights, and identify both convergences and inconsistencies across the literature (Guo et al., 2024).

Furthermore, the rapid evolution of digital technologies means that results from earlier MALL investigations may have limited applicability to present-day contexts (Hu et al., 2023; Li et al., 2023). By targeting research published between 2015 and 2025, this review encompasses the most up-to-date evidence while ensuring sufficient scope to detect emerging trends and developments.

The decision to focus specifically on tertiary-level EFL education is driven by the unique nature of this learner group. University students differ significantly from younger learners in terms of their autonomy, metacognitive capacities, and motivational orientations (Alamer & Al Khateeb, 2023; Peng et al., 2024; Rashid et al., 2021). Moreover, tertiary EFL contexts are typically shaped by explicit academic and professional objectives that influence both learning strategies and expected outcomes (Wang et al., 2024).

This Systematic Literature Review (SLR) aims to consolidate empirical studies on the use of MALL in tertiary-level English language education published between 2015 and 2025. The objectives of the study are as follows: to explore trends in the integration of MALL in English language instruction at the university level; to examine the theoretical and instructional frameworks that underpin MALL practices in higher education settings. Based on these objectives, the following research questions guide this review:

RQ1: What are the key themes and trends in empirical studies on MALL in tertiary English education over the past decades?

RQ2: Which pedagogical strategies, theoretical models, and research methods are most commonly employed in MALL research at the tertiary level?

By addressing these questions, this review offers a decade-long synthesis of empirical evidence (2015–2025) that maps developmental trajectories, identifies dominant theoretical and pedagogical trends, and critically evaluates methodological rigour across studies. In doing so, it contributes not only a comprehensive evidence base but also an intellectual framework that clarifies how mobile technologies have reshaped communicative language learning in higher education contexts.

Literature Review

To situate the present study within the broader field, this section reviews the theoretical foundations and historical trajectory of MALL. It also identifies critical gaps in existing systematic reviews, thereby clarifying the necessity and contribution of the current research.

Theoretical Foundations of Mobile-Assisted Language Learning

Mobile-Assisted Language Learning (MALL) involves the use of handheld devices to support language learning both inside and outside the classroom, through applications, multimedia,

and communication tools (Kukulska-Hulme, 2009). The conceptual basis of MALL is deeply interdisciplinary, integrating perspectives from educational psychology, applied linguistics, and technology studies. This breadth reflects the diverse ways in which mobile technologies mediate language acquisition and create opportunities for both formal and informal learning (Godwin-Jones, 2017b; Kukulska-Hulme & Shield, 2008). Current MALL scholarship frequently draws upon several influential theoretical perspectives that help explain how mobile tools can enhance second language learning processes.

Technology Acceptance Models

One of the most widely employed frameworks in MALL research is the Technology Acceptance Model (TAM) proposed by Davis (1989). TAM suggests that learners' intention to adopt technology is influenced primarily by two perceptions: its usefulness and its ease of use. In mobile learning contexts, TAM has been extensively applied to investigate how students evaluate and accept mobile applications for language learning, often showing robust explanatory power across different cultural and institutional environments (Venkatesh & Davis, 2000).

As TAM has evolved, researchers have developed extended models such as the Unified Theory of Acceptance and Use of Technology (UTAUT), which introduces variables like social influence and facilitating conditions (Venkatesh et al., 2003). UTAUT has gained traction in MALL research, especially in tertiary education, where peer influence and institutional infrastructure play significant roles in adoption. For instance, Almaiah et al. (2022) found that institutional support and peer encouragement strongly predict university students' willingness to engage with MALL platforms. These findings suggest that acceptance models remain crucial for understanding the sociotechnical dimensions of MALL. Empirical studies on MALL have consistently supported the assumptions of TAM, showing that learners' perceived usefulness and ease of use significantly influence their willingness to adopt mobile technologies for English language learning.

Constructivist Learning Theory

Constructivist theories, particularly those of Piaget (1977) and Vygotsky (1978), emphasize learners' active role in constructing knowledge through engagement with tasks and social interactions. This perspective resonates strongly with MALL, where learners interact with authentic, real-life materials and contexts via mobile devices (Kukulska-Hulme, 2020).

Social constructivism is especially relevant to MALL because of its focus on collaborative learning and peer interaction. Mobile platforms facilitate communication across time and space, enabling learners to work in communities of practice, share experiences, and receive feedback (Lai et al., 2022).The collaborative affordances of messaging apps, voice-recording tools, and online forums align closely with constructivist principles, supporting both individual reflection and collective meaning-making (Ma & Chiu, 2024). Evidence from recent MALL research aligns with constructivist principles, demonstrating that mobile-based collaborative tasks and interactive activities foster active knowledge construction and meaningful learner engagement.

Motivation Theory

A third theoretical pillar in MALL research is motivation theory, particularly Self-Determination Theory (SDT) by Deci and Ryan (2013). SDT posits that intrinsic motivation flourishes when learners experience autonomy, competence, and relatedness. Applied to MALL, SDT has been used to explain how mobile tools can enhance learner engagement and persistence (Ryan & Deci, 2000). Mobile applications that provide immediate feedback on performance or allow learners to set their own pace can strengthen their sense of competence and autonomy (Shen et al., 2023). Similarly, collaborative mobile tasks create opportunities for relatedness by connecting learners socially (Oga-Baldwin et al., 2023; Zakaria et al., 2025). Empirical findings further confirm the relevance of SDT in MALL contexts, indicating that mobile-supported learning environments enhance learners' intrinsic motivation by promoting autonomy, competence, and relatedness.

Self-Regulated Learning

Self-regulated learning (SRL) offers a useful lens for understanding how MALL supports tertiary EFL development. Building on classic models in which learners cyclically forethink, enact strategies, and self-reflect (Pintrich, 2000; Zimmerman, 2002), recent empirical studies provide convergent evidence of mobile technologies' facilitative role in fostering SRL skills: for example, Boroughani et al. (2023) demonstrated that mobile-assisted SRL interventions significantly enhanced university students' academic vocabulary acquisition compared to traditional methods. Similarly, a longitudinal quantitative study found that sustained MALL involvement improved both vocabulary learning attitudes and self-regulatory capacity (Xiao et al., 2022). A 2025 systematic review further confirmed that MALL environments afford interactive, adaptive, and collaborative supports that reinforce learner autonomy and SRL, leading to improved motivation and language proficiency (Pan et al., 2025). Collectively, these findings suggest that MALL not only provides language exposure, but also scaffolds learners' ability to regulate their own learning. Empirical evidence supports the application of SRL theory in MALL, as studies reveal that mobile tools facilitate goal-setting, strategy use, and self-monitoring, thereby strengthening learners' capacity for independent language learning.

Evolution of MALL Research in Tertiary Level

The trajectory of MALL research can be divided into distinct phases, each reflecting shifts in technological capability, pedagogical orientation, and methodological rigor.

Early Phase (2000–2010)

The first decade of MALL scholarship was largely exploratory. Mobile devices of this period were technologically limited, leading to small-scale feasibility studies focused on SMS-based vocabulary practice or audio listening tasks (Kukulska-Hulme & Shield, 2008). Most studies lacked strong theoretical grounding and involved small samples, but they laid the foundation by establishing mobility as a distinctive affordance of mobile learning (Sharples et al., 2005). These

pioneering efforts demonstrated that learning could extend beyond classrooms and laboratories, sparking broader interest in mobile pedagogy.

Expansion Phase (2011–2020)

With the widespread availability of smartphones and tablets, MALL research entered a more sophisticated phase. Scholars began integrating established language learning theories with mobile technologies, exploring multimedia applications, adaptive learning systems, and gamification (Burston, 2015; Pegrum, 2014). Research methodologies became more rigorous, with randomized controlled trials and large-scale surveys increasingly common (Sung et al., 2016). At the same time, practical concerns such as digital equity, privacy, and institutional integration emerged as central issues (Godwin-Jones, 2017a; Habbash, 2020; Stockwell, 2010).

Maturation Phase (2021–Present)

The current phase of MALL is marked by methodological rigor and theoretical depth (Naveed et al., 2023). Recent research has shifted from simple adoption studies to nuanced analyses of learning outcomes, contextual variables, and learner differences (Khezrlou et al., 2025). Mixed-methods and longitudinal designs are now more frequently used to capture complex patterns of learner engagement and skill development (Bai, 2024). Additionally, studies increasingly consider cultural and contextual diversity, recognizing that MALL effectiveness varies across different educational systems and learner populations (Hossain, 2024). More recently, emerging research has incorporated artificial intelligence (AI) and adaptive learning systems to personalize instruction, provide real-time feedback, and enhance learner autonomy, marking a new stage of technological sophistication in MALL (Gligorea et al., 2023; Warschauer & Xu, 2024).

Although systematic literature reviews on technology-enhanced and mobile-assisted language learning have been conducted, several critical gaps remain that justify the present study.

1. Skill-Specific Focus: Much of the existing synthesis has focused on specific language skill (Nawaz et al., 2025; Shavkatovna et al., 2024) rather than addressing mobile-assisted language learning in tertiary English as a foreign language contexts. Reviews often cover broad educational settings, which limits their applicability to higher education where learner autonomy, institutional policies, and curricular integration play distinctive roles (Naveed et al., 2023).

2. Limited Temporal Scope: Prior reviews have been constrained by relatively narrow timeframes, often stopping before the recent acceleration of mobile learning research driven by rapid technological innovation and the global shift toward blended and online learning (Zhou et al., 2023). Consequently, more recent empirical studies published between 2020 and 2025, which reflect the impact of new applications such as AI-powered platforms, gamification, and adaptive systems (Khosrawi-Rad et al., 2025; Ma et al., 2025; Ren & Tan, 2025), have not been systematically synthesized.

3. Lack of Contextual Analysis: Earlier reviews have tended to focus on learning outcomes without sufficiently examining contextual moderators such as cultural differences, learner

characteristics, and institutional support mechanisms (Khezrlou et al., 2025; Okumuş Dağdeler, 2023).

4. Theoretical Narrowness: Existing reviews rarely integrate multiple theoretical perspectives, instead relying on single frameworks such as the Technology Acceptance Model, which provides only partial explanations of the complex pedagogical processes in mobile-assisted learning (Almaiah et al., 2022).

Addressing these gaps, this review synthesizes empirical evidence from 2015 to 2025, offering a comprehensive and context-specific understanding of mobile-assisted language learning in tertiary English as a foreign language education.

Methodology

This review was conducted following the **Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)** framework (Page et al., 2021), which guided the systematic identification, screening, eligibility, and inclusion of studies. A detailed checklist and flow diagram were applied to document the selection process.

The protocol for this review was developed in advance to specify the research objectives, eligibility criteria, search and selection strategies, data extraction methods, and synthesis procedures (Gough et al., 2012). These procedures ensured consistency and traceability throughout the review.

To strengthen reliability and validity, databases were searched using systematic strategies, and all records were independently screened. Standardized data extraction forms were used to compile and compare study characteristics. Following recommendations for methodological transparency and replicability in educational research (Torgerson, 2003), the review process was documented in detail to facilitate verification.

Databases and Search Strategies

Research was conducted using two premier academic databases, Scopus and Web of Science (WoS), chosen for their extensive and complementary coverage of peer-reviewed literature across disciplines. As literature in scientometrics affirms, combining these “titans of bibliographic information” enhances completeness and reduces bias when performing systematic reviews (Pranckutė, 2021). From this extensive search, 42 empirical studies were ultimately selected for inclusion in the systematic review. This number was deemed sufficient to achieve data saturation, as it reflects a comprehensive coverage of relevant MALL research within the defined timeframe and meets the sample adequacy standards recommended for systematic literature reviews (Petticrew & Roberts, 2005).

The search strategy was refined through iterative testing and pilot searches. The final query integrated three concept clusters with Boolean operators. See Figure 1 below.

Three Concept Clusters with Boolean Operators	
• **Mobile Technology**	(mobile OR smartphone OR tablet OR "mobile device" OR "mobile application" OR "mobile app" OR MALL OR "mobile learning" OR "m-learning")
• **Language Learning**	(language AND (learning OR acquisition OR education OR instruction) OR "English as a foreign language" OR EFL OR "English language learning" OR "second language" OR L2)
• **Educational Level**	(university OR college OR "higher education" OR tertiary OR undergraduate OR "post-secondary")

Figure 1. Three concept clusters with Boolean operators

Search terms were derived from prior reviews and recurrent keywords in seminal MALL publications (Burston, 2015; Khezrlou et al., 2025; Naveed et al., 2023; Zhou et al., 2023). Synonyms and variations were included to maximize sensitivity, while Boolean and proximity operators were applied to maintain precision.

To maintain relevance and quality, several filters were applied. See Figure 2 below.

Search Limitations and Filters	
· **Time frame**	2015–2025 (reflecting the maturation of mobile technologies and ensuring inclusion of the latest research)
· **Language**	English (to ensure comparability and accessibility)
· **Document Type**	Peer-reviewed journal articles
· **Subject Areas**	Social sciences, computer science, arts and humanities

Figure 2. Search limitations and filters

The cutoff starting at 2015 aligns with the widespread adoption of smartphones and the expansion of mobile learning applications, which makes research from this period highly applicable to contemporary practices (Burston, 2015; Metruk, 2024). These parameters ensured the review concentrated on high-quality, directly relevant research. In total, 320 **articles** were retrieved through databases searches and subjected to preliminary screening.

Screening and Selection Process

Study selection followed a two-step process consistent with PRISMA. First, titles and abstracts were screened against eligibility criteria. 169 articles deemed potentially relevant at this stage were retained for full-text review to avoid premature exclusion. The second stage of the review involved a detailed full-text assessment to ensure that studies aligned with the population, intervention, outcomes, and design criteria of the research scope. A standardized screening form was employed to record reasons for exclusion, thereby maintaining transparency and consistency in decision-making. See Figure 3 below.

Inclusion Criteria	Exclusion Criteria
· **Population:**University or college students learning English as a foreign language	Focused on primary or secondary school learners, or adults outside formal tertiary education
· **Setting:** Higher education institutions (colleges, universities, or equivalents)	Addressed languages other than English, or ESL in predominantly English-speaking contexts
· **Intervention:** Use of mobile devices or apps for language learning purposes	Relied on non-mobile or desktop-based technologies
· **Design:**Empirical studies (quantitative, qualitative, or mixed methods)	Consisted of literature reviews, conceptual papers, or editorials without empirical data
· **Outcomes:** Assessment of language proficiency, learner perceptions, or implementation processes	Provided insufficient methodological detail or demonstrated significant flaws in design
	Were inaccessible in full text or not available in English

Figure 3. Inclusion/Exclusion criteria

Following this, the methodological quality of the included studies was appraised using criteria adapted from established educational research evaluation frameworks (Thomas & Harden, 2008). Key dimensions of assessment included the clarity of research objectives, appropriateness of study design, adequacy of sample size and recruitment procedures, validity and reliability of instruments, suitability of data analysis methods, and transparency in reporting.

Through this rigorous two-step process of screening and appraisal, a total of 42 studies were identified as meeting the inclusion criteria and were retained for the final review. See Figure 4 below.

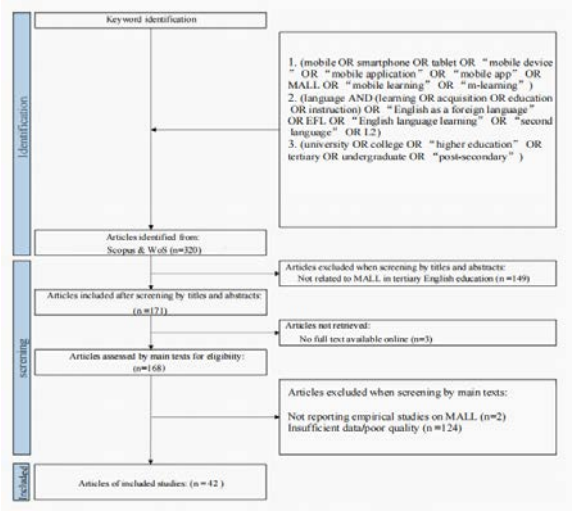


Figure 4. PRISMA flow diagram

Data Extraction and Synthesis

A structured data extraction form was developed to ensure consistency and comparability across studies included in this review. The form systematically captured study characteristics (author(s), year, country, research design, and participant demographics), intervention characteristics (type of device, application or platform employed, duration, intensity, and underlying theoretical foundations), outcomes (primary and secondary measures, along with major findings), and challenges (institutional barriers, limitations in integration with formal instruction,

cultural constraints, and issues related to technical support). This structured approach aligns with best practices in systematic reviewing, as it facilitates transparency, reduces reviewer bias, and enables meaningful synthesis of heterogeneous evidence (Petticrew & Roberts, 2005; Thomas & Harden, 2008).

Due to the diversity of methodologies, interventions, and outcome measures, the review employed **narrative synthesis** rather than meta-analysis. Following Popay et al. (2006), in synthesizing the findings, a structured multi-dimensional approach was adopted to ensure rigor and transparency. Thematic analysis was employed to group studies according to intervention type, theoretical orientation, and reported outcomes, while temporal analysis was conducted to trace changes in research foci and methodological approaches over time (Braun & Clarke, 2006; Thomas & Harden, 2008). In this process, thematic codes were derived through a hybrid approach that combined inductive and deductive reasoning: initial codes emerged inductively from patterns identified within the empirical data, while subsequent refinement and categorization were guided by established theoretical frameworks in mobile-assisted and technology-enhanced language learning. In addition, a quality-weighted interpretation strategy was used, assigning greater weight to studies with stronger methodological rigor, consistent with established guidance for systematic reviews in educational research (Gough et al., 2012; Petticrew & Roberts, 2005). The synthesis process was iterative, involving continuous reference back to primary sources to maintain accuracy, while tables and visual representations were developed to systematically summarize study characteristics and highlight emerging trends. This structured methodological process ensured that the review produced a comprehensive and credible synthesis of the evidence on MALL in tertiary EFL contexts.

Findings

The systematic review yielded **42 empirical studies** that satisfied the eligibility criteria. Collectively, these works reflect a wide-ranging exploration of MALL within tertiary-level EFL contexts between 2015 and 2025. They span diverse theoretical foundations, methodological traditions, and types of mobile learning tools, forming a solid base for understanding the current landscape of MALL research in higher education.

Publication Trends and Key Themes

In addressing RQ1, the findings are structured across two dimensions. The analysis first considers temporal and regional distributions, and then explores thematic variations.

Temporal and Regional Distribution of Publications

An analysis of publication patterns reveals a clear upward trajectory in MALL-related research within tertiary EFL education over the review period (2015–2025). See Figure 5 below. The number of eligible empirical studies increased markedly over time, reflecting growing scholarly interest and the accelerating integration of mobile technologies in language education. Only one qualifying study was published in 2015, and in 2016, no studies met the inclusion criteria,

although this does not necessarily indicate an absence of MALL-related publications during that year. Rather, it reflects the stringent eligibility requirements of this review, which excluded studies that did not align with the defined population, intervention, outcomes, or methodological standards. The volume began to rise from 2017 (n = 3) and peaked in 2024 (n = 10). Given that the cut-off date for the literature search was May 14, 2025, only a limited number of studies from 2025 (n = 3). were included. Notably, 2021 (n = 6), 2022 (n = 7), and 2023 (n = 5) also demonstrated heightened research activity. This pattern aligns with the broader post-pandemic shift toward digital and mobile learning, particularly in higher education contexts (Lazaro & Duart, 2023). The consistent increase over the decade signals not only the maturation of the MALL research field but also its sustained relevance in the evolving landscape of technology-enhanced language learning.

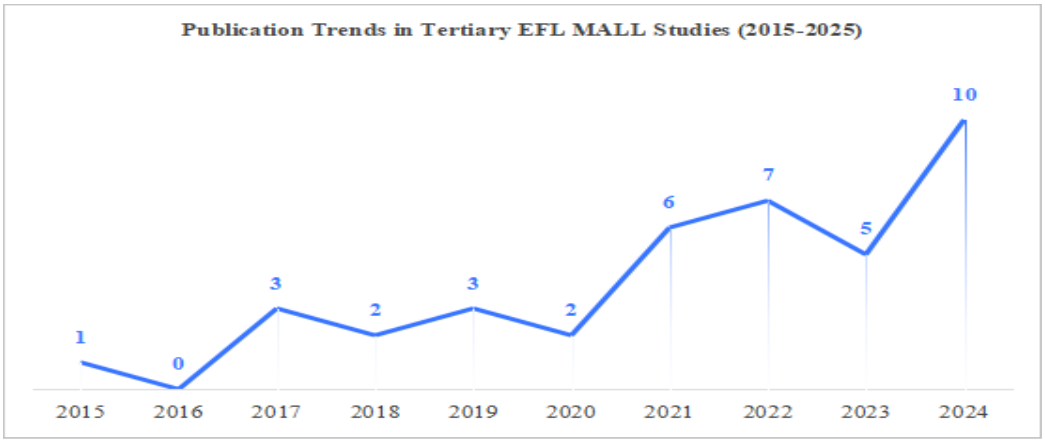


Figure 5. Publication trends of the included studies (2015–2025)

The most active years were **2021–2024**, which accounted for the bulk of the reviewed studies. This surge reflects both the **maturing of mobile technologies** and the acceleration of digital learning prompted by the COVID-19 pandemic (Hasumi & Chiu, 2024).

The studies were geographically diverse (See Figure 6 below.), though concentrated in **East Asia and the Middle East**. China dominated the sample, contributing over one-third of the studies (38%), followed by Turkey (16.7%). Other represented regions included Hongkong, Iran, Japan, Pakistan, Saudi Arabia and Vietnam. The prominence of Asian contexts is notable, as these regions combine high mobile penetration rates, strong policy support for educational technologies, and culturally specific challenges that influence the uptake and success of MALL interventions (Guo et al., 2024).

Taken together, the temporal and regional patterns reveal that MALL research in tertiary EFL education has evolved from sporadic early efforts into a sustained and geographically concentrated scholarly movement. The steady rise in publications mirrors both technological advancement and pedagogical adaptation, suggesting that MALL has moved beyond experimentation toward institutional integration. The dominance of Asian contexts indicates that regional educational priorities and infrastructure play a crucial role in shaping the global research

landscape. Collectively, these findings highlight that while MALL has achieved maturity as a research domain, its continued expansion depends on cross-regional collaboration and comparative inquiry to capture diverse pedagogical and cultural dynamics.

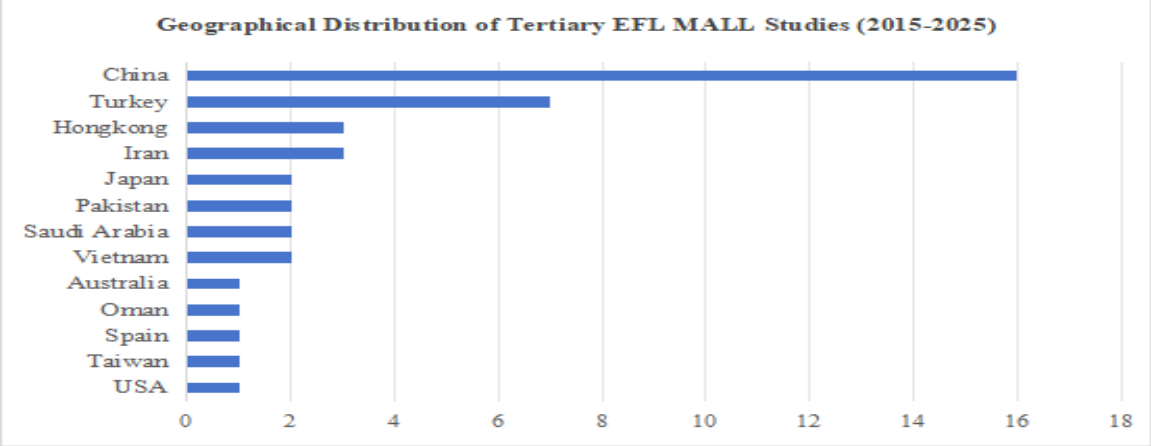


Figure 6. Geographical distribution of the included studies

Temporal and Regional Distribution of Themes

Early MALL studies (2015-2020) in the context of university English instruction predominantly focused on vocabulary, with subsequent attention given to speaking, listening, reading, writing, translation and interpretation. Nearly half of the reviewed studies investigated how mobile tools enhanced **specific aspects of English proficiency**. See Figure 4 below.

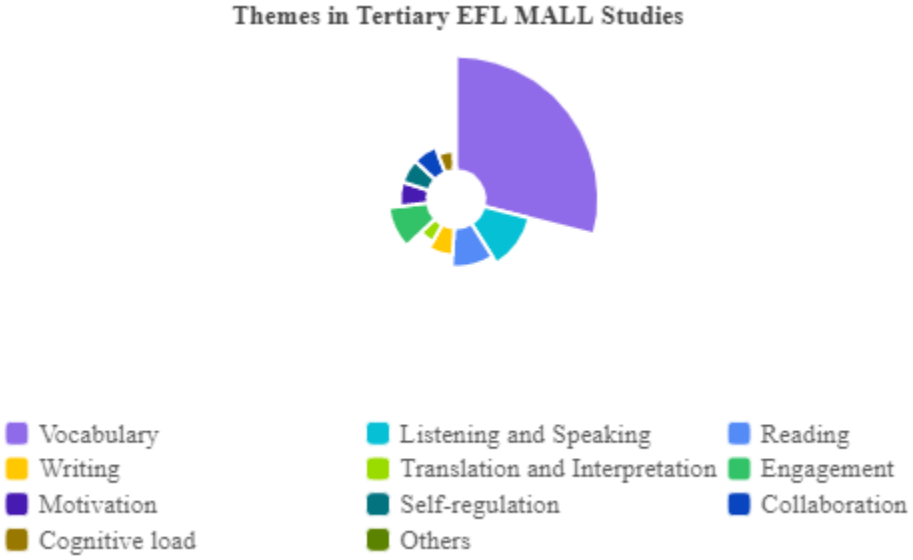


Figure 7. Thematic Diversity of the Included Studies

Vocabulary is the most widely examined skill, with approximately 29% of the reviewed studies explicitly focusing on vocabulary acquisition and retention. Wu (2015), and Zakian et al. (2022) reported medium-to-large effect sizes for mobile vocabulary learning. Effective strategies included **SRS-supported think-pair-share (Çelik & Baran, 2022)**, contextualized

reading support (Uz Bilgin & Tokel, 2019), **gamification** (Önal et al., 2022), and multimedia-enhanced vocabulary apps (Alhujaylan, 2024).

With respect to speaking and listening, the reviewed studies employed a range of approaches, with approximately 12% of the analyzed articles focusing explicitly on communication and oral expression skills. These included the use of speech recognition technologies for pronunciation practice (Dai & Wu, 2023), mobile applications designed for listening and speaking comprehension(Fang et al., 2021; Teeter, 2017), video-based oral tasks to promote interactive communication skills (Aşık et al., 2024; Huang, 2021) and authentic communication practices via chat platforms (Andujar & Salaberri-Ramiro, 2021; Jiang & Li, 2018).

In the domain of reading comprehension, approximately 10% of the reviewed studies specifically investigated how mobile technologies support learners' reading processes, highlighting tools such as e-dictionaries (Zhang & Pérez-Paredes, 2024), annotation functions (Yang & Hu, 2024), and scaffolded tasks that facilitated comprehension and engagement with authentic texts , and consistently reported positive effects, particularly for lower-proficiency learners who benefited from instant glossaries and comprehension scaffolds. Moreover, customizable support features were found to exert a notably strong impact on learning outcomes.

In terms of writing, with approximately 7% of the reviewed studies specifically addressing writing skills as a core outcome, research indicates that MALL applications effectively enhance proficiency by fostering greater learner engagement (Rashid et al., 2021). Providing real-time feedback, and supporting iterative revision processes (Aghajani & Adloo, 2018), such mobile-mediated scaffolding was particularly beneficial in improving academic writing accuracy and fluency. However, findings from a Japanese study revealed that students who wrote by hand demonstrated significantly greater recollection of content than those who composed on smartphones (Lee, 2021).

Notably, in the area of translation and interpretation, only about 5% of the reviewed studies focused on these skills, underscoring translation as the least explored dimension in this field. Studies show that MALL tools contribute positively to skill development, mainly through mobile dictionaries, collaborative platforms, and context-sensitive resources (Lai et al., 2022). Learners reported improved accuracy, efficiency, and confidence when performing interpreting-related tasks in authentic settings (Chan, 2024).

In the most recent five-year period (2021–2025), the thematic orientation of MALL research in tertiary EFL education has demonstrated a noticeable shift, with increasing scholarly attention directed toward learner-centered constructs such as engagement (Yang & Hu, 2023), motivation (Alamer & Al Khateeb, 2023), self-regulation (Lai et al., 2022; Li & Kim, 2024), collaboration (Hu et al., 2023; Huang, 2021), and cognitive load (Lee, 2021; Yang & Hu, 2024). This trend reflects a move beyond purely technological or skill-based investigations toward a more nuanced exploration of psychological, social, and cognitive dimensions that shape the effectiveness of MALL in higher education contexts.

Apart from the temporal evolution of research themes, geographical contexts further shaped the focus of MALL scholarship, highlighting the interplay between regional priorities and thematic orientations. Studies conducted in East Asia tended to foreground adoption patterns and the influence of cultural factors (Lee, 2021; Ma, 2017; Zhang et al., 2025). Research from the Middle East placed stronger emphasis on language skill development and context-driven instructional practices (Aghajani & Adloo, 2018; Uz Bilgin & Tokel, 2019; Zakian et al., 2022). By contrast, investigations in Western contexts highlighted learner experiences and individual differences as central themes (Bikowski & Casal, 2018; Li & Kim, 2024).

Synthesizing these thematic and regional patterns, it becomes evident that MALL research in tertiary EFL education has transitioned from a skill-specific orientation toward a more holistic understanding of learning processes. The early dominance of vocabulary and discrete language skills reflects the initial phase of technological experimentation, whereas the recent emphasis on engagement, motivation, and self-regulation signifies a paradigmatic shift toward learner-centered and theory-driven inquiry. Moreover, regional distinctions suggest that sociocultural and institutional factors significantly shape how MALL is conceptualized and implemented. Collectively, these developments demonstrate that the field is moving toward greater theoretical maturity and contextual sensitivity, positioning MALL as a multidimensional construct that integrates linguistic, cognitive, and affective dimensions of language learning.

Pedagogical, Theoretical, and Methodological Dimensions of MALL Research

To address RQ2, the findings are organized into three interrelated dimensions. The first examines the pedagogical strategies and instructional design approaches employed in MALL research. The second highlights the theoretical models that have provided conceptual grounding for these studies. The third considers the methodological approaches most frequently adopted, reflecting how empirical investigations in this field have been conducted.

Pedagogical Strategies and Instructional Design

The reviewed studies revealed a variety of pedagogical strategies for integrating mobile technologies into tertiary EFL instruction. Blended learning emerged as a widely adopted approach, where mobile applications complemented classroom-based teaching (Ali et al., 2024; Avci & Adiguzel, 2017; Rashid et al., 2021). Evidence indicated that such integration enhanced both learner satisfaction and learning outcomes, particularly when instructional tasks were well aligned, teachers received appropriate training, feedback was continuous, and scheduling remained flexible (Chen, 2022; Nami, 2020; Yu et al., 2019). Collaborative learning was another prominent strategy, with findings showing that peer-to-peer interaction through social networks or mobile-supported platforms fostered greater engagement and improved language performance, provided that collaborative tasks and tools were purposefully designed (Aghajani & Adloo, 2018; Avci & Adiguzel, 2017; Huang, 2021). In addition, game-based learning gained attention for its potential to stimulate learner motivation through gamified features such as rewards and narrative elements. However, results on its effectiveness for improving language proficiency were mixed, suggesting

that outcomes depended largely on the quality of game design and its alignment with instructional objectives (Önal et al., 2022; Taskiran, 2019).

Overall, the reviewed evidence suggests that successful MALL implementation in tertiary EFL education hinges on intentional design rather than mere technological adoption. The consistent finding across blended, collaborative, and game-based learning studies is that pedagogical alignment — between mobile affordances, learning objectives, and instructional scaffolding — is the key determinant of effectiveness. While technology enhances flexibility, interactivity, and motivation, its pedagogical value materializes only when integrated within coherent instructional frameworks that actively engage learners in authentic and collaborative communication. This synthesis underscores that MALL is most effective when grounded in sound pedagogical design principles that balance innovation with purposeful learning intent.

Theoretical Models

The theoretical underpinnings of the reviewed studies reveal a wide-ranging but interconnected body of frameworks guiding MALL research in tertiary EFL education. Broadly, these can be organized into four clusters: technology adoption models, cognitive and self-regulation frameworks, constructivist and sociocultural perspectives, and pedagogical/task-based approaches.

The most prominent cluster comprises technology adoption models, which focus on explaining learners' acceptance and sustained use of mobile technologies. The Technology Acceptance Model (TAM) has provided a foundational lens for investigating how perceived usefulness and ease of use shape learners' behavioral intentions (García Botero et al., 2018; Hoi & Mu, 2021; Zhang et al., 2025). Building upon TAM, the Unified Theory of Acceptance and Use of Technology (UTAUT) extends this perspective by incorporating social influence and facilitating conditions, thus offering a more comprehensive account of adoption in diverse educational settings (Hoi, 2020; Hu et al., 2023; Ren & Tan, 2025). Closely aligned, the Theory of Planned Behavior (TPB) has been employed to capture the attitudinal and motivational determinants of learners' engagement with MALL (Ren & Tan, 2025; Wang et al., 2022). Together, these frameworks form a coherent strand of research concerned with understanding technology acceptance and behavioral intention in higher education contexts.

A second major cluster centers on constructivist and sociocultural perspectives, which foreground the social and contextual dimensions of learning. Central here is Vygotsky's Zone of Proximal Development (ZPD), frequently invoked to illustrate how mobile technologies can scaffold learners' progression through dynamic support (Fang et al., 2021; Huang, 2021; Ma, 2017). Broader constructivist orientations, including Social Constructivism and General Constructivist Theory, frame MALL as a process of collaborative knowledge construction (Chan, 2024; Rashid et al., 2021). Related perspectives such as Sociocultural Theory, Social-Contextual Theory, and Situated Learning Theory further stress that learning is inseparable from cultural, institutional, and interactional contexts (Chen, 2022; Lai & Zheng, 2018; Uz Bilgin & Tokel, 2019).

The third cluster draws on cognitive and self-regulation frameworks, reflecting interest in how learners process, regulate, and sustain learning in mobile environments. Self-Determination Theory and Self-Regulated Learning models highlight the importance of autonomy, goal-setting, and metacognitive strategies, positioning learners as active agents in their mobile learning journeys (Alamer & Al Khateeb, 2023; Lai et al., 2022; Li & Kim, 2024). Cognitive Load Theory provides complementary insights into instructional design, emphasizing the need to manage cognitive resources and avoid overload when tasks are delivered through mobile platforms (Yang & Hu, 2024; Yu et al., 2019). The Flow Theory, though less frequently applied, contributes by explaining how immersive and intrinsically motivating experiences can sustain learner engagement (Hu et al., 2023).

Finally, a smaller but significant cluster involves pedagogical and task-based approaches, which translate theoretical insights into instructional design. Task-Based Language Teaching (TBLT) highlights the alignment of mobile learning with authentic, communicative tasks (Fang et al., 2021), while approaches such as the Lexical Approach, and Collaborative or Personalized Vocabulary Learning focus on targeted skill development (Al-Abri et al., 2025; Ma & Chiu, 2024). Earlier conceptual contributions adopted by researchers, including the Socio-Educational Model of SLA (Teeter, 2017), the M-Learning Framework (Lai & Zheng, 2018), Mobile-blended collaborative learning (Avci & Adiguzel, 2017), cultures-of-use conceptual framework (Bikowski & Casal, 2018), emerged in the late 2010s as attempts to theorize the growing role of digital technologies in language education; however, these frameworks have received comparatively limited attention in more recent MALL research.

From a broader perspective, the theoretical landscape of MALL research in tertiary EFL education reflects a gradual evolution from technology-focused adoption models toward more holistic frameworks that integrate cognitive, social, and pedagogical dimensions. This convergence suggests that effective MALL implementation is best understood not merely through learners' acceptance of technology, but through the interplay between motivation, interaction, and instructional design. As such, the field is moving toward a more integrated understanding in which theoretical models complement rather than compete, collectively illuminating how mobile technologies mediate both the process and context of language learning.

Research Methods

Analysis of methodological trends reveals both temporal evolution and regional differentiation in MALL research within tertiary EFL education. Temporally, the field has demonstrated a clear trajectory of increasing rigor and sophistication. Early studies, conducted by Wu (2015), Avci and Adiguzel (2017) and Aghajani and Adloo (2018), were largely exploratory, characterized by small samples and simple pre–post designs. Between 2019 and 2020, research designs became more robust, with the inclusion of larger sample sizes, control groups, and explicit theoretical grounding (Hoi, 2020; Yu et al., 2019) More recent investigations (2021–2025) reflect a mature stage of methodological development, with increasing emphasis on sustainability, learner diversity, and contextual implementation. These studies integrated designs with Rasch-based path

model (Hoi & Mu, 2021), structural equation modeling (Shen et al., 2023; Yang & Hu, 2024), and mixed-methods triangulation (Zhang et al., 2025).

Regionally, both the mobile tools adopted and the research methodologies employed vary considerably. In East Asian contexts. MALL research predominantly relied on large-scale surveys and mixed-methods approaches, with frequent use of mobile dictionaries, English learning applications, and social communication tools such as WeChat to support vocabulary acquisition and collaborative learning (Dai & Wu, 2023; Ma & Chiu, 2024; Zhang et al., 2025). In contrast, studies from the Middle East placed greater emphasis on exploring the development of specific language skills and investigating whether widely used social networking platforms such as WhatsApp and Facebook could effectively support learner motivation and skill enhancement, often using experimental or quasi-experimental designs to assess effectiveness (Alamer & Al Khateeb, 2023; Avci & Adiguzel, 2017). Meanwhile, research conducted in Western contexts showed a stronger tendency toward qualitative inquiry and smaller-scale interventions, employing innovative technologies such as mobile feedback systems, and specialized language learning apps to explore learner experiences and individual differences (Andujar & Salaberri-Ramiro, 2021; Bikowski & Casal, 2018; Li & Kim, 2024).

Viewed collectively, these methodological patterns illustrate how MALL research in tertiary EFL education has progressed from exploratory pilot studies to increasingly rigorous and contextually grounded designs. The diversity of methodological choices across regions reflects varying educational priorities and resource environments—East Asian research emphasizes scalability and quantitative validation, while Western studies often prioritize depth and learner-centered qualitative insights. This methodological plurality demonstrates the maturation of the field, highlighting an ongoing shift toward integrative, mixed-methods, and context-sensitive investigations that better capture the complex realities of mobile-assisted language learning.

Discussion

This systematic review of 42 empirical studies provides robust evidence for the promise of MALL in higher education EFL contexts, while also highlighting important contextual dynamics that shape its success. The following section responds to the two research questions and offer insights that extend beyond individual studies, advancing broader understanding of the field.

RQ1: *What are the key themes and trends in empirical studies on MALL in tertiary English education over the past decades?*

From Early Trends to Emerging Directions

This section reflects on the key themes and trends identified in MALL research (RQ1), considering how these developments advance our understanding of mobile-assisted language learning in tertiary English education. The synthesized evidence indicates that MALL research in tertiary EFL contexts has moved from exploratory demonstrations of technical feasibility toward

a more mature, pedagogically oriented literature. Two interrelated developments are particularly salient.

Initially, the temporal trajectory maps clearly onto three phases. In the early phase (2000–2010), studies were largely exploratory, focusing on the technical feasibility of using mobile devices for language practice (Kukulska-Hulme & Shield, 2008). The expansion phase (2011–2020) was characterized by increasing smartphone penetration and the rapid growth of mobile applications, leading to broader investigations of pedagogical affordances and learner acceptance (Sun & Long, 2025). The maturation phase (2021–present) marks a significant consolidation of research, with studies increasingly embedding MALL in institutional settings and exploring advanced technologies such as Augmented Reality (AR) and Virtual Reality (VR) to create immersive and authentic learning experiences (Huang et al., 2021). This temporal pattern extends and updates prior syntheses that documented the field's early optimism and recommended methodological strengthening (Burston, 2015; Sung et al., 2016).

Second, thematic emphases have shifted. Early work concentrated on discrete, technology-centered tasks (e.g., vocabulary drills, SMS-based activities). During the expansion phase, uptake and usability became prominent concerns; adoption studies and investigations of learner attitudes featured heavily. More recent research has re-oriented toward learner-centred questions—motivation, engagement, self-regulation, and collaborative practice—and toward multimodal and context-rich applications (e.g., short-video, AR-supported tasks). This thematic evolution both confirms and refines earlier review conclusions: whereas Burston (2015) and Sung et al. (2016) noted promising but fragmented evidence, the current corpus reveals consolidation around pedagogical value and learner experience, consistent with observations that MALL has moved from novelty to curricular consideration (Dou et al., 2024).

This thematic and temporal evolution underscores both theoretical and practical relevance. The shift toward learner-centred, context-rich applications suggests that MALL is not only a technological innovation but also a catalyst for rethinking instructional design and learner engagement strategies in tertiary EFL settings. The integration of immersive technologies and multimodal tasks highlights opportunities for advancing theoretical models of language acquisition, particularly those that account for social, cognitive, and motivational dimensions in digitally mediated environments.

RQ2: *Which pedagogical strategies, theoretical models, and research methods are most commonly employed in MALL research at the tertiary level?*

Pedagogical, Theoretical, and Methodological Convergence

In response to RQ2, this section examines how MALL research at the tertiary level integrates pedagogical strategies, theoretical frameworks, and methodological approaches, revealing patterns of convergence that characterize current scholarly practice. The findings of this review indicate that pedagogical strategies for integrating mobile technologies into tertiary EFL instruction have matured significantly, moving beyond feasibility studies to more systematic instructional design.

Earlier reviews emphasized the novelty of blended learning and mobile-based tasks but also pointed out that empirical validation was limited and inconsistent (Avci & Adiguzel, 2017). By contrast, recent evidence demonstrates that blended approaches, when supported by teacher training, clear task alignment, and continuous feedback, consistently enhance both learner satisfaction and language outcomes (Huang, 2021). This confirms earlier claims of potential while providing robust empirical support in university contexts (Zhou et al., 2023). Collaborative learning represents another area of development. Prior syntheses highlighted the promise of peer-to-peer collaboration in mobile environments but often noted weak evidence for sustained language gains (Avci & Adiguzel, 2017; Huang, 2021). More recent studies, however, show that well-structured collaborative tasks conducted via social or mobile platforms can foster engagement and improve language performance (Hu et al., 2023; Ma & Chiu, 2024). Similarly, game-based learning has evolved from early optimism about motivation to a more nuanced understanding: while gamified features enhance enjoyment, their effectiveness for language proficiency depends heavily on design quality and alignment with instructional goals (Önal et al., 2022; Taskiran, 2019).

The theoretical framing of MALL has also become more diverse and integrated. Earlier meta-analyses noted that research frequently lacked explicit theoretical grounding or relied narrowly on technology adoption models (Burston & Giannakou, 2022). This review shows that while models such as TAM remain widely used to explain learners' behavioral intentions (Zhang et al., 2025), recent studies extend adoption research through UTAUT and the Theory of Planned Behavior, adding dimensions of social influence and attitudinal factors (Hu et al., 2023; Ren & Tan, 2025). This confirms Pérez-Paredes and Zhang (2022) argument that robust adoption frameworks are necessary but insufficient on their own. Complementary perspectives are now increasingly incorporated. Constructivist and sociocultural theories, particularly Vygotsky's Zone of Proximal Development, situate mobile learning within interactive and contextualized practice (Fang et al., 2021). Compared to earlier research, which often treated mobile devices as delivery tools, current studies frame them as mediational resources for collaborative knowledge construction (Luo & Watts, 2024). This reflects a broader disciplinary shift toward learner-centered design. Likewise, motivational and self-regulation frameworks have gained prominence. Self-Determination Theory (Ryan & Deci, 2000) and Self-Regulated Learning models (Pintrich, 2000) underscore the centrality of autonomy, goal-setting, and metacognitive strategies in sustaining engagement—dimensions largely absent in earlier reviews but now confirmed by empirical work (Alamer & Al Khateeb, 2023; Li & Kim, 2024). Cognitive Load Theory further strengthens the instructional dimension by addressing design efficiency, ensuring that mobile tasks remain cognitively manageable (Yang & Hu, 2024). Together, these frameworks bridge the gap between technology adoption and language learning outcomes. Task-based approaches such as TBLT and lexical or personalized vocabulary learning demonstrate how theory translates into practice, emphasizing authentic, communicative tasks in mobile settings (Fang et al., 2021; Ma & Chiu, 2024). Earlier frameworks, such as the M-Learning Framework (Ozdamli, 2012) and the

Socio-Educational Model of SLA (Gardner et al., 1985) , were important for establishing conceptual ground but are now less visible, suggesting that MALL scholarship increasingly prioritizes empirically validated and context-sensitive models.

Methodological sophistication has increased. The earliest studies commonly used case studies and pre–post designs with small samples; the expansion phase added larger surveys and quasi-experimental comparisons. In the maturation phase, researchers more frequently employ randomized or controlled trials, mixed-methods triangulation, advanced statistical modelling, and learning-analytics approaches. This methodological progression addresses long-standing critiques in the literature about limited generalizability and weak causal inference (Burston et al., 2024), although gaps remain—notably the need for more longitudinal work assessing sustained learning gains and cross-cultural generalizability. Compared with earlier reviews, the present evidence shows both quantitative growth and qualitative consolidation: MALL has evolved from experimental curiosities into a field that interrogates pedagogy, learner experience, and scalable implementation. Nevertheless, the concentration of studies in particular regions and persistent methodological limitations caution against broad generalizations and underscore the need for comparative, multi-context research to test the transferability of current findings.

Overall, the contrast with earlier reviews is clear: where Burston (2015) and Sung et al. (2016) identified theoretical gaps and limited evidence, recent research demonstrates theoretical diversification, pedagogical sophistication, and stronger empirical grounding. This reflects the field's broader maturation, integrating adoption, constructivist, and motivational models with diverse and increasingly sophisticated methodology into a more coherent and multi-dimensional account of MALL in tertiary EFL contexts. These developments highlight that MALL research has progressed from isolated technology adoption to a comprehensive view of how mobile tools shape learning. The integration of pedagogical, theoretical, and methodological sophistication informs instructional design and provides a stronger empirical basis for refining language acquisition theories. This convergence also opens avenues for scalable, context-sensitive implementations in diverse tertiary EFL settings.

Limitations

Although this review followed rigorous procedures aligned with PRISMA recommendations (Page et al., 2021), several constraints must be acknowledged that may influence the interpretation and transferability of the results (Gough et al., 2012).

One major constraint relates to the scope of the search process. Restricting the search to English-language publications may also have created linguistic and cultural bias by omitting significant contributions published in other languages. This is a noteworthy issue given that much mobile learning research originates from non-English speaking regions, as seen in recent work from Zhang et al. (2025) and Al-Abri et al. (2025). Furthermore, the decision to limit the corpus to peer-reviewed journals may have excluded other potentially valuable sources such as dissertations, technical reports, or book chapters. While this restriction ensures quality and

accessibility, it may also reduce exposure to novel or experimental findings often found in emerging research outlets. Restricting the review to studies published between 2015 and 2025 ensured a contemporary focus but may not capture the full trajectory of mobile learning development. Including 2025 publications, although forward-looking, was also challenging because the year had not concluded at the time of data collection, creating the risk of incomplete coverage of the latest research.

Implications for Theory, Practice, and Policy

Theoretical Implications

The findings call for integrated theoretical frameworks that combine insights from technology adoption, learning sciences, and SLA theory (Ryan & Deci, 2000; Venkatesh et al., 2003). While single theories offer explanatory value, they cannot fully account for the multifaceted dynamics of MALL. Additionally, context—cultural, institutional, and individual—emerges as a decisive factor, suggesting that context-responsive frameworks are more appropriate than universal models (Pérez-Paredes & Zhang, 2022).

Practical Implications

For educators and instructional designers, the evidence emphasizes the effectiveness of blended designs that integrate MALL with classroom teaching (Yang & Hu, 2024; Zhang et al., 2025). Careful alignment of mobile activities with curricular goals, along with continuous monitoring and feedback, is essential (Li et al., 2024). Individual learner profiles should guide application selection and instructional design.

Policy Implications

At the policy level, institutions and governments should establish clear guidelines for mobile use in education, invest in robust technical infrastructure, and provide training programs for faculty. Attention to privacy and data protection in mobile contexts is also critical.

Collectively, these implications provide a roadmap for translating empirical evidence into theory-building, pedagogical practice, and informed policymaking, ensuring sustainable and effective integration of mobile technologies in tertiary EFL education.

Conclusion

This systematic review advances previous work by synthesizing empirical studies from 2015–2025 specifically in tertiary EFL contexts, integrating temporal, thematic, pedagogical, theoretical, and methodological dimensions. It highlights learner-centred constructs, emerging technologies, and regional variations that prior reviews had underexplored, providing a more comprehensive and context-sensitive understanding of MALL. This review demonstrates that MALL has become a pivotal component of tertiary EFL education, reshaping the ways learners interact with language, peers, and instructional content. The body of research from 2015 to 2025

highlights that MALL is not simply a technological supplement but a pedagogical innovation capable of promoting learner autonomy, authentic engagement, and access to diverse communicative contexts.

The findings collectively suggest that the effectiveness of MALL depends less on the availability of technology and more on its meaningful pedagogical integration, theoretical grounding, and contextual adaptation. This indicates that sustainable implementation in higher education requires alignment with curriculum goals, institutional support structures, and professional development opportunities for instructors. Without these considerations, the potential of MALL may remain underutilized, reducing its role to a transient trend rather than a long-term educational advancement.

Looking ahead, future research should employ more rigorous and diverse methodologies, including longitudinal and mixed-method designs, to capture the evolving impact of mobile learning on language acquisition. Expanding research beyond dominant geographical contexts will be essential to ensure greater generalizability and inclusivity. Moreover, theoretical development that integrates insights from language acquisition, motivation, and technology acceptance is needed to provide a more comprehensive framework for understanding MALL.

In conclusion, MALL presents both opportunities and challenges. Its capacity to support flexible, engaging, and personalized learning experiences positions it as a valuable approach in tertiary EFL education. However, realizing its full potential requires ongoing empirical investigation, theoretical refinement, and deliberate pedagogical integration, ensuring that mobile technologies contribute meaningfully and sustainably to language learning in diverse higher education contexts.

Funding: This study was not funded.
Acknowledgements: Not Applicable.
Conflict of Interest: The authors declare no conflict of interest.
Authenticity: The authors confirm that this manuscript is original, has not been published previously, and is not under consideration for publication elsewhere. All sources are duly cited and referenced.
Artificial Intelligence Statement: During the drafting process, AI tools such as Grammarly and ChatGPT were used only for language improvement and idea refinement. All intellectual contributions, analysis, and literature synthesis were made by the authors. The final manuscript reflects the authors' academic interpretation and critical evaluation of the reviewed studies.

About the authors:
Li Chengli is a PhD candidate in TESL at the Faculty of Education, Universiti Kebangsaan Malaysia (UKM). Her research focuses on Mobile-Assisted Language Learning (MALL) in higher education. She is currently a lecturer at Sichuan University Jinjiang College, where she has been teaching English for nearly ten years. She has published several articles on English language teaching. https://orcid.org/0009-0008-4999-9106

Associate Professor Dr. Harwati Hashim is a faculty member at the Faculty of Education, Universiti Kebangsaan Malaysia (UKM). She is also a member of the University Research Group on EduXcellence: Development of Innovative Curriculum & Pedagogy. Her research interests include mobile-assisted language learning, instructional design, and innovative pedagogical approaches in ESL education. Dr. Harwati has extensive experience in curriculum development and is committed to enhancing educational practices through technology integration. https://orcid.org/0000-0002-8817-427X

Professor Supyan Hussin is Principal Fellow at the Institute of Ethnic Studies, Universiti Kebangsaan Malaysia (UKM). He specializes in language, education, and technology, with research interests in teacher education, teaching methodology, e-learning, mobile learning, cultural heritage, Malay civilization, ethno-medical science, Malay martial arts, and Big Data development. With over 35 years of teaching experience, he has led numerous research projects and has published seven books and more than 200 scholarly works in books, journals, and proceedings. https://orcid.org/0000-0003-0544-0649

References

Aghajani, M., & Adloo, M. (2018). The effect of online cooperative learning on students' writing skills and attitudes through telegram application. *International Journal of Instruction*, *11*(3), 433-448. https://doi.org/10.12973/iji.2018.11330a

Al-Abri, A., Ranjbaran Madiseh, F., & Morady Moghaddam, M. (2025). Exploring Learning-Oriented Assessment in Enhancing Students' Lexical Fluency Through MALL. *Asia-Pacific Education Researcher*, *34*(1), 1-13, Article 103893. https://doi.org/10.1007/s40299-024-00832-7

Alamer, A., & Al Khateeb, A. (2023). Effects of using the WhatsApp application on language learners motivation: a controlled investigation using structural equation modelling. *Computer Assisted Language Learning*, *36*(1-2), 149-175. https://doi.org/10.1080/09588221.2021.1903042

Alhujaylan, H. (2024). Gauging the effectiveness of a mobile application for learning English phrasal verbs. *ReCALL*. https://doi.org/10.1017/S0958344024000223

Ali, M. M., Alaa, A. M., & Shahnaz, A. (2024). The impact of mobile learning in English language classrooms in Pakistan. *Asian-Pacific Journal of Second and Foreign Language Education*, *9*(1), Article 62. https://doi.org/10.1186/s40862-024-00274-0

Almaiah, M. A., Alfaisal, R., Salloum, S. A., Hajjej, F., Thabit, S., El-Qirem, F. A., Lutfi, A., Alrawad, M., Al Mulhem, A., Alkhdour, T., Awad, A. B., & Al-Maroof, R. S. (2022). Examining the Impact of Artificial Intelligence and Social and Computer Anxiety in E-Learning Settings: Students' Perceptions at the University Level. *Electronics (Switzerland)*, *11*(22), Article 3662. https://doi.org/10.3390/electronics11223662

An, Z., Wang, C., Li, S., Gan, Z., & Li, H. (2021). Technology-Assisted Self-Regulated English Language Learning: Associations With English Language Self-Efficacy, English Enjoyment, and Learning Outcomes. *Frontiers in Psychology*, *11*, Article 558466. https://doi.org/10.3389/fpsyg.2020.558466

Andujar, A., & Salaberri-Ramiro, M. S. (2021). Exploring chat-based communication in the EFL class: computer and mobile environments. *Computer Assisted Language Learning*, *34*(4), 434-461. https://doi.org/10.1080/09588221.2019.1614632

Aşık, A., Sert, O., & Miller, P. (2024). The affordances of a mobile video-tagging tool for evaluating presentation skills in a second language. *Reflective Practice*, *25*(2), 145-163. https://doi.org/10.1080/14623943.2024.2305878

Avci, H., & Adiguzel, T. (2017). A case study on mobile-blended collaborative learning in an english as a foreign language (EFL) context. *International Review of Research in Open and Distributed Learning*, *18*(7), 45-58. https://doi.org/10.19173/irrodl.v18i7.3261

Bai, Y. (2024). A Mixed Methods Investigation of Mobile-Based Language Learning on EFL Students' Listening, Speaking, Foreign Language Enjoyment, and Anxiety. *SAGE Open*, *14*(2). https://doi.org/10.1177/21582440241255554

Bikowski, D., & Casal, J. E. (2018). Interactive digital textbooks and engagement: A learning strategies framework. *Language Learning and Technology*, *22*(1), 119-136. https://doi.org/10.64152/10125/44584

Boroughani, T., Xodabande, I., & Karimpour, S. (2023). Self-regulated learning with mobile devices for university students: exploring the impacts on academic vocabulary development. *Discover Education*, *2*(1), 5. https://doi.org/10.1007/s44217-023-00028-z

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Burston, J. (2015). Twenty years of MALL project implementation: A meta-analysis of learning outcomes. *ReCALL*, *27*(1), 4-20. https://doi.org/10.1017/S0958344014000159

Burston, J., & Giannakou, K. (2022). MALL language learning outcomes: A comprehensive meta-analysis 1994–2019. *ReCALL*, *34*(2), 147-168. https://doi.org/10.1017/S0958344021000240

Çelik, S., & Baran, E. (2022). Student response system: its impact on EFL students' vocabulary achievement. *Technology, Pedagogy and Education*, *31*(2), 141-158. https://doi.org/10.1080/1475939X.2021.1986125

Chan, V. (2024). Using a virtual reality mobile application for interpreting learning: listening to the students' voice. *Interactive Learning Environments*, *32*(6), 2438-2451. https://doi.org/10.1080/10494820.2022.2147958

Chen, Y. M. (2022). Understanding foreign language learners' perceptions of teachers' practice with educational technology with specific reference to Kahoot! and Padlet: A case from China. *Education and Information Technologies*, *27*(2), 1439-1465. https://doi.org/10.1007/s10639-021-10649-2

Dai, Y., & Wu, Z. (2023). Mobile-assisted pronunciation learning with feedback from peers and/or automatic speech recognition: a mixed-methods study. *Computer Assisted Language Learning*, *36*(5-6), 861-884. https://doi.org/10.1080/09588221.2021.1952272

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly: Management Information Systems*, *13*(3), 319-339. https://doi.org/10.2307/249008

Deci, E. L., & Ryan, R. M. (2013). *Intrinsic motivation and self-determination in human behavior*. Springer Science & Business Media.

Dou, K., Halim, H. A., & Saad, M. R. M. (2024). A Bibliometric Study of Mobile-Assisted Language Learning from 2013 to 2023: Research Themes and Trends. *Journal of Curriculum and Teaching*. https://doi.org/10.5430/jct.v13n5p431

Dou, K., Halim, H. A., & Saad, M. R. M. (2025). Factors Determining the Continuance Intention to Use Mobile English Learning Applications: An Application and Extension of Expectation Confirmation Model in the Mobile-Assisted Language Learning Context. *European Journal of Education*, *60*(2), Article e70082. https://doi.org/10.1111/ejed.70082

Fang, W. C., Yeh, H. C., Luo, B. R., & Chen, N. S. (2021). Effects of mobile-supported task-based language teaching on EFL students' linguistic achievement and conversational interaction. *ReCALL*, *33*(1), 71-87. https://doi.org/10.1017/S0958344020000208

García Botero, G., Questier, F., Cincinnato, S., He, T., & Zhu, C. (2018). Acceptance and usage of mobile assisted language learning by higher education students. *Journal of Computing in Higher Education*, *30*(3), 426-451. https://doi.org/10.1007/s12528-018-9177-1

Gardner, R. C., Lalonde, R. N., & Moorcroft, R. (1985). THE ROLE OF ATTITUDES AND MOTIVATION IN SECOND LANGUAGE LEARNING: CORRELATIONAL AND EXPERIMENTAL CONSIDERATIONS. *Language Learning*, *35*(2), 207-227. https://doi.org/10.1111/j.1467-1770.1985.tb01025.x

Gligorea, I., Cioca, M., Oancea, R., Gorski, A.-T., Gorski, H., & Tudorache, P. (2023). Adaptive Learning Using Artificial Intelligence in e-Learning: A Literature Review [Review]. *Education Sciences*, *13*(12), Article 1216. https://doi.org/10.3390/educsci13121216

Godwin-Jones, R. (2017a). Emerging technologies: Scaling up and zooming in: Big data and personalization in language learning. *Language Learning and Technology*, *21*(1), 4-15. https://doi.org/llt.msu.edu/issues/february2017/emerging.pd

Godwin-Jones, R. (2017b). Smartphones and language learning. *Language Learning and Technology*, *21*(2), 3-17. https://doi.org/llt.msu.edu/issues/june2017/emerging.pdf

Gough, D., Thomas, J., & Oliver, S. (2012). *An introduction to systematic reviews*. SAGE.

Guo, P., Jeyaraj, J. J., & Razali, A. B. (2024). A systematic review of collaborative mobile-assisted language learning (C-MALL) practices using bibliometric, content, and scientometric analyses. *Humanities and Social Sciences Communications*, *11*(1), Article 1435. https://doi.org/10.1057/s41599-024-03940-3

Habbash, M. (2020). Computer assisted language learning integration challenges in Saudi Arabian English as a foreign language classes: The case of EFL teachers at Tabuk University. *Arab World English Journal (AWEJ) 11*(4), 474-489. https://doi.org/10.24093/awej/vol11no4.30

Hasumi, T., & Chiu, M. S. (2024). Technology-enhanced language learning in English language education: Performance analysis, core publications, and emerging trends. *Cogent Education*, *11*(1), Article 2346044. https://doi.org/10.1080/2331186X.2024.2346044

Hema, S., & Gandhimathi, S. N. S. (2025). Unravelling the tapestry of M-learning in English language teaching: a bibliometric analysis. *Cogent Education*, *12*(1), Article 2521162. https://doi.org/10.1080/2331186X.2025.2521162

Hoi, V. N. (2020). Understanding higher education learners' acceptance and use of mobile devices for language learning: A Rasch-based path modeling approach. *Computers and Education*, *146*, Article 103761. https://doi.org/10.1016/j.compedu.2019.103761

Hoi, V. N., & Mu, G. M. (2021). Perceived teacher support and students' acceptance of mobile-assisted language learning: Evidence from Vietnamese higher education context. *British Journal of Educational Technology*, *52*(2), 879-898. https://doi.org/10.1111/bjet.13044

Hossain, K. I. (2024). Reviewing the role of culture in English language learning: Challenges and opportunities for educators. *Social Sciences and Humanities Open*, *9*, Article 100781. https://doi.org/10.1016/j.ssaho.2023.100781

Hu, L., Hei, D., Wang, H., & Dai, X. (2023). Chinese college students collaborative mobile-assisted language learning experience and flow as a key factor for further adoption. *Frontiers in Psychology*, *14*, Article 1165332. https://doi.org/10.3389/fpsyg.2023.1165332

Huang, H. W. (2021). Effects of smartphone-based collaborative vlog projects on EFL learners' speaking performance and learning engagement. *Australasian Journal of Educational Technology*, *37*(6), 18-40. https://doi.org/10.14742/ajet.6623

Huang, X. Y., Zou, D., Cheng, G., & Xie, H. R. (2021). A Systematic Review of AR and VR Enhanced Language Learning. *Sustainability 13*(9), 4639. https://doi.org/10.3390/su13094639

Hwang, G. J., & Fu, Q. K. (2019). Trends in the research design and application of mobile language learning: a review of 2007–2016 publications in selected SSCI journals. *Interactive Learning Environments*, *27*(4), 567-581. https://doi.org/10.1080/10494820.2018.1486861

Jeong, K. O. (2023). Integrating Technology into Language Teaching Practice in the Post-COVID-19 Pandemic Digital Age: From a Korean English as a Foreign Language Context. *RELC Journal*, *54*(2), 394-409. https://doi.org/10.1177/00336882231186431

Jiang, W., & Li, W. (2018). Linking up learners of Chinese with native speakers through WeChat in an Australian tertiary CFL curriculum [Article]. *Asian-Pacific Journal of Second and Foreign Language Education*, *3*(1), Article 14. https://doi.org/10.1186/s40862-018-0056-0

Kacetl, J., & Klímová, B. (2019). Use of smartphone applications in english language learning—A challenge for foreign language education. *Education Sciences*, *9*(3), Article 179. https://doi.org/10.3390/educsci9030179

Khezrlou, S., Stockwell, G., & (2025). A Synthetic Review of MALL Research: Artifact, Environment, and Task. *Digital Studies in Language and Literature*,

. https://doi.org/10.1515/dsll-2024-0026

Khosrawi-Rad, B., Keller, P. F., Benner, D., Grogorick, L., Borchers, A., Janson, A., Leimeister, J. M., & Robra-Bissantz, S. (2025). Promoting students' motivation in language education with gamified pedagogical conversational agents. *Computers and Education*, *238*, Article 105374. https://doi.org/10.1016/j.compedu.2025.105374

Kukulska-Hulme, A. (2009). Will mobile learning change language learning? [; Proceedings Paper]. *ReCALL*, *21*, 157-165. https://doi.org/10.1017/s0958344009000202

Kukulska-Hulme, A. (2020). Mobile and personal learning for newcomers to a city. *Electronic Journal of Foreign Language Teaching*, *17*, 93-103. https://doi.org/10.56040/agkh1717

Kukulska-Hulme, A., & Shield, L. (2008). An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL*, *20*(3), 271-289. https://doi.org/10.1017/S0958344008000335

Lai, C., & Zheng, D. (2018). Self-directed use of mobile devices for language learning beyond the classroom. *ReCALL*, *30*(3), 299-318. https://doi.org/10.1017/S0958344017000258

Lai, Y., Saab, N., & Admiraal, W. (2022). University students' use of mobile technology in self-directed language learning: Using the integrative model of behavior prediction. *Computers and Education*, *179*, Article 104413. https://doi.org/10.1016/j.compedu.2021.104413

Lazaro, G. R. D., & Duart, J. M. (2023). Moving Learning: A Systematic Review of Mobile Learning Applications for Online Higher Education. *Journal of New Approaches in Educational Research*, *12*(2), 198-224. https://doi.org/10.7821/naer.2023.7.1287

Lee, B. J. (2021). Comparing factual recall of tapped vs. handwritten text. *Acta Psychologica*, *212*, Article 103221. https://doi.org/10.1016/j.actpsy.2020.103221

Li, F., Fan, S., & Wang, Y. (2022). Mobile-assisted language learning in Chinese higher education context: a systematic review from the perspective of the situated learning theory. *Education and Information Technologies*, *27*(7), 9665-9688. https://doi.org/10.1007/s10639-022-11025-4

Li, L., & Kim, M. (2024). It is like a friend to me: Critical usage of automated feedback systems by self-regulating English learners in higher education. *Australasian Journal of Educational Technology*, *40*(1), 1-18. https://doi.org/10.14742/ajet.8821

Li, R., Zou, D., Reynolds, B. L., & Vazquez-Calvo, B. (2023). Editorial: Mobile assisted language learning: developments, affordances, and solutions. *Frontiers in Psychology*, *14*, Article 1293483. https://doi.org/10.3389/fpsyg.2023.1293483

Li, X., Wang, K., Wang, Y., He, J., & Zhang, J. (2024). Assessing User Experiences in Virtual Reality Learning Environments. *Asia-Pacific Education Researcher*, *33*(5), 1149-1160. https://doi.org/10.1007/s40299-023-00785-3

Luo, Y., & Watts, M. (2024). Exploration of university students' lived experiences of using smartphones for English language learning. *Computer Assisted Language Learning*, *37*(4), 608-633. https://doi.org/10.1080/09588221.2022.2052904

Ma, M., Noordin, N., & Razali, A. B. (2025). Improving EFL speaking performance among undergraduate students with an AI-powered mobile app in after-class assignments: an empirical investigation. *Humanities and Social Sciences Communications*, *12*(1), Article 370. https://doi.org/10.1057/s41599-025-04688-0

Ma, Q. (2017). A multi-case study of university students' language-learning experience mediated by mobile technologies: a socio-cultural perspective. *Computer Assisted Language Learning*, *30*(3-4), 183-203. https://doi.org/10.1080/09588221.2017.1301957

Ma, Q., & Chiu, M. M. (2024). Self-regulated and collaborative personalised vocabulary learning approach in MALL. *Language Learning and Technology*, *28*(1), 1-28. https://doi.org/hdl.handle.net/10125/73579

Metruk, R. (2024). Mobile-assisted language learning and pronunciation instruction: A systematic literature review. *Education and Information Technologies*, *29*(13), 16255-16282. https://doi.org/10.1007/s10639-024-12453-0

Mihaylova, M., Gorin, S., Reber, T. P., & Rothen, N. (2022). A Meta-Analysis on Mobile-Assisted Language Learning Applications: Benefits and Risks. *Psychologica Belgica*, *62*(1), 252-271. https://doi.org/10.5334/pb.1146

Nami, F. (2020). Educational smartphone apps for language learning in higher education: Students' choices and perceptions. *Australasian Journal of Educational Technology*, *36*(4), 82-95. https://doi.org/10.14742/ajet.5350

Naveed, Q. N., Choudhary, H., Ahmad, N., Alqahtani, J., & Qahmash, A. I. (2023). Mobile Learning in Higher Education: A Systematic Literature Review. *Sustainability*, *15*(18), 13566. https://doi.org/10.3390/su151813566

Nawaz, M., Ganapathy, M., Manzoor, S., Ezzy, T., & Farooqi, S. U. H. (2025). A SYSTEMATIC REVIEW OF MOBILE-ASSISTED VOCABULARY LEARNING AND TEACHING IN ESL/EFL CONTEXT. *Malaysian Journal of Learning and Instruction*, *22*(1), 40-58. https://doi.org/10.32890/mjli2025.22.1.3

Oga-Baldwin, Quint, W., Parrish, A., & Noels, K. (2023). Taking root: Self-determination in language education. *Journal for the Psychology of Language Learning 4*(1). https://doi.org/10.52598/jpll/4/1/9

Okumuş Dağdeler, K. (2023). A systematic review of Mobile-Assisted Vocabulary Learning research. *Smart Learning Environments*, *10*(1), Article 19. https://doi.org/10.1186/s40561-023-00235-z

Önal, N., Çevik, K. K., & Şenol, V. (2022). The effect of SOS Table learning environment on mobile learning tools acceptance, motivation and mobile learning attitude in English language learning. *Interactive Learning Environments*, *30*(5), 834-847. https://doi.org/10.1080/10494820.2019.1690529

Ozdamli, F. (2012). Pedagogical framework of m-learning. *Procedia - Social and Behavioral Sciences*, *31*, 927-931. https://doi.org/10.1016/j.sbspro.2011.12.171

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372. https://doi.org/eresourcesptsl.ukm.remotexs.co/10.1136/bmj.n71

Pan, M., Lai, C., & Guo, K. (2025). Technology-enhanced self-regulated second language learning interventions: A systematic review. *System*, *131*, 103675. https://doi.org/eresourcesptsl.ukm.remotexs.co/10.1016/j.system.2025.103675

Pegrum, M. (2014). *Mobile learning: Languages, literacies and cultures*. Springer. https://doi.org/10.1057/9781137309815

Peng, C., Rashid, S. M., & Jiang, L. (2024). The Use of Metacognitive Strategies within a Mobile-Assisted Language Learning Context: Enhancing Chinese EFL Learner's Listening Skills. *Forum for Linguistic Studies*, *6*(6), 410-428. https://doi.org/10.30564/fls.v6i6.7439

Pérez-Paredes, P., & Zhang, D. (2022). Mobile assisted language learning: scope, praxis and theory. *Porta Linguarum An International Journal of Foreign Language Teaching and Learning*, 11-25. https://doi.org/10.30827/portalin.vi.21424

Petticrew, M., & Roberts, H. (2005). *Systematic reviews in the social sciences: A practical guide* (1st ed.). Wiley-Blackwell.

Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures* (A. Rosin, Trans.). Viking Press.

Pintrich, P. R. (2000). The Role of Goal Orientation in Self-Regulated Learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 451-502). Academic Press. https://doi.org/eresourcesptsl.ukm.remotexs.co/10.1016/B978-012109890-2/50043-3

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006). *Guidance on the conduct of narrative synthesis in systematic reviews: A product from the ESRC Methods Programme*.

Pranckutė, R. (2021). Web of Science (WoS) and Scopus: The titans of bibliographic information in today's academic world. *Publications*, *9*(1), 12. https://doi.org/DOI:10.3390/publications9010012

Rashid, S., Howard, J., Cunningham, U., & Watson, K. (2021). Learner training in MALL: a Pakistani case study. *Innovation in Language Learning and Teaching*, *15*(2), 181-194. https://doi.org/10.1080/17501229.2020.1737076

Ren, Y., & Tan, X. (2025). The influence mechanism of planned behavior and self-consciousness on learners' willingness to use language learning Apps: Validation based on structural equation modeling. *PLoS ONE*, *20*(3 March), Article e0319300. https://doi.org/10.1371/journal.pone.0319300

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68-78. https://doi.org/10.1037/0003-066X.55.1.68

Sharples, M., Taylor, J., & Vavoula, G. (2005). Towards a theory of mobile learning.

Shavkatovna, D. K., Hashim, H., & Jamaludin, K. A. (2024). Mobile Learning for Transforming English Writing Skills: A Systematic Review (2013–2023). *International Journal of Interactive Mobile Technologies*, *18*(19), 4-17. https://doi.org/10.3991/ijim.v18i19.47729

Shen, F., Liang, L., & Feng, Y. (2023). A study on factors influencing Chinese undergraduate EFL learners' self-directed use of mobile English learning resources. *Frontiers in Psychology*, *14*, Article 1189055. https://doi.org/10.3389/fpsyg.2023.1189055

Šramová, B. (2024). University students' experience with mobile learning during COVID-19 pandemic. *Interactive Learning Environments*, *32*(9), 5537-5551. https://doi.org/10.1080/10494820.2023.2220362

Stockwell, G. (2010). Using mobile phones for vocabulary activities: Examining the effect of the platform. *Language Learning and Technology*, *14*(2), 95-110. https://www.scopus.com/inward/record.uri?eid=2-s2.0-77956413909&partnerID=40&md5=171a7aef8028997889eafe8ff32af5ec

Stockwell, G. (2013). Technology and motivation in English-language teaching and learning. In E. Ushioda (Ed.), *International perspectives on motivation: Language learning and professional challenges* (pp. 156-175). Palgrave Macmillan UK. https://doi.org/10.1057/9781137000873

Sun, C., & Long, R. (2025). Advances and challenges in online language learning: a bibliometric analysis. *Innovation in Language Learning and Teaching*, 1-23. https://doi.org/10.1080/17501229.2025.2457552

Sung, Y. T., Chang, K. E., & Liu, T. C. (2016). The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis and research synthesis. *Computers and Education*, *94*, 252-275. https://doi.org/10.1016/j.compedu.2015.11.008

Taskiran, A. (2019). The effect of augmented reality games on English as foreign language motivation. *E-Learning and Digital Media*, *16*(2), 122-135. https://doi.org/10.1177/2042753018817541

Teeter, J. L. (2017). Improving motivation to learn english in japan with a self-study shadowing application. *Languages*, *2*(4), Article 19. https://doi.org/10.3390/languages2040019

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, *8*(1), 45. https://doi.org/10.1186/1471-2288-8-45

Torgerson, C. (2003). *Systematic reviews* Continuum.

Uz Bilgin, C., & Tokel, S. T. (2019). Facilitating Contextual Vocabulary Learning in a Mobile-Supported Situated Learning Environment. *Journal of Educational Computing Research*, *57*(4), 930-953. https://doi.org/10.1177/0735633118779397

Venkatesh, V., & Davis, F. D. (2000). Theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science*, *46*(2), 186-204. https://doi.org/10.1287/mnsc.46.2.186.11926

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly: Management Information Systems*, *27*(3), 425-478. https://doi.org/10.2307/30036540

Vygotsky, L. S., & Cole, M. . (1978). *Mind in society : the development of higher psychological processes*. Harvard University Press.

Wang, F., Geng, X., & Han, J. (2024). Chinese university EFL learners' English for General Academic Purposes: relationships between target needs and self-efficacy. *Humanities and Social Sciences Communications*, *11*(1), Article 215. https://doi.org/10.1057/s41599-024-02735-w

Wang, Q., Zhao, G., & Cheng, Z. (2022). Examining the moderating role of technostress and compatibility in EFL Learners' mobile learning adoption: A perspective from the theory of planned behaviour. *Frontiers in Psychology*, *13*, Article 919971. https://doi.org/10.3389/fpsyg.2022.919971

Warschauer, M., & Xu, Y. (2024). Artificial intelligence for language learning: Entering a new era [Editorial Material]. *Language Learning & Technology*, *28*(2), 20-20. <Go to ISI>://WOS:001249118200001

Wu, Q. (2015). Pulling mobile assisted language learning (MALL) into the mainstream: MALL in broad practice. *PLoS ONE*, *10*(5), Article e0128762. https://doi.org/10.1371/journal.pone.0128762

Xiao, L., J., F., S., N., & M., R. (2022). The Impact of Mobile-Assisted Language Learning on English as a Foreign Language Learners' Vocabulary Learning Attitudes and Self-Regulatory Capacity. *Front Psychol*. https://doi.org/10.3389/fpsyg.2022.872922

Yang, X., & Hu, J. (2023). Chinese college students' preferences for mobile-assisted language learning classes and their effects on student engagement in English language classrooms. *Innovation in Language Learning and Teaching*, *17*(5), 953-966. https://doi.org/10.1080/17501229.2023.2207564

Yang, X., & Hu, J. (2024). Distinctions between mobile-assisted and paper-based EFL reading comprehension performance: reading cognitive load as a mediator. *Computer Assisted Language Learning*, *37*(7), 2051-2082. https://doi.org/10.1080/09588221.2022.2143527

Yu, Z. G., Ying, Z., Zhichun, Y., & Wentao, C. (2019). Student satisfaction, learning outcomes, and cognitive loads with a mobile learning platform. *Computer Assisted Language Learning*, *32*(4), 323-341. https://doi.org/10.1080/09588221.2018.1517093

Zakaria, N. Y. K., Hashim, H., & Jamaludin, K. A. (2025). Mobile Technology for Enhanced ESL Writing Education: A Systematic Review of Benefits and Pedagogical Innovations [Article]. *Arab World English Journal*, *2025*(Special Issue), 210-222. https://doi.org/10.24093/awej/call11.13

Zakian, M., Xodabande, I., Valizadeh, M., & Yousefvand, M. (2022). Out-of-the-classroom learning of English vocabulary by EFL learners: investigating the effectiveness of mobile assisted learning with digital flashcards. *Asian-Pacific Journal of Second and Foreign Language Education*, *7*(1), Article 16. https://doi.org/10.1186/s40862-022-00143-8

Zhang, D., Hennessy, S., & Pérez-Paredes, P. (2025). An investigation of Chinese EFL learners' acceptance of mobile dictionaries in English language learning. *Computer Assisted Language Learning*, *38*(3), 317-341. https://doi.org/10.1080/09588221.2023.2189915

Zhang, D., & Pérez-Paredes, P. (2024). Chinese EFL learners' use of mobile dictionaries in reading comprehension tasks. *System, 121*, Article 103221. https://doi.org/10.1016/j.system.2024.103221

Zhou, Q., Hashim, H., & Sulaiman, N. A. (2023). A Systemic Review (2014–2023) on the Mobile-Assisted Blended Learning for English as a Foreign Language Education with a Focus on Empirical Studies. *International Journal of Interactive Mobile Technologies, 17*(24), 153-172. https://doi.org/10.3991/IJIM.V17I24.42019

Zimmerman, B. J. (2002). Becoming a Self-Regulated Learner: An Overview. *Theory Into Practice, 41*(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2

Zou, C., Li, P., & Jin, L. (2021). Online college English education in Wuhan against the COVID-19 pandemic: Student and teacher readiness, challenges and implications. *PLoS ONE, 16*(10 October), Article e0258137. https://doi.org/10.1371/journal.pone.0258137

Appendices

Appendix A

Coding Scheme

Code	Country/Region	Theme	Theoretical Framework/Pedagogical Strateg	Resear	Mobile Technology/Tool	Analytical Tools	Findings	Challenges	New format or approach
2015-1	China	vocabulary	no	quanti	designed		- The smartpho	- The study highlights the lack of user c	
2017-1	Turkey	communication skills,vocabulary	collaborative	mix	whatsapp		- The study fou	- Students faced difficulties in agreeing	
2017-2	Hongkong	personalized learning	socio-cultural framework	quali	many		- The study rev	- The paper highlights pedagogical limit	
2017-3	Japan	listening speaking	motivation	quanti	shadowing		- The study hig	- The scarcity of opportunities to commu	
2018-1	Iran	writing	cooperative learning	quanti	telegram		- The study inv	- Students relied too much on online cor	
2018-2	USA	reading	cultures-of-use conceptual framework	mix	designed interactive digital textbook		- The study fou	- Many students find digital textbooks c	
2019-1	Turkey	AR	game-based learning	quanti	desinged AR game		- EFL learners l	Not addressed in the paper.	
2019-2	Turkey	contextual vocabulary learning	situated learning theory	mix	designed app		- The study fou	- Novice learners struggle with transfer a	
2019-3	China	proficiency, satisfaction,cognitive loads	cognitive load	mix	designed a mobile learning platform		- The study val	- The study faced challenges due to a la	
2020-1	Vietnam	learners' attitude, behavioral intention	UTAUT	quanti	many	Rasch-based path model	- The study sup	- The study's cross-sectional nature limi	
2020-2	Iran	choices and perceptions	mobile learning	quanti	many	ANOVA	- The study fou	- Challenges in mobile learning include s	
2021-1	Spain	cognitive engagement,emotional engagement	chat-based communication. CMC MMC	mix	whatsapp,facebook		- The study fou	- The lack of curriculumintegration pose	
2021-2	Taiwan	vocabulary,grammar,oral communication	TBLT,ZPD	mix	designed app for IG activities		- Mobile-suppo	- Large classroomsizes limit learners' act	
2021-3	Vietnam	teacher support, students' acceptance	TAM	quanti	many	Rasch-based path model	- The study fou	- Technical limitations include small scre	
2021-4	China	speaking, engagement	Social constructivism,Collaborative learning	mix	smartphone		- The study fou	- Students found video editing to be tim	
2021-5	Japan	writing,cognitive	memory and notetaking	quanti	Google Doc		- Handwritten n	- Controlling independent variables, suc	
2021-6	Pakistan	writing blogpost	learner-training framework	quali	designed MALL-based programme		- The study fou	- Technical difficulties hindered student	
2022-1	Turkey	vocabulary	TPACK, Student response systems	mix	SRS app – Socrative		- The study fou	- SRS technology is rarely utilized in lang	
2022-2	China	learners' perceptions	mlearning	quali	Kahoot,Padlet		- The study con	- Internet connectivity issues were a sig	
2022-3	China	vocabulary,translation,students' intention,Self regulation Self-directed	integrative model of behavior prediction	quanti	many	SEM	- The study fou	- Vietnamese and Japanese learners rarel	
2022-4	Turkey	tense,acceptance, motivation	no	mix	designed SOS Table mobile game		The SOS Tabl	Participants reported negative features	
2022-5	Turkey	listening speaking	mlearning	quanti	smartphone		- Participants s	- The research sample is limited to two u	
2022-6	Chian	adoption intention	theory of planned behaviour (TPB)	quanti	many	Mplus	- The study fou	- The study's construct of compatibility	
2022-7	Iran	vocabulary	no	quanti	flashcard app	SPSS,ANOVA	- The study fou	- The motivational impact of mobile devi	
2023-1	Saudi Arabia	motivation	self-determination theory (SDT)	quanti	whatsapp	ANOVA,SEM	- The study fou	- WhatsApp requires well-thought imple	
2023-2	China	pronunciation,collaboration, peer feedback	skill acquisition theory, sociocultural theory	mix	wechat,automatic speech recognition		All three grou	Students expressed concerns about pe	
2023-3	China	Collaborative MALL	UTAUT and flow theory	quanti	many	AMOS 24.0,SPSS 26.0	- The study rev	- High perceived costs can reduce users	
2023-4	China	behavioral intention to use MELR	UTAUT	quanti	many	SEM	- The study ide	- The study had a limited sample size of	
2023-5	China	engagement,preference	no	mix	Xuexitong		- The study fou	- The study's data collection at the seme	
2024-1	Saudi Arabia	phrasal verbs,vocabulary	constructive learning theory	quanti	PHRASAL VERBS app	ANOVA	- The experimer	- EFL learners struggle with the complex	
2024-2	Pakistan	grammar	Connectivism	quanti	whatsapp and many other	SPSS 23.0	- The study fou	Not addressed in the paper.	
2024-3	Turkey	presentation skills	no	quali	Video-Enhanced Observation		- The study con	- The paper highlights the issue of distra	
2024-4	Hongkong	Interpreting	situated learning	quali	desinged Virtual Interpreting Practice,	VR	- The study fou	- Participants experienced anxiety due to	
2024-5	Australia	engagement	self-regulated learning (SRL)	mix	automated feedback systems (AFSs)		- Participants re	- The study identified constraints relate	
2024-6	China	vocabulary	no	quanti	VR,computer, mobile	G*Power 3.1.9.4 software,	- The study hig	- Participants noted limitations in the val	
2024-7	China	self-regulation, motivation	Sociocultural theory (SCT)	quali	many		- The study rev	- Limited sac smartphone-assisted Englis	
2024-8	Hongkong	vocabulary,self-regulation, collaborative personalised lea	self-regulation, collaborative learning, person	mix	Schoology		- The study pro	- The study Self-regulated, Collaborativ	
2024-9	China	reading,cognitive load	cognitive load framework	mix	waiyanyuedu	ANOVA,SEM	- The study fou	- VR's software and hardware systems w	
2024-10	China	dictionary, reading	no	mix	three mobile dictionaries	SPSS (version 28.0)	- The study fou	- Participants perceived greater cognitiv	
2025-1	Oman	vocabulary retention, lexical fluency, speaking performanc	Digital Learning Ecology (DLE) framework	mix	Mentimeter, a Student Response System		- The integratio	- The quizzes designed through Mentim	
2025-2	China	learners' willingness	TAM-TPR framework	quanti	many	SEM,AMOS 23.0	- The study val	Not addressed extended TAM-TPR framew	
2025-3	China	dictionary,reading, vocabulary	TAM	mix	three mobile dictionaries	MANOVA	- The study fou	- Participants reported excessive definiti	

Appendix B

List of Included Studies

Code (Ref. ID)	Full Reference (APA 7th)
01	Wu, Q. (2015). Pulling mobile assisted language learning (MALL) into the mainstream: MALL in broad practice. *PLoS ONE*, *10*(5), Article e0128762. https://doi.org/10.1371/journal.pone.0128762
02	Avci, H., & Adiguzel, T. (2017). A case study on mobile-blended collaborative learning in an english as a foreign language (EFL) context. *International Review of Research in Open and Distributed Learning, 18*(7), 45-58. https://doi.org/10.19173/irrodl.v18i7.3261
03	Ma, Q. (2017). A multi-case study of university students' language-learning experience mediated by mobile technologies: a socio-cultural perspective. *Computer Assisted Language Learning*, *30*(3-4), 183-203. https://doi.org/10.1080/09588221.2017.1301957
04	Teeter, J. L. (2017). Improving motivation to learn English in Japan with a self-study shadowing application. *Languages*, *2*(4), Article 19. https://doi.org/10.3390/languages2040019

05	Aghajani, M., & Adloo, M. (2018). The effect of online cooperative learning on students' writing skills and attitudes through telegram application. *International Journal of Instruction*, *11*(3), 433-448. https://doi.org/10.12973/iji.2018.11330a
06	Bikowski, D., & Casal, J. E. (2018). Interactive digital textbooks and engagement: A learning strategies framework. *Language Learning and Technology*, *22*(1), 119-136. https://doi.org/10.64152/10125/44584
07	Taskiran, A. (2019). The effect of augmented reality games on English as foreign language motivation. *E-Learning and Digital Media*, *16*(2), 122-135. https://doi.org/10.1177/2042753018817541
08	Uz Bilgin, C., & Tokel, S. T. (2019). Facilitating Contextual Vocabulary Learning in a Mobile-Supported Situated Learning Environment. *Journal of Educational Computing Research*, *57*(4), 930-953. https://doi.org/10.1177/0735633118779397
09	Yu, Z. G., Ying, Z., Zhichun, Y., & Wentao, C. (2019). Student satisfaction, learning outcomes, and cognitive loads with a mobile learning platform. *Computer Assisted Language Learning*, *32*(4), 323-341. https://doi.org/10.1080/09588221.2018.1517093
10	Hoi, V. N. (2020). Understanding higher education learners' acceptance and use of mobile devices for language learning: A Rasch-based path modeling approach. *Computers and Education*, *146*, Article 103761. https://doi.org/10.1016/j.compedu.2019.103761
11	Nami, F. (2020). Educational smartphone apps for language learning in higher education: Students' choices and perceptions. *Australasian Journal of Educational Technology*, *36*(4), 82-95. https://doi.org/10.14742/ajet.5350
12	Andujar, A., & Salaberri-Ramiro, M. S. (2021). Exploring chat-based communication in the EFL class: computer and mobile environments. *Computer Assisted Language Learning*, *34*(4), 434-461. https://doi.org/10.1080/09588221.2019.1614632
13	Fang, W. C., Yeh, H. C., Luo, B. R., & Chen, N. S. (2021). Effects of mobile-supported task-based language teaching on EFL students' linguistic achievement and conversational interaction. *ReCALL*, *33*(1), 71-87. https://doi.org/10.1017/S0958344020000208
14	Hoi, V. N., & Mu, G. M. (2021). Perceived teacher support and students' acceptance of mobile-assisted language learning: Evidence from Vietnamese higher education context. *British Journal of Educational Technology*, *52*(2), 879-898. https://doi.org/10.1111/bjet.13044
15	Huang, H. W. (2021). Effects of smartphone-based collaborative vlog projects on EFL learners' speaking performance and learning engagement. *Australasian Journal of Educational Technology*, *37*(6), 18-40. https://doi.org/10.14742/ajet.6623
16	Lee, B. J. (2021). Comparing factual recall of tapped vs. handwritten text. *Acta Psychologica*, *212*, Article 103221. https://doi.org/10.1016/j.actpsy.2020.103221
17	Rashid, S., Howard, J., Cunningham, U., & Watson, K. (2021). Learner training in MALL: a Pakistani case study. *Innovation in Language Learning and Teaching*, *15*(2), 181-194. https://doi.org/10.1080/17501229.2020.1737076
18	Çelik, S., & Baran, E. (2022). Student response system: its impact on EFL students' vocabulary achievement. *Technology, Pedagogy and Education*, *31*(2), 141-158. https://doi.org/10.1080/1475939X.2021.1986125
19	Chen, Y. M. (2022). Understanding foreign language learners' perceptions of teachers' practice with educational technology with specific reference to Kahoot! and Padlet: A case from China. *Education and Information Technologies*, *27*(2), 1439-1465. https://doi.org/10.1007/s10639-021-10649-2
20	Lai, Y., Saab, N., & Admiraal, W. (2022). University students' use of mobile technology in self-directed language learning: Using the integrative model of behavior

	prediction. *Computers and Education*, *179*, Article 104413. https://doi.org/10.1016/j.compedu.2021.104413
21	Önal, N., Çevik, K. K., & Şenol, V. (2022). The effect of SOS Table learning environment on mobile learning tools acceptance, motivation and mobile learning attitude in English language learning. *Interactive Learning Environments*, *30*(5), 834-847. https://doi.org/10.1080/10494820.2019.1690529
22	Şad, S. N., Özer, N., Yakar, Ü., & Öztürk, F. (2022). Mobile or hostile? Using smartphones in learning English as a foreign language. *Computer Assisted Language Learning*, *35*(5-6), 1031-1057. https://doi.org/10.1080/09588221.2020.1770292
23	Wang, Q., Zhao, G., & Cheng, Z. (2022). Examining the moderating role of technostress and compatibility in EFL Learners' mobile learning adoption: A perspective from the theory of planned behaviour. *Frontiers in Psychology*, *13*, Article 919971. https://doi.org/10.3389/fpsyg.2022.919971
24	Zakian, M., Xodabande, I., Valizadeh, M., & Yousefvand, M. (2022). Out-of-the-classroom learning of English vocabulary by EFL learners: investigating the effectiveness of mobile assisted learning with digital flashcards. *Asian-Pacific Journal of Second and Foreign Language Education*, *7*(1), Article 16. https://doi.org/10.1186/s40862-022-00143-8
25	Alamer, A., & Al Khateeb, A. (2023). Effects of using the WhatsApp application on language learners motivation: a controlled investigation using structural equation modelling. *Computer Assisted Language Learning*, *36*(1-2), 149-175. https://doi.org/10.1080/09588221.2021.1903042
26	Dai, Y., & Wu, Z. (2023). Mobile-assisted pronunciation learning with feedback from peers and/or automatic speech recognition: a mixed-methods study. *Computer Assisted Language Learning*, *36*(5-6), 861-884. https://doi.org/10.1080/09588221.2021.1952272
27	Hu, L., Hei, D., Wang, H., & Dai, X. (2023). Chinese college students collaborative mobile-assisted language learning experience and flow as a key factor for further adoption. *Frontiers in Psychology*, *14*, Article 1165332. https://doi.org/10.3389/fpsyg.2023.1165332
28	Shen, F., Liang, L., & Feng, Y. (2023). A study on factors influencing Chinese undergraduate EFL learners' self-directed use of mobile English learning resources. *Frontiers in Psychology*, *14*, Article 1189055. https://doi.org/10.3389/fpsyg.2023.1189055
29	Yang, X., & Hu, J. (2023). Chinese college students' preferences for mobile-assisted language learning classes and their effects on student engagement in English language classrooms. *Innovation in Language Learning and Teaching*, *17*(5), 953-966. https://doi.org/10.1080/17501229.2023.2207564
30	Alhujaylan, H. (2024). Gauging the effectiveness of a mobile application for learning English phrasal verbs. *ReCALL*. https://doi.org/10.1017/S0958344024000223
31	Ali, M. M., Alaa, A. M., & Shahnaz, A. (2024). The impact of mobile learning in English language classrooms in Pakistan. *Asian-Pacific Journal of Second and Foreign Language Education*, *9*(1), Article 62. https://doi.org/10.1186/s40862-024-00274-0
32	Aşık, A., Sert, O., & Miller, P. (2024). The affordances of a mobile video-tagging tool for evaluating presentation skills in a second language. *Reflective Practice*, *25*(2), 145-163. https://doi.org/10.1080/14623943.2024.2305878
33	Chan, V. (2024). Using a virtual reality mobile application for interpreting learning: listening to the students' voice. *Interactive Learning Environments*, *32*(6), 2438-2451. https://doi.org/10.1080/10494820.2022.2147958
34	Li, L., & Kim, M. (2024). It is like a friend to me: Critical usage of automated feedback systems by self-regulating English learners in higher education. *Australasian Journal of Educational Technology*, *40*(1), 1-18. https://doi.org/10.14742/ajet.8821

35	Li, X., Wang, K., Wang, Y., He, J., & Zhang, J. (2024). Assessing User Experiences in Virtual Reality Learning Environments. *Asia-Pacific Education Researcher*, *33*(5), 1149-1160. https://doi.org/10.1007/s40299-023-00785-3
36	Luo, Y., & Watts, M. (2024). Exploration of university students' lived experiences of using smartphones for English language learning. *Computer Assisted Language Learning*, *37*(4), 608-633. https://doi.org/10.1080/09588221.2022.2052904
37	Ma, Q., & Chiu, M. M. (2024). Self-regulated and collaborative personalised vocabulary learning approach in MALL. *Language Learning and Technology*, *28*(1), 1-28. https://doi.org/hdl.handle.net/10125/73579
38	Yang, X., & Hu, J. (2024). Distinctions between mobile-assisted and paper-based EFL reading comprehension performance: reading cognitive load as a mediator. *Computer Assisted Language Learning*, *37*(7), 2051-2082. https://doi.org/10.1080/09588221.2022.2143527
39	Zhang, D., & Pérez-Paredes, P. (2024). Chinese EFL learners' use of mobile dictionaries in reading comprehension tasks. *System*, *121*, Article 103221. https://doi.org/10.1016/j.system.2024.103221
40	Al-Abri, A., Ranjbaran Madiseh, F., & Morady Moghaddam, M. (2025). Exploring Learning-Oriented Assessment in Enhancing Students' Lexical Fluency Through MALL. *Asia-Pacific Education Researcher*, *34*(1), 1-13, Article 103893. https://doi.org/10.1007/s40299-024-00832-7
41	Ren, Y., & Tan, X. (2025). The influence mechanism of planned behavior and self-consciousness on learners' willingness to use language learning Apps: Validation based on structural equation modeling. *PLoS ONE*, *20*(3 March), Article e0319300. https://doi.org/10.1371/journal.pone.0319300
42	Zhang, D., Hennessy, S., & Pérez-Paredes, P. (2025). An investigation of Chinese EFL learners' acceptance of mobile dictionaries in English language learning. *Computer Assisted Language Learning*, *38*(3), 317-341. https://doi.org/10.1080/09588221.2023.2189915

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp.378-399
DOI: https://dx.doi.org/10.24093/awej/vol16no4.21

Half a Century of Research on Strategies for Vocabulary Learning: A Scientometric Analysis

Xuan Wang
School of Foreign Studies
Northwestern Polytechnical University
Corresponding Author: xuan.emily.wang@nwpu.edu.cn

Xinyi Shen
School of Foreign Studies
Northwestern Polytechnical University
Email: shenxinyi@mail.nwpu.edu.cn

Feiyu Liu
School of Foreign Studies
Northwestern Polytechnical University
Email: liufeiyu@mail.nwpu.edu.cn

Received: 08/25/2025 Accepted: 11/19/2025 Published:12/10/2025

Abstract

In the past five decades, a significant amount of research has focused on strategies for vocabulary learning, given its crucial role in language acquisition. This study utilized CiteSpace to systematically review 340 studies on strategies for vocabulary learning from 1971 (when they were first indexed) to 2024, in the Social Sciences Citation Index and the Arts & Humanities Citation Index collection. The review focused on three aspects: development stages, key scholarly foundations, and main research streams. The findings reveal that the field of strategies for vocabulary learning research has undergone four distinct phases of development: the embryonic stage (1971-1990), the growth stage (1991-2005), the expansion stage (2006-2015), and the rapid development stage (2016-2024). Notably, there has been a substantial increase in annual publications, particularly since 2006, which can be attributed to the prolific publication of scholarly monographs and the impact of modern technology. The primary focus of strategies for vocabulary learning research lies in motivational belief, mobile technology, and digital game-based learning, with meta-analysis serving as an important research methodology. Furthermore, the findings highlight the interdisciplinary nature of vocabulary learning strategies, with contributions from computer science, psychology, and education.

Keywords: Citespace, scientometric review, strategic learning, strategies for vocabulary learning, technology-assisted language learning

How to cite: Wang, X., Shen, X., & Liu, F. (2025). Half a Century of Research on Strategies for Vocabulary Learning: A Scientometric Analysis. *Arab World English Journal, 16*(4):378-399. https://dx.doi.org/10.24093/awej/vol16no4.21

Introduction

Vocabulary knowledge significantly influences language acquisition due to the irreplaceable role of such knowledge in language skills development, particularly for listening, speaking, reading, and writing skills (Chen et al., 2025; Rueda-Sánchez et al., 2025; Wang-Taylor & Clenton, 2022; Zhang, 2022). However, vocabulary acquisition and learning pose considerable challenges for language learners (Babayiğit & Trenkic, 2025; Mangold et al., 2025; Ueno & Takeuchi, 2025) due to the complex cognitive processes within the multifaceted construct of vocabulary knowledge (González-fernández & Schmitt, 2020; Mujtaba et al., 2021; Wang et al., 2025). Strategies for vocabulary learning that enable self-regulated learning capacity, behavioral engagement, and individual pace choice are crucial for overcoming these difficulties (Chacón-Beltrán, 2018; Gan et al., 2025; Gay, 2022).

Existing research has substantially contributed to conceptualizing and measuring vocabulary learning strategies (Ali, 2020; Shamsan et al., 2021), as well as to examining their relationships with learning outcomes (Al Zahrani & Chaudhary, 2022; Wang & Feng, 2025). However, very few literature reviews focus on strategies for vocabulary learning using scientometrics for a systematic and dynamic analysis; instead, reviews have relied mainly on traditional literature review methods, which are prone to bias. Scientometrics examine emerging trends, thematic patterns, and the intellectual landscape by analyzing metrics and indicators in bibliographic records, including citation connections. Through this approach, researchers can understand how scientific fields evolve, recognize insightful patterns from visually encoded information, and synthesize diverse evidence, ultimately aiming, in Chen's terms, "to capture the state of the art of the domain" (2017, p. 2). In the field of linguistics, scientometric reviews typically include more than 100 articles (Wang et al., 2023). Thus, the present study aims to employ scientometric analysis to review studies on strategies for vocabulary learning. To systematically explore the intellectual landscape of strategies for vocabulary learning, we seek to answer the following research questions:

(1) What are the stages of research in the field of strategies for vocabulary learning?
(2) Which scholars, references, and sources of publications are most influential in the field of strategies for vocabulary learning?
(3) What are the key areas of research on strategies for vocabulary learning?

Literature review

Over the past five decades, significant research attention has focused on strategies for vocabulary learning, with scope changes from vocabulary learning strategies that emphasize outcomes and tasks to strategic vocabulary learning that emphasizes processes and regulation (Chen & Yeh, 2025; Hsu et al., 2024; Wang & Cohen, 2023). Vocabulary learning strategies are task-based learning strategies that help language learners acquire and retain vocabulary knowledge (Gu & Johnson, 1996; Oxford, 1990; Schmitt, 1997). Based on the form, meaning, and use of vocabulary knowledge, Nation (2001) taxonomized vocabulary learning strategies into four categories: planning, resources, process, and skill in use. Strategic vocabulary learning brought the concept of strategic learning from educational psychology into the field of vocabulary research and considers such learning an active and constructive process through which learners first set goals for learning and then monitor, regulate, and control their own cognition, motivation, and behavior (LaBontee, 2019; Murphy et al., 2023; Ziegler, 2015). It has five categories: effort

control, metacognitive control, fatigue control, emotion control, and environment control (Tseng et al., 2006; Tseng & Schmitt, 2008). Although vocabulary learning strategies and strategic vocabulary learning developed in different periods, both act as strategies for vocabulary learning and make complementary contributions to vocabulary knowledge development (Alharbi, 2024; Wang & Cohen, 2021; Wang & Cohen, 2023). Therefore, this study refers to the two as complementary strategies under the umbrella of *strategies for vocabulary learning*.

Given the development of the theoretical framework for strategies for vocabulary learning, instruments for their measurement, and numerous empirical studies have concentrated on three themes. Scales have been developed and validated for observing and describing the strategies, including the Vocabulary Learning Questionnaire (Gu, 2018; Gu & Johnson, 1996) and Self-Regulating Capacity in Vocabulary Learning Scale (Mizumoto & Takeuchi, 2012; Tseng et al., 2006). In empirical studies, researchers have explored the following themes: learners' personal preference on strategies (Fan, 2003; Fang-Chuan Ou & Wen-Chi Vivian, 2015; Rahimi & Allahyari, 2019), the relationship between strategy use and vocabulary learning effectiveness (Chiang et al., 2023; Fan, 2020; Kakihana, 2024; Shi, 2025), and differences in learners' strategy use due to individual difference factors such as gender, motivation, and language proficiency (Clarke, 2018; Lee, 2020; Tam & Kim, 2024). Advancements in educational technology have also led to the investigation of strategies on mobile-assisted vocabulary learning and digital game-based vocabulary learning; with the help of the new devices or platforms, strategies also influence learning approach tactics (Ou-Yang & Wu, 2016; Shi & Tsai, 2024; Wang & Feng, 2025; Zhang et al., 2023, 2024).

Primarily through empirical studies or book chapters, the existing literature has produced key findings in conceptualizing, measuring, and establishing relationships to learning outcomes based on strategies for vocabulary learning, with very few dedicated reviews on the subject. However, reviews are necessary to the development of a discipline because they help transfer tacit knowledge to codified and routinized knowledge (Chen, 2017). The evolution of a scientific discipline has four stages: introducing concepts, developing tools, broadening the scope, and codifying knowledge through systematic reviews (Chen, 2017; Shneider, 2009). By discussing concepts, developing questionnaires, and exploring strategic learning, the extant literature on strategies for vocabulary learning has already progressed through the first three stages. Now is thus an opportune time to conduct a comprehensive review of the field and explore future developments.

Method

Data Sources

For this study, the data sources were articles and reviews from the Social Sciences Citation Index (SSCI) and the Arts & Humanities Citation Index (AHCI), widely recognized as databases with collections of high-quality peer-reviewed articles for literature retrieval and research (Larivière et al., 2006; Larsen & von Ins, 2010). Both citation indexes are part of the Web of Science (WoS) Core Collection managed by Clarivate Analytics. By reviewing literature manually and piloting the WoS search, we found that search terms and keywords for strategies for vocabulary learning have changed alongside the development of research. The terms used in the early stage somewhat differ from those in the later stage, and the field has been widely debated for the past two decades, with scope changes from vocabulary learning strategies to strategic vocabulary learning regulation

(Gu & Johnson, 1996; LaBontee, 2019; Mizumoto & Takeuchi, 2009; Zhang & Lu, 2015). Thus, to ensure scientifically accurate data files, we used the combined phrases "vocabulary learning strateg*[1]" or "self-regulat* AND vocabulary learning" or "vocabulary learning AND strateg*" or "self-regulat* AND vocabulary acquisition" or "vocabulary strateg*" or "vocabulary related strateg*" as the subject keywords for the search. Manzo and Sherk (1971) first proposed some generalizations and strategies to guide vocabulary learning, so we set the time span to 54 years (1971–2024) to facilitate a comprehensive understanding of the research status and evolution from the initial stage of strategies for vocabulary learning research to the present. Then, to further refine the data and ensure its accuracy, we set the language as "English" and the document types as "article" and "review article." The above search formula yielded 340 search results for the SSCI and the AHCI data set.

Data Analysis

In this study, we employed the CiteSpace 6.4.R2 Advanced version to visualize and review literature in the field of vocabulary learning strategies. CiteSpace is an information visualization and analysis tool for bibliometric calculus to help users analyze problems through computer algorithms and interactive visualizations. The tool also visualizes literature information, data, and relationships via scientific mapping based on the needs of researchers, helping them understand the development history, research status, research hotspots, and research frontiers of a field more scientifically and precisely (Chen, 2016). CiteSpace's visual analysis of research dynamics has been widely used in applied linguistics research, including in studies on multiword expressions (Yi et al., 2025), translanguaging (Lu & Gu, 2024), writing assessment (Dong et al., 2024), and language assessment literacy (Wang et al., 2023). Therefore, scientometric review by CiteSpace has been employed to explore three main fields of the current study: research development stages, influential scholars, articles and journals, and key areas. We employed CiteSpace to process and analyze the data

Results

Research Development Stages

The annual number of publications in a field can reflect the academic circle within the field, which is a significant factor for analyzing the development status of a field and predicting its future development trend (Hanssen et al., 2018). As shown in Figure 1, the research dynamics of strategies for vocabulary learning are active and overall show a fluctuating rise. The figure also shows the theoretical base attention via the annual cited line, which further assists the analysis. Generally, the publication of research on strategies for vocabulary learning has gone through four stages: the embryonic stage (1971–1990), the growth stage (1991–2005), the expansion stage (2006–2015), and the rapid development stage (2016–2024).

[1] A wildcard "*" is a special symbol that can be used to represent any set of characters with no limit on the number of characters, including no characters. The wildcard "*" can be used in retrieval to perform an advanced search. Since both "vocabulary learning strategy" and "vocabulary learning strategies" have been used in the literature, we used the wildcard "*" to ensure the rigor and accuracy of the data. We thus used the following subject keywords for the search: "vocabulary learning strateg*".

In the first stage (1971–1990), there were a few publications on strategies for vocabulary learning, which indicates that researchers paid less attention to this field and that strategies for vocabulary learning research was still in its infancy. Researchers at the time developed an interest in language learning strategies (Manzo & Sherk, 1971; Pressley et al., 1982). Within the research on language learning strategies, strategies for vocabulary learning have also been investigated, with researchers mainly focused on examining the learning effects of specific strategies through experiments. In 1990, Batia Laufer, Robin Goulden, Paul Nation, and John Read—prominent figures in the field—published articles on vocabulary in top journals, including in *Applied Linguistics* and *Foreign Language Annals*; the articles are not directly about vocabulary learning strategies but did discuss not only the difficulty and width of vocabulary knowledge, but also discussed how to make strategies a great potential research topic (Goulden et al., 1990; Laufer, 1990).

In the second stage (1991–2005), the number of publications on strategies for vocabulary learning research increased but fluctuated, suggesting researchers' escalating focus on the field. Some significant research has been conducted since 1995. Researchers developed a range of taxonomies for vocabulary learning strategies through interviews, observations, and other qualitative research methods (e.g., Nation 2001; Schmitt, 1997). Concurrently, researchers have carried out numerous empirical studies to explore different issues, including the relationship between vocabulary learning strategies use and learning efficiency (e.g., Gu & Johnson, 1996; Kojic-Sabo & Lightbown, 1999) and the factors influencing the use of strategies (e.g., Catalan, 2003; Gu, 2002). Yet, even though strategies for vocabulary learning research has existed for a long time and became a well-established academic discipline, there was still no acknowledged theory of strategies for vocabulary learning due to the complexity of vocabulary learning. The classification and measurement of strategies for vocabulary learning were also controversial. Many scholars have, as a result, expressed concerns about the validity and rigor of strategy research. For example, Dörnyei (2005) ramped up criticism of vocabulary learning strategies and introduced self-regulated learning into the field of strategies for vocabulary learning research. The emergence of research challenges ultimately ushered in a period of instability in the field.

In the third stage (2006–2015), the increased publication volume of research results remained relatively stable. New research perspectives were introduced into the strategies for vocabulary learning field during this stage, which aroused more academic attention and discussion. Tseng et al. (2006) developed a new measurement instrument based on self-regulated learning (i.e., the Self-Regulating Capacity in Vocabulary Learning Scale) and advocated for vocabulary learning research from the perspective of self-regulation. Since then, the field of strategies for learning vocabulary has been in a state of change as the researchers' views provoked intense discussions. The developments of this period led to a breakthrough in strategies for learning vocabulary research. In addition, Pavii Taka (2008) published *Vocabulary Learning Strategies and Foreign Language Acquisition*, the first book entirely focused on strategies for vocabulary learning. The book provided a detailed review of strategies for vocabulary learning research findings, offering timely, thoughtful, and welcome guidance and assistance to researchers. Moreover, the author's review of future research trends provided proper direction for the development of the field. Collectively, these efforts laid the foundation for the field to grow.

In the fourth stage (2016–2024), following the continuous advancement of the underlying theory and the introduction of new perspectives, strategies for vocabulary learning publications showed rapid and relatively stable growth, except during the period around the COVID-19

pandemic. Notably, the publication skyrocketed from 2021 to 2022 and from 2023 to 2024; the latter saw explosive growth with almost twice the number of articles published than in the previous year. One important event occurred in mid-2016: The American Association for Applied Linguistics (AAAL) announced that vocabulary and lexical studies would become a separate guest strand for AAAL 2017. Since AAAL 2018, the studies have become a permanent fixture. AAAL is one of the most influential and popular annual conferences for applied linguistics, so the shift also attracted more researchers and scholars focused on vocabulary research and strategies for vocabulary learning. In addition, the significant impact of modern technology on learners' strategies for learning vocabulary cannot be ignored during this period. First, the advancement of Massive Open Online Courses (MOOC) lessons, mobile-assisted learning approach, and digital games promoted self-learning on vocabulary and changed learners' strategies. Second, during the COVID-19 period, especially the lockdowns of cities and schools, learners changed how they learned vocabulary. Computer-assisted language learning (CALL) and mobile-assisted language learning (MALL) education technologies, for instance, offered interesting and effective ways to facilitate vocabulary acquisition, replacing monotonous methods such as word lists and rote memorization (Daly, 2022). In addition, learners today grow up in a technologically advanced society, which has altered their learning habits. Technology-integrated curriculum can shape creative teaching processes and stimulate learners' engagement in learning materials, and improve learners' language practice (Liu & Chen, 2023). Finally, the combination of mobile communication technology and modern educational technology has promoted the development of mobile learning. The COVID-19 pandemic resulted in the prevalence of online learning, so the use of mobile applications for vocabulary learning has since become an indispensable part of language learners' study routines (Guo & Li, 2022). Together, these factors have had a significant and positive impact on strategies for modern language learners, resulting in rapid increases in the exploration of strategies for learning vocabulary in new contexts. The rise in attention also indicates the field's substantial potential for development. In the coming years, strategies for learning vocabulary will thus arguably be one of the focused issues in the field of language acquisition research.

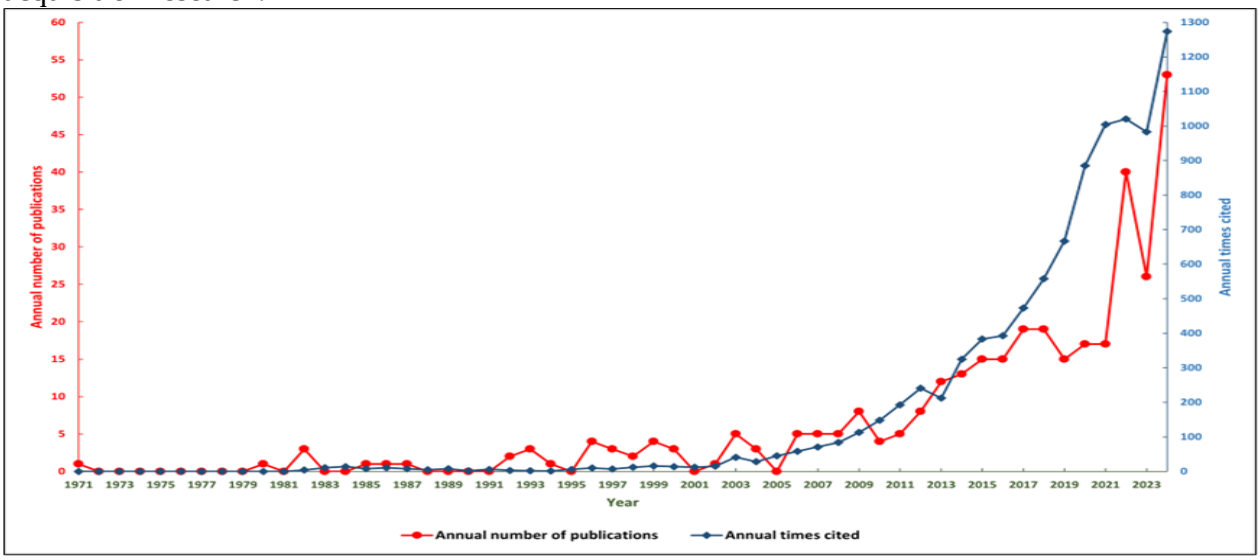


Figure 1. Quantitative distribution of strategies for vocabulary learning literature from 1971 to 2024

ISSN: 2229-9327

Influential Scholars, References, and Sources of Publications

Influential Scholars

Table 1 lists the most-cited authors identified using CiteSpace, broken down by those with the highest citation counts and the strongest citation bursts. While citation counts are accumulative counts of an author's articles, citation bursts are a dynamic index indicating a quick rise in the citation of an author's articles. High citation counts mean an author's articles are fundamental to a field; strong citation bursts attract attention from researchers in the field and indicate an author is pioneering a new area in the field. Paul Nation (Emeritus Professor at Victoria University of Wellington), Norbert Schmitt (Emeritus Professor at University of Nottingham), and Batia Laufer (Emeritus Professor at University of Haifa) are the three most-cited authors based on counts. They are all founders in the field of second language vocabulary acquisition. Meanwhile, Michael Pressley (University of Western Ontario), Richard Atkinson (Emeritus Professor at University of California, San Diego), and Batia Laufer are the three most-cited authors based on bursts, with the latter appearing on both lists. While the top three authors on the citation count list are from the second vocabulary learning field, Michael Pressley and Richard Atkinson are from the cognitive science field and, according to CiteSpace burst analysis, pioneered new fields between 1982 and 2015. Most remaining authors in both lists have conducted long-term research on vocabulary; however, interdisciplinary or cross-field characteristics have also been evident in the authors' research fields. For example, Gordon Bower and Fergus Craik are interdisciplinary scholars from the field of cognitive psychology, which focuses on how semantic processing influences memory; Zoltán Dörnyei and Barry Zimmerman are cross-field scholars from the fields of motivation and self-regulated learning.

Table 1. *Most-cited authors*

Authors with the highest citation counts		Authors with the strongest citation bursts	
Cited authors	Counts	Cited authors	Strength
Paul Nation	159	Michael Pressley	13.65
Norbert Schmitt	120	Richard Atkinson	10.11
Batia Laufer	86	Batia Laufer	9.11
Rebecca Oxford	73	Joel Levin	8.64
Gu Yongqi	61	Andrew Cohen	7.47
Stuart Webb	54	Michael Lawson	6.95
William Nagy	50	Jan Hulstijn	6.53
Joe Barcroft	44	William Nagy	6.51
Wen-Ta Tseng	42	Paul Meara	5.66
Fergus Craik	42	John Read	5.62

Influential References

Table 2 describes the most-cited articles in the CiteSpace network (340 total articles) and the WoS literature. While the CiteSpace data includes articles from SSCI and AHCI from WoS, the WoS literature includes other databases (e.g., SCI, ESCI, etc.) that did not appear in the CiteSpace data. The overlapping number of articles between the two is thus relatively large. However, SSCI and AHCI are core collections of high-quality articles in the social sciences and humanities, so we cultivated lists based on CiteSpace data but provided the WOS counts for reference. All articles (N = 17) with no fewer than five citations in the network were included. These articles represent foundational research in the fields and include empirical studies (N = 11),

meta-analyses (N = 4), and general reviews (N = 2); one of each study type appears in the top three cited studies. Regarding areas of focus, the general review and meta-analysis articles mainly provide insights related to technology-assisted learning, such as mobile-assisted learning (Lin & Lin, 2019; Mahdi, 2018) and digital game-based learning (Zou et al., 2021). The three main topics in the empirical studies were mobile-assisted learning for vocabulary strategies (García Botero et al., 2019), self-directed learning (Chen et al., 2019; García Botero et al., 2019; Tseng et al., 2006), and young learners (Andrä et al., 2020; Dickinson et al., 2019; Hadley et al., 2019; Teng, 2022). Notably, all the empirical articles employ quantitative research methods, including hypothesis testing, experiment design, and causal-comparative research.

Sources of Publications

Academic journals disseminate papers and academic research results. Thus, analyzing journal information helps researchers identify the most authoritative and important journals in the field and focus more on details and information about the research status and research basis. The number of journal publications partially reflects the research frontier and any emerging disciplines in the field. We therefore listed the top 10 journals in the field of vocabulary learning strategies based on the number of publications from 1971 to 2024, as shown in the left column of Table 3. Meanwhile, the cited journals reflect the research base of the field, so we selected the top 10 cited journals for vocabulary learning research from 1971 to 2024, as shown in the right column of Table 3. Five journals appear in both lists: *Language Learning*, *Modern Language Journal*, *System*, *TESOL Quarterly*, and *Language Teaching Research*. This overlap implies that strategies for vocabulary learning research is both a foundational topic and one garnering significant attention in applied linguistics and TESOL journals. In addition, *Computer Assisted Language Learning*, *Frontiers in Psychology*, *Educational Technology Society*, and *Computers & Education* are top journals for the field based on publication data. This finding suggests that research on strategies for vocabulary learning reflects some interdisciplinary focus. In other words, the language and education fields interact with the psychology, computer, and sociology fields alongside the development of technology and society.

Main Research Streams

In research, co-citation is employed to describe the frequency with which two documents are cited together in other documents, a type of analysis considered to be a core profile for a specialty research field that helps elucidate the pattern of linkages among key publications (Small, 1973). The co-citation relationship of literature might change over time, so literature co-citation analysis allows researchers to explore how a field develops and evolves (Chen, 2020). To understand the thematic patterns of strategies for vocabulary learning and their knowledge base, we use co-citation analysis and the clustering function to visualize the co-citation relationship using CiteSpace, whose co-citation analysis function has been shown to be most effective in highlighting co-citation patterns (Chen, 2016). We identified 36 main clusters using CiteSpace, indicating base research themes within the field. Based on our analysis, the 0.96 modularity Q value suggests the clusters in the co-citation network are well-separated, and the 0.98 weighted mean silhouette score suggests the references within each cluster are coherent in the research theme and distinct from other clusters (Newman, 2006). Figure 2 illustrates 19 major clusters, and Table 4 presents the top nine clusters.

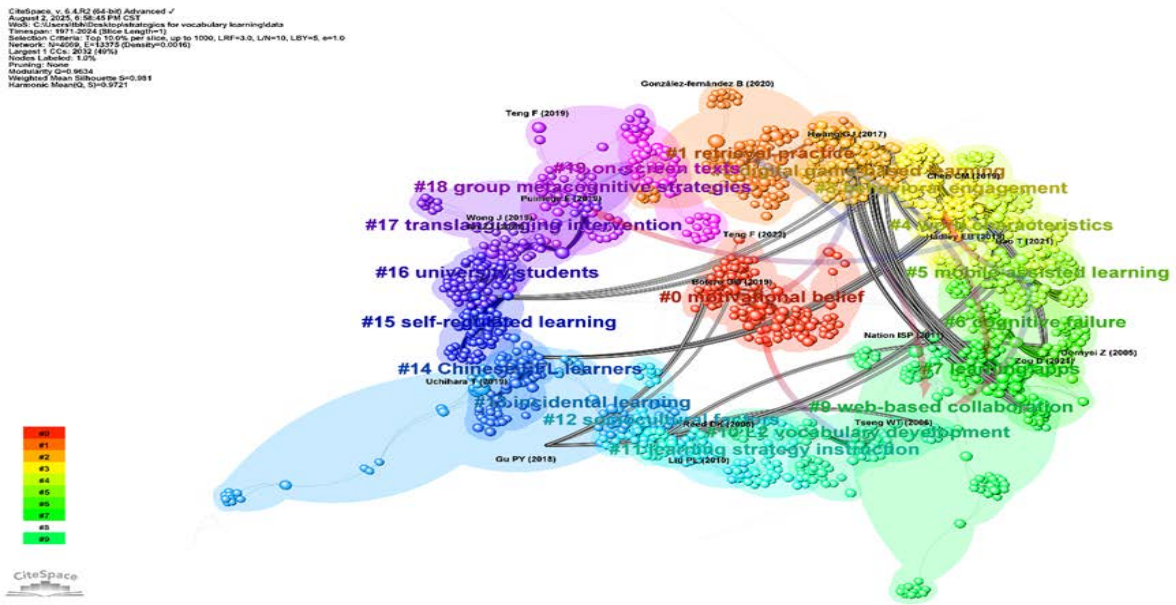


Figure 2. Co-Citation network clustersTable 2. *Most-cited references (1971–2024)*

Number of citations in the network/WoS literature	Cited reference and year	Source	Title	Type of paper
12/75	(Chen et al., 2019)	Computer Assisted Language Learning	An English vocabulary learning app with self-regulated learning mechanism to improve learning performance and motivation	quasi-experimental design
12/156	(Lin & Lin, 2019)	Computer Assisted Language Learning	Mobile-assisted ESL/EFL vocabulary learning: a systematic review and meta-analysis	meta-analysis
11/151	(Zou et al., 2021)	Computer Assisted Language Learning	Digital game-based vocabulary learning: where are we and where are we going?	review
7/147	(Uchihara et al., 2019)	Language Learning	The effects of repetition on incidental vocabulary learning: a meta-analysis of correlational studies	meta-analysis
6/45	(Andrä et al., 2020)	Educational Psychology Review	Learning foreign language vocabulary with gestures and pictures enhances vocabulary memory for several months post-learning in eight-year-old school children	experiment design
6/108	(González-fernández & Schmitt, 2020)	Applied Linguistics	Word knowledge: exploring the relationships and order of acquisition of vocabulary knowledge components	causal-comparative research
5/33	(Dickinson, Nesbitt, et al., 2019)	Early Childhood Research Quarterly	Effects of language on initial reading: direct and indirect associations between code and language from preschool to first grade	hypothesis testing
5/24	(Fan, 2020)	Frontiers in Psychology	Strategy use in second language vocabulary learning and its relationships with the breadth and depth of vocabulary	causal-comparative research

			knowledge: a structural equation modelling study	
5/109	(García Botero et al., 2019)	Computer Assisted Language Learning	Self-directed language learning in a mobile-assisted, out-of-class context: do students walk the talk	causal-comparative research
5/52	(Hadley et al., 2019)	Reading Research Quarterly	Building semantic networks: the impact of a vocabulary intervention on preschoolers' depth of word knowledge	experiment design
5/42	(Hao et al., 2021)	Journal of Research on Educational Effectiveness	Technology-assisted vocabulary learning for EFL learners: a meta-analysis	meta-analysis
5/155	(Hwang et al., 2017)	Computers & Education	Interaction of problem-based gaming and learning anxiety in language students' English listening performance and progressive behavioral patterns	quasi-experiment
5/69	(Mahdi, 2018)	Journal of Educational Computing Research	Effectiveness of mobile devices on vocabulary learning: a meta-analysis	meta-analysis
5/88	(Puimège & Peters, 2019)	Language Learning	Learners' English vocabulary knowledge prior to formal instruction: the role of learner-related and word-related variables	causal-comparative research
5/19	(Teng, 2022)	Language Awareness	Exploring awareness of metacognitive knowledge and acquisition of vocabulary knowledge in primary grades: a latent growth curve modelling approach	causal-comparative research
5/421	(Tseng et al., 2006)	Applied Linguistics	A new approach to assessing strategic learning: the case of self-regulation in vocabulary acquisition	causal-comparative research
5/190	(Zhang & Zou, 2022)	Computer Assisted Language Learning	Types, purposes, and effectiveness of state-of-the-art technologies for second and foreign language learning	review

Table 3. *Most influential journals of strategies for vocabulary learning from 1971 to 2024*

Most Published Journals			Most Cited Journals		
Rank	Journals	Counts	Rank	Journals	Counts
1	Computer Assisted Language Learning	15	1	Language Learning	175
2	Frontiers in Psychology	15	2	Modern Language Journal	158
3	Language Teaching Research	12	3	Journal of Educational Psychology	151
4	Educational Technology Society	11	4	System	137
5	Sage Open	11	5	Applied Linguistics	126
6	Computers & Education	9	6	TESOL Quarterly	115
7	Modern Language Journal	9	7	Reading Research Quarterly	93
8	System	9	8	Language Teaching Research	88
9	Language Learning	8	9	Studies in Second Language Acquisition	80
10	International Journal of Applied Linguistics	6	10	The Review of Educational Research	79
10	Journal of Educational Computing Research	6			
10	Porta Linguarum	6			
10	TESOL Quarterly	6			

ISSN: 2229-9327

The top clusters in Table 4 can be summarized into two main areas. Area 1 established the basic theoretical strategies for vocabulary learning from the perspectives of cognition and behavior. Based on the timeline, Cluster #1 retrieval practice and Cluster #4 word characteristics relate to cognition and appeared earlier in averaged 2016 than clusters about behavior in averaged 2019. Retrieval practice is a learning strategy that enables learners to actively recall information from memory, which ultimately contributes to long-term vocabulary knowledge retention and a deeper understanding (Rowland, 2014; van den Broek et al., 2018). These positive effects are considered testing effects, whose conditions can be flashcards, cloze tasks, dictations, or multimedia environments (Elgort et al., 2018; Goossens et al., 2016; van den Broek et al., 2021). On the other hand, word characteristics describe how words are perceptually accessible to learners and are characterized based on four elements: shape, individuation, concreteness, and imageability (Hadley et al., 2021; Hadley et al., 2016). For most language learners, vocabulary learning represents the initial stage of learning as a foundational reading skill, so researchers have investigated vocabulary learning strategies based on word characteristics and the relevant instructions among pre-schoolers, especially in the first-language context (Dickinson et al., 2019; Dwyer & Schachter, 2020; Hindman et al., 2019). The two clusters marked the transition from theory to practice—from memory-based cognition to in-classroom vocabulary learning (Elgort et al., 2018; van den Broek et al., 2018).

Table 4. *Top clusters based on the CiteSpace analysis*

ClusterID	Label	Size	Silhouette	Mean (year)
0	motivational belief	133	0.949	2019
1	retrieval practice	131	0.979	2016
2	digital game-based learning	124	0.95	2019
3	behavioral engagement	104	0.945	2019
4	word characteristics	97	0.993	2016
5	mobile-assisted learning	85	0.995	2020
6	cognitive failure	84	0.975	2007
7	learning apps	84	0.984	2017
9	web-based collaboration	77	0.986	2015

Regarding Cluster #0 motivational belief and Cluster #3 behavioral engagement in the same average year of 2019, Area 1 establishes both from the perspective of behavior, with motivational belief being the largest among all clusters. Cluster #0 motivational belief marked a shift in strategy for vocabulary learning research to the field of psychology. That shift has led to research on the role of self-efficacy and motivation in vocabulary size and depth (Lee et al., 2022), and the second language context is one of the most obvious settings for empirical studies (Guo et al., 2022; Li & Wang, 2024). Behavior engagement emphasizes the learning process instead of final outcomes—for example, by focusing on how learners self-regulate their vocabulary learning strategies to achieve learning goals (Lei et al., 2022) and considering behavior a dynamic development influenced by social and technological contexts (Hwang & Fu, 2019; Wei, 2023).

Area 2 focused on the role of technology in vocabulary learning strategies, and the shifts over time illustrate how the clusters capture changing attitudes toward technology in the scholarship—primarily a shift from rejection and limited applications to widespread acceptance and use. In addition, the average year of these clusters demonstrated the shift from web-based to mobile-assisted technology development. Cluster #6 cognitive failure appeared the earliest in averaged 2007, and focuses on learners' tiredness, exhaustion, and helplessness in technology-assisted learning. With the development of technology use in vocabulary learning, learners'

internet cognitive failure gained attention in the later stages of research on this topic, with cognitive failure being found to be negatively related to learners' self-regulation in vocabulary learning (Hong et al., 2015). Cluster #9 web-based collaboration was the second in average 2015. Sarah Hsueh-Jui et al. (2014) explored how English as a foreign language (EFL) learners' web-based collaboration affects vocabulary improvements, finding that willingness to engage in web-based collaboration was positively related to learners' vocabulary knowledge. Meanwhile, Cluster #7 learning apps appeared in 2017, and Cluster #5 mobile-assisted learning was the latest in average 2020. These two clusters focus on mobile-assisted vocabulary learning research on smartphones and tablets, which exerted a far-reaching influence on vocabulary learning strategies. *Baicizhan*, *Kingsoft Vocabulary*, *New Oriental Fun Vocabulary*, and *Shanbay* were some of the apps mentioned in research, and learner age was deemed a crucial factor for game design and application (Guo et al., 2022; Li, 2021; Yu, 2018). Among the clusters, Cluster #2 digital game-based learning acted as the fourth in time sequences in average 2019. To increase students' engagement in learning and improve learning motivation, problem-based games and competitive gaming scenarios have been employed (Hwang et al., 2017; Wei et al., 2018).

Burst detection investigates instances where the citation frequency of a particular document experiences a sudden and significant increase within a short period, which helps identify unanticipated but widely distributed documents and reveals the evolution of research hotspots and frontier trends (Chen, 2006; Kleinberg, 2003). Different from other fields in applied linguistics, monographs burst frequently in the field of vocabulary development research due to their unique nature and role being increasingly noticed by scholarship in applied linguistics. Three monographs and seven articles with the strongest bursts detected are shown in Table 5. Specifically, Nation (2001), Nation and Webb (2011), and Webb and Nation (2017) are the three monographs that laid the fundamentals of vocabulary learning strategies, including conceptualization, taxonomy, principles for teaching, and instruments for measure. The other seven documents are research articles: Sandberg et al. (2014) in Cluster #7 learning apps and Hadley et al. (2016) in Cluster #4 word characteristics are from the top clusters, while Oxford and Scarcella (1994), Fan (2003), Huang et al. (2012), Smith et al. (2013), and Lin and Lin (2019) are from relatively small clusters. The documents in both top clusters and bursts indicate that the documents are both theoretical foundations and emerging trends, while articles only in the top burst imply their main contribution is pioneering new fields in vocabulary learning strategies. For example, Smith et al. (2013) investigated vocabulary learning experiences with inference-based computer games embedded in eBooks and the advantages of the approach compared to traditional learning, offering a new orientation from the perspectives of both research and technology.

Table 5. *Documents with the strongest bursts*

Bursts	Author	Title	Source/Publisher	Cluster ID
2.86	Webb & Nation (2017)	How vocabulary is learned	Oxford University Press	13
2.70	Lin and Lin (2019)	Mobile-assisted ESL/EFL vocabulary learning: A systematic review and meta-analysis	Computer Assisted Language Learning	365
2.45	Smith et al. (2013)	Play games or study? Computer games in eBooks to learn English vocabulary	Computers & Education	171
2.43	Nation & Webb (2011)	Researching and analysing vocabulary	Heinle Cengage Learning	9

2.41	Sandberg et al. (2014)	The added value of a gaming context and intelligent adaptation for a mobile learning application for vocabulary learning	Computers & Education	7
2.31	Huang et al. (2012)	A ubiquitous English vocabulary learning system: Evidence of active/passive attitudes vs. usefulness/ease-of-use	Computers & Education	21
2.10	Hadley et al. (2016)	Examining the acquisition of vocabulary knowledge depth among preschool students	Reading Research Quarterly	4
2.05	Nation (2001)	Learning vocabulary in another language	Cambridge University Press	81
2.02	Oxford & Scarcella (1994)	Second language vocabulary learning among adults: State of the art in vocabulary instruction	System	113
2.02	Fan (2003)	Frequency of use, perceived usefulness, and actual usefulness of second language vocabulary strategies: A study of Hong Kong learners	The Modern Language Journal	250

Discussion

A total of 340 documents in the field of vocabulary learning strategies were analyzed using bibliometric analysis. First, we examined the stages of research on vocabulary learning strategies. Based on the annual publication numbers (Figure 1), the curve representing the trend of publication volume started soaring in 2016 after a long period of ups and downs. The increase reflects how the development of strategies for vocabulary learning research started to flourish after a long period of volatile exploration. Arguably, then, scholars' efforts in the exploratory phase laid the foundation for the current spurt of strategies for vocabulary learning research. During the growth stage (1991–2005), scholars' contributions to the taxonomy of strategies established the theoretical basis for the study of strategies for vocabulary learning and drove its development. The subsequent publication of the book *Vocabulary Learning Strategies and Foreign Language Acquisition* during the expansion phase (2006–2015) made significant contributions to the development of the field. The book not only reviews past research on strategies for vocabulary learning but also highlights future research directions. At the same time, the introduction of self-regulated learning has brought new perspectives to strategies for vocabulary learning research.

Second, the disciplinary nature of strategies for vocabulary learning research has also been found in influential authors' works, articles, and journals, and analyses of key topic analyses. Based on the analysis of journals, strategies for vocabulary learning research has obvious interdisciplinary characteristics, including those aligned with the linguistics, psychology, computer, technology, education, and other disciplines. The language, education, and linguistics disciplines are the basic fields within strategies for vocabulary learning research, while the psychology and computer disciplines are frontier fields. Collectively, these results reveal a clear interdisciplinary trend in research on vocabulary learning strategies. The introduction of self-regulated learning has also led to a strong connection between educational psychology and strategies for vocabulary learning research. Over the past two decades, the rapid expansion of digital technology has provided new tools and resources to meet learners' diverse vocabulary learning needs as well (Boroughani et al., 2023). Technology has become one of the most important tools for vocabulary learning, so the strategies used by learners have also changed (An et al., 2020). As a result, researchers have begun to explore the characteristics and efficiency of

learners' use of vocabulary learning strategies in the context of technology-assisted learning, which has led to more interaction between disciplines such as computer science and artificial intelligence. For instance, Hsu et al. (2024) investigated the effects of artificial intelligence (AI) image recognition technology on the vocabulary acquisition, learning anxiety, and learning behaviors of English language learners. An et al. (2020) developed and validated the Technology-Based Self-Regulated English Learning Strategies Scale, an evaluation tool used to assess EFL students' technology-based self-regulated English language learning experience.

Third, key topics within and the knowledge base of strategies for vocabulary learning research were also examined. The co-citation strategies for vocabulary learning network clusters (Figure 2) revealed that #0 motivational belief, #1 retrieval practice, #2 digital game-based learning, #3 behavioral engagement, #4 word characteristics, #5 mobile-assisted learning, #6 cognitive failure, #7 learning apps, and #9 web-based collaboration have formed a core research network. Strategies for vocabulary learning research presents significant features in terms of research object, research content, and research methods. The research mainly focuses on primary language learners as the research object. Meanwhile, due to the high number of Chinese English language learners (Huang et al., 2021), scholars have also explored extensive strategies for vocabulary learning research on this group. In addition, in terms of research content, mobile technology and teacher beliefs are important concerns. The rapid development of technology and its increasing integration with education have made mobile technology an integral part of strategies for vocabulary learning research (An et al., 2020). In second language vocabulary research, other topics of significant interest are how to support learners in learning vocabulary and how teachers can be better prepared to teach vocabulary (Chung, 2018). Researchers have thus studied teaching beliefs to gain insights into vocabulary instruction and teacher development to assess how to benefit vocabulary learners best. In terms of research methods, the many benefits of meta-analysis, including increased statistical power and improved accuracy, have led to this method being used in numerous strategies for vocabulary learning studies (Chan & Arvey, 2012). As shown in Table 5, the book *How vocabulary is learned* leads the way in co-citation bursts. In the book, Webb and Nation (2017) capitalized on research and theory to provide comprehensive and authoritative coverage of important elements of vocabulary instruction, providing clear and practical information for teachers and learners. The book thus serves as an important guide to research in the field of strategies for vocabulary learning.

Conclusion

This study employed the visualization tool CiteSpace to analyze strategies for vocabulary learning literature from the SSCI and AHCI databases, focusing on research trends, influential authors, articles, and journals, and key topic areas. Currently, discussions about vocabulary learning strategies and strategic learning are deepening, and the interdisciplinary integration between technology and education continues to expand. The findings thus suggest that the vigorous development stage of strategies for vocabulary learning research will persist in the long term. Among the selected articles, the most frequently researched subjects were motivational belief, mobile technology, meta-analysis, and digital game-based learning. Additionally, while not key clusters, social and cultural factors have emerged as a new area, and the strategies for vocabulary learning research indicates an expansion toward translanguaging acquisition. Overall, this study highlights the knowledge structure and research dynamics in this field by adopting

scientometric and visual analytic tools, providing valuable references for future strategies for vocabulary learning research.

Funding: This research was funded by the NPU Research Scheme for Education and Teaching Reform, grant number 2024JGWZ09.
Acknowledgments: Not applicable.
Conflicts of Interest: The authors declare no conflict of interest.
Authenticity: This manuscript is an original work.
Artificial Intelligence Statement: The authors of this manuscript used AI only as basic writing support tools, such as for checking grammar, vocabulary, and spelling.

About the authors

Dr. Xuan Wang is currently Associate Professor at the School of Foreign Studies, Northwestern Polytechnical University, China. She completed her PhD from the University of Sydney in 2018 and her research interest is second language vocabulary development, language testing and teaching, and language policy and planning. ORCID: https://orcid.org/0000-0002-5741-0065

Ms. Xinyi Shen is a postgraduate research student in applied linguistics at School of Foreign Studies, Northwestern Polytechnical University. Her areas of interest include curriculum development, language pedagogy, and tacit knowledge in education. ORCID: https://orcid.org/0000-0002-9111-1819

Ms. Feiyu Liu is an undergraduate student in English major at School of Foreign Studies, Northwestern Polytechnical University. Her research interests are applied linguistics, Teaching English to Speakers of Other Languages (TESOL). ORCID: https://orcid.org/ 0000-0001-5052-7626

References

Alharbi, K. (2024). Enhancing technical vocabulary acquisition: A multimedia strategic approach for English for specific purposes during COVID-19. *Interactive Learning Environments*, *32*(9), 4812-4830. https://doi.org/10.1080/10494820.2023.2205902

Ali, M. A. (2020). Investigation of Vocabulary Learning Strategies to Identify Word Meanings for Saudi EFL Students in Reading Context. *Arab World English Journal, 11* (3) 149-169. https://dx.doi.org/10.24093/awej/vol11no3.9

Al Zahrani, S. M. &, Chaudhary, A. (2022). Vocabulary learning strategies in ESP context: Knowledge and implication. *Arab World English Journal, 13* (1), 382- 393. https://dx.doi.org/10.24093/awej/vol13no1.25

An, Z., Gan, Z., & Wang, C. (2020). Profiling Chinese EFL students' technology-based self-regulated English learning strategies. *PLOS One*, *15*(10), 1-16. https://doi.org/10.1371/journal.pone.0240094

Andrä, C., Mathias, B., Schwager, A., Macedonia, M., & von Kriegstein, K. (2020). Learning foreign language vocabulary with gestures and pictures enhances vocabulary memory for several months post-learning in eight-year-old school children. *Educational Psychology Review*, *32*(3), 815-850. https://doi.org/10.1007/s10648-020-09527-z

Babayiğit, S., & Trenkic, D. (2025). Vocabulary knowledge is key to understanding and addressing disparities in higher education. *British Journal of Educational Psychology*, Advance online publication. https://doi.org/10.1111/bjep.70000

Boroughani, T., Behshad, N., & Xodabande, I. (2023). Mobile-assisted academic vocabulary learning with digital flashcards: Exploring the impacts on university students' self-regulatory capacity. *Frontiers in Psychology*, *14*, 1-8. https://doi.org/10.3389/fpsyg.2023.1112429

Catalan, R. (2003). Sex differences in L2 vocabulary strategies. *Journal of Applied Linguistics*, *13*, 54-77. http://dx.doi.org/10.1111/1473-4192.00037

Chacón-Beltrán, R. (2018). Vocabulary learning strategies outside the classroom context: What adults learn in a technology-based learner-centered environment. *The Language Learning Journal*, *46*(5), 583-593. https://doi.org/10.1080/09571736.2018.1503135

Chan, M. E., & Arvey, R. D. (2012). Meta-analysis and the development of knowledge. *Perspectives on Psychological Science*, *7*(1), 79-92. https://doi.org/10.1177/1745691611429355

Chen, C. H., & Yeh, H. C. (2025). Advancing English for specific purposes (ESP) pedagogy in engineering education: Infusing generative learning strategies into AR game design for enhanced vocabulary acquisition and critical thinking. *Computer Assisted Language Learning*, Advance online publication. https://doi.org/10.1080/09588221.2025.2518410

Chen, C. M., Chen, L. C., & Yang, S. M. (2019). An English vocabulary learning app with self-regulated learning mechanism to improve learning performance and motivation. *Computer Assisted Language Learning*, *32*(3), 237-260. https://doi.org/10.1080/09588221.2018.1485708

Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, *57*(3), 359-377. https://doi.org/10.1002/asi.20317

Chen, C. (2016). *CiteSpace: A practical guide for mapping scientific literature*. Nova Science Publishers.

Chen, C. (2017). Science mapping: A systematic review of the literature. *Journal of Data and Information Science*, *2*(2), 1-40. https://doi.org/10.1515/jdis-2017-0006

Chen, C. (2020). A glimpse of the first eight months of the COVID-19 literature on Microsoft academic graph: themes, citation contexts, and uncertainties. *Frontiers in Research Metrics and Analytics*, *5*, 1-15. https://doi.org/10.3389/frma.2020.607286

Chen, S., Ye, S. X., Weng, F., & Chen, W. (2025). Individual differences factors and listening comprehension: examining the roles of aural vocabulary knowledge and foreign language listening anxiety. *International Journal of Applied Linguistics*, Advance online publication. https://doi.org/10.1111/ijal.12805

Chiang, K. M., Chang, C. H., & Lee, I. (2023). Vocabulary learning strategy instruction in the EFL secondary classroom: An exploratory study. *Language Awareness*, *32*(1), 94-113. https://doi.org/10.1080/09658416.2021.1979566

Chung, E. (2018). Revisiting second language vocabulary teaching: Insights from Hong Kong in-service teachers. *The Asia-Pacific Education Researcher*, *27*(6), 499-508. https://doi.org/10.1007/s40299-018-0412-3

Clarke, D. C. (2018). Student responses to vocabulary learning strategies on an ESAP course. *ELT journal*, *72*(3), 319-328. https://doi.org/10.1093/elt/ccy001

Daly, N. P. (2022). Investigating learner autonomy and vocabulary learning efficiency with MALL. *Language Learning & Technology*, *26*(1), 1-30. https://doi.org/10.64152/10125/73469

Dickinson, D. K., Collins, M. F., Nesbitt, K., Toub, T. S., Hassinger-Das, B., Hadley, E. B., Hirsh-Pasek, K., & Golinkoff, R. M. (2019). Effects of teacher-delivered book reading and play on vocabulary learning and self-regulation among low-income preschool children. *Journal of Cognition and Development*, *20*(2), 136-164. https://doi.org/10.1080/15248372.2018.1483373

Dickinson, D. K., Nesbitt, K. T., & Hofer, K. G. (2019). Effects of language on initial reading: Direct and indirect associations between code and language from preschool to first grade. *Early Childhood Research Quarterly*, *49*, 122-137. https://doi.org/10.1016/j.ecresq.2019.04.005

Dong, J., Zhao, Y., & Buckingham, L. (2024). Thirty years of writing assessment: A bibliometric analysis of research trends and future directions. *Assessing Writing*, *61*, 1-16. https://doi.org/10.1016/j.asw.2024.100862

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

Dwyer, J., & Schachter, R. E. (2020). Going beyond defining: Preschool educators' use of knowledge in their pedagogical reasoning about vocabulary instruction. *Dyslexia*, *26*(2), 173-199. https://doi.org/10.1002/dys.1637

Elgort, I., Candry, S., Boutorwick, T. J., Eyckmans, J., & Brysbaert, M. (2018). Contextual word learning with form-focused and meaning-focused elaboration. *Applied Linguistics*, *39*(5), 646-667. https://doi.org/10.1093/applin/amw029

Fan, M. Y. (2003). Frequency of use, perceived usefulness, and actual usefulness of second language vocabulary strategies: A study of Hong Kong learners. *The Modern Language Journal*, *87*(2), 222-241. https://doi.org/10.1111/1540-4781.00187

Fan, N. (2020). Strategy use in second language vocabulary learning and its relationships with the breadth and depth of vocabulary knowledge: A structural equation modelling study. *Frontiers in Psychology*, *11*, 1-14. https://doi.org/10.3389/fpsyg.2020.00752

Fang-Chuan Ou, Y., & Wen-Chi Vivian, W. (2015). Using mixed-modality learning strategies via e-learning for second language vocabulary acquisition. *Journal of Educational Technology & Society*, *18*(3), 309-322. http://www.jstor.org/stable/jeductechsoci.18.3.309

Gan, H., Zhou, S., Thomas, N., & Zhang, D. (2025). Vocabulary learning through online learning communities: University students' acceptance of technology and their self-regulated strategy use. *Innovation in Language Learning and Teaching*, Advance online publication. https://doi.org/10.1080/17501229.2025.2501611

García Botero, G., Questier, F., & Zhu, C. (2019). Self-directed language learning in a mobile-assisted, out-of-class context: Do students walk the talk? *Computer Assisted Language Learning*, *32*(1-2), 71-97. https://doi.org/10.1080/09588221.2018.1485707

Gay, F. (2022). Investigating the effects of technology-enhanced vocabulary learning strategy instruction on supporting mixed-ability EMI learners studying a journalism and communication major: An action research project at a university in China. *Journal of English for Academic Purposes*, *55*, 1-14. https://doi.org/10.1016/j.jeap.2021.101068

González-fernández, B., & Schmitt, N. (2020). Word knowledge: Exploring the relationships and order of acquisition of vocabulary knowledge components. *Applied Linguistics*, *41*(4), 481-505. https://doi.org/10.1093/applin/amy057

Goossens, N. A. M. C., Camp, G., Verkoeijen, P. P. J. L., Tabbers, H. K., Bouwmeester, S., & Zwaan, R. A. (2016). Distributed practice and retrieval practice in primary school vocabulary learning: A multi-classroom study. *Applied Cognitive Psychology*, *30*(5), 700-712. https://doi.org/10.1002/acp.3245

Goulden, R., Nation, P., & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics*, *11*(4), 341-363. https://doi.org/10.1093/applin/11.4.341

Gu, P. Y. (2018). Validation of an online questionnaire of vocabulary learning strategies for ESL learners. *Studies in Second Language Learning and Teaching*, *8*(2), 325-350. https://doi.org/10.14746/ssllt.2018.8.2.7

Gu, Y., & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning*, *46*(4), 643-679. https://doi.org/10.1111/j.1467-1770.1996.tb01355.x

Gu, Y. (2002). Gender, academic major, and vocabulary learning strategies of Chinese EFL learners. *RELC Journal*, *33*(1), 35-54. https://doi.org/10.1177/003368820203300102

Guo, H., & Li, Z. (2022). An analysis of the learning effects and differences of college students using English vocabulary APP. *Sustainability*, *14*(15), 1-12. https://doi.org/10.3390/su14159240

Guo, F., Zhang, Y., & Wu, Z. (2022). Mobile-assisted vocabulary learning through the Shanbay App outside the classroom: Effects of self-regulation and peer scaffolding. *Frontiers in Psychology*, *13*, 1-16. https://doi.org/10.3389/fpsyg.2022.993224

Hadley, E. B., Dedrick, R. F., Dickinson, D. K., Kim, E., Hirsh-Pasek, K., & Golinkoff, R. M. (2021). Exploring the relations between child and word characteristics and preschoolers' word-learning. *Journal of Applied Developmental Psychology*, *77*, 1-12. https://doi.org/10.1016/j.appdev.2021.101332

Hadley, E. B., Dickinson, D. K., Hirsh-Pasek, K., & Golinkoff, R. M. (2019). Building semantic networks: The impact of a vocabulary intervention on preschoolers' depth of word knowledge. *Reading Research Quarterly*, *54*(1), 41-61. https://doi.org/10.1002/rrq.225

Hadley, E. B., Dickinson, D. K., Hirsh-Pasek, K., Golinkoff, R. M., & Nesbitt, K. T. (2016). Examining the acquisition of vocabulary knowledge depth among preschool students. *Reading Research Quarterly*, *51*(2), 181-198. https://doi.org/10.1002/rrq.130

Hanssen, T.-E. S., Jørgensen, F., & Larsen, B. (2018). The relation between the quality of research, researchers' experience, and their academic environment. *Scientometrics*, *114*(3), 933-950. https://doi.org/10.1007/s11192-017-2580-y

Hao, T., Wang, Z., & Ardasheva, Y. (2021). Technology-assisted vocabulary learning for EFL learners: A meta-analysis. *Journal of Research on Educational Effectiveness*, *14*(3), 645-667. https://doi.org/10.1080/19345747.2021.1917028

Hindman, A. H., Wasik, B. A., & Bradley, D. E. (2019). How classroom conversations unfold: Exploring teacher-child exchanges during shared book reading. *Early Education and Development*, *30*(4), 478-495. https://doi.org/10.1080/10409289.2018.1556009

Hong, J. C., Hwang, M. Y., Chang, H. W., Tai, K. H., Kuo, Y. C., & Tsai, Y. H. (2015). Internet cognitive failure and fatigue relevant to learners' self-regulation and learning progress in English vocabulary with a calibration scheme. *Journal of Computer Assisted Learning*, *31*(5), 450-461. https://doi.org/10.1111/jcal.12103

Hsu, T. C., Chang, C., & Jen, T. H. (2024). Artificial Intelligence image recognition using self-regulation learning strategies: Effects on vocabulary acquisition, learning anxiety, and learning behaviors of English language learners. *Interactive Learning Environments*, *32*(6), 3060-3078. https://doi.org/10.1080/10494820.2023.2165508

Huang, F., Teo, T., & Guo, J. (2021). Understanding English teachers' non-volitional use of online teaching: A Chinese study. *System*, *101*, 1-11. https://doi.org/10.1016/j.system.2021.102574

Huang, Y. M., Huang, Y. M., Huang, S. H., & Lin, Y. T. (2012). A ubiquitous English vocabulary learning system: Evidence of active/passive attitudes vs. usefulness/ease-of-use. *Computers & Education*, *58*(1), 273-282. https://doi.org/10.1016/j.compedu.2011.08.008

Hwang, G. J., & Fu, Q. K. (2019). Trends in the research design and application of mobile language learning: A review of 2007–2016 publications in selected SSCI journals. *Interactive Learning Environments*, *27*(4), 567-581. https://doi.org/10.1080/10494820.2018.1486861

Hwang, G. J., Hsu, T. C., Lai, C. L., & Hsueh, C. J. (2017). Interaction of problem-based gaming and learning anxiety in language students' English listening performance and progressive behavioral patterns. *Computers & Education*, *106*, 26-42. https://doi.org/10.1016/j.compedu.2016.11.010

Kakihana, S. (2024). Does the word-writing strategy in vocabulary learning increase learning outcomes or mislead learners into overestimation? *International Journal of Applied Linguistics*, *34*(3), 936-951. https://doi.org/10.1111/ijal.12544

Kleinberg, J. (2003). Bursty and hierarchical structure in streams. *Data Mining and Knowledge Discovery*, *7*(4), 373-397. https://doi.org/10.1023/A:1024940629314

Kojic-Sabo, I., & Lightbown, P. M. (1999). Students' approaches to vocabulary learning and their relationship to success. *The Modern Language Journal*, *83*(2), 176-192. https://doi.org/10.1111/0026-7902.00014

LaBontee, R. (2019). Questionnaire instrumentation for strategic vocabulary learning in the Swedish as a second language learning context. *Studies in Second Language Learning and Teaching*, *9*(2), 313-348. https://doi.org/10.14746/ssllt.2019.9.2.4

Larivière, V., Archambault, É., Gingras, Y., & Vignola-Gagné, É. (2006). The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities. *Journal of the American Society for Information Science and Technology*, *57*(8), 997-1004. https://doi.org/10.1002/asi.20349

Larsen, P. O., & von Ins, M. (2010). The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index. *Scientometrics*, *84*(3), 575-603. https://doi.org/10.1007/s11192-010-0202-z

Laufer, B. (1990). Ease and difficulty in vocabulary learning: Some teaching implications. *Foreign Language Annals*, *23*(2), 147-155. https://doi.org/10.1111/j.1944-9720.1990.tb00355.x

Lee, J. H., Ahn, J. J., & Lee, H. (2022). The role of motivation and vocabulary learning strategies in L2 vocabulary knowledge: A structural equation modeling analysis. *Studies in Second Language Learning and Teaching*, *12*(3), 435-458. https://doi.org/10.14746/ssllt.2022.12.3.5

Lee, S. (2020). Examining the roles of aptitude, motivation, strategy use, language processing experience, and gender in the development of the breadth and depth of EFL learners' vocabulary knowledge. *SAGE Open*, *10*(4), 1-15. https://doi.org/10.1177/2158244020977883

Lei, X., Fathi, J., Noorbakhsh, S., & Rahimi, M. (2022). The Impact of mobile-assisted language learning on English as a foreign language learners' vocabulary learning attitudes and self-regulatory capacity. *Frontiers in Psychology*, *13*, 1-14. https://doi.org/10.3389/fpsyg.2022.872922

Li, J., & Wang, C. (2024). Mediating roles of motivational beliefs and vocabulary learning strategies for the relationship between self-regulation and vocabulary proficiency. *European Journal of Education*, *59*(4), 1-16. https://doi.org/10.1111/ejed.12706

Li, R. (2021). Does game-based vocabulary learning APP influence Chinese EFL learners' vocabulary achievement, motivation, and self-confidence? *SAGE Open*, *11*(1), 1-12. https://doi.org/10.1177/21582440211003092

Lin, J. J., & Lin, H. (2019). Mobile-assisted ESL/EFL vocabulary learning: A systematic review and meta-analysis. *Computer Assisted Language Learning*, *32*(8), 878-919. https://doi.org/10.1080/09588221.2018.1541359

Liu, P. L., & Chen, C. J. (2023). Using an AI-Based object detection translation application for English vocabulary learning. *Educational Technology & Society*, *26*(3), 5-20. https://www.jstor.org/stable/48734318

Lu, C., & Gu, M. M. (2024). Review of research on digital translanguaging among teachers and students: A visual analysis through CiteSpace. *System*, *123*, 1-14. https://doi.org/10.1016/j.system.2024.103314

Mahdi, H. S. (2018). Effectiveness of mobile devices on vocabulary learning: A meta-analysis. *Journal of Educational Computing Research*, *56*(1), 134-154. https://doi.org/10.1177/0735633117698826

Mangold, M., Lenhard, W., Schulz, D., Schindler, J., & Weiner-Bühler, J. C. (2025). Bilingual advantage in vocabulary development: A longitudinal look at Turkish-German preschoolers. *Language Learning and Development*, *21*(2), 193-213. https://doi.org/10.1080/15475441.2024.2429033

Manzo, A. V., & Sherk, J. K. (1971). Critical perspectives in reading: Some generalizations and strategies for guiding vocabulary learning. *Journal of Reading Behavior*, *4*(1), 78-89. https://doi.org/10.1080/10862967109546978

Mizumoto, A., & Takeuchi, O. (2009). Examining the effectiveness of explicit instruction of vocabulary learning strategies with Japanese EFL university students. *Language Teaching Research*, *13*(4), 425-449. https://doi.org/10.1177/1362168809341511

Mizumoto, A., & Takeuchi, O. (2012). Adaptation and validation of self-regulating capacity in vocabulary learning scale. *Applied Linguistics*, *33*(1), 83-91. https://doi.org/10.1093/applin/amr044

Mujtaba, S. M., Kamyabi Gol, A., & Parkash, R. (2021). A study on the relationship between language aptitude, vocabulary size, working memory, and L2 writing accuracy. *Foreign Language Annals*, *54*(4), 1059-1081. https://doi.org/10.1111/flan.12584

Murphy, D. H., Hargis, M. B., & Castel, A. D. (2023). Younger and older adults' strategic use of associative memory and metacognitive control when learning foreign vocabulary words of varying importance. *Psychology and Aging*, *38*(2), 103-116. https://doi.org/10.1037/pag0000730

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

Nation, I. S. P., & Webb, S. A. (2011). *Researching and analyzing vocabulary*. Cengage Learning.

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, *103*(23), 8577-8582. https://doi.org/10.1073/pnas.0601602103

Ou-Yang, F. C., & Wu, W. C. V. (2016). Using mixed-modality vocabulary learning on mobile devices: Design and evaluation. *Journal of Educational Computing Research*, *54*(8), 1043-1069. https://doi.org/10.1177/0735633116648170

Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Newbury House.

Oxford, R. L., & Scarcella, R. C. (1994). Second language vocabulary learning among adults: State of the art in vocabulary instruction. *System*, *22*(2), 231-243. https://doi.org/10.1016/0346-251X(94)90059-0

Pavii Taka, V. (2008). *Vocabulary learning strategies and foreign language acquisition*. Multilingual Matters.

Pressley, M., Levin, J. R., Kuiper, N. A., Bryant, S. L., & Michener, S. (1982). Mnemonic versus nonmnemonic vocabulary-learning strategies: Additional comparisons. *Journal of Educational Psychology*, *74*(5), 693–707. https://doi.org/10.1037/0022-0663.74.5.693

Pressley, M., Levin, J. R., & Miller, G. E. (1982). The keyword method compared to alternative vocabulary-learning strategies. *Contemporary Educational Psychology*, *7*(1), 50-60. https://doi.org/10.1016/0361-476X(82)90007-8

Puimège, E., & Peters, E. (2019). Learners' English vocabulary knowledge prior to formal instruction: The role of learner-related and word-related variables. *Language Learning*, *69*(4), 943-977. https://doi.org/10.1111/lang.12364

Rahimi, M., & Allahyari, A. (2019). Effects of multimedia learning combined with strategy-based instruction on vocabulary learning and strategy use. *SAGE Open*, *9*(2), 1-14. https://doi.org/10.1177/2158244019844081

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, *140*(6), 1432-1463. https://doi.org/10.1037/a0037559

Rueda-Sánchez, M. I., García-Gómez, L., & González-Fernández, L. (2025). Effects of depth vocabulary, word reading and metalinguistic competence on reading comprehension in Spanish. *Educational Psychology*, *45*(5), 495-519. https://doi.org/10.1080/01443410.2025.2486323

Sandberg, J., Maris, M., & Hoogendoorn, P. (2014). The added value of a gaming context and intelligent adaptation for a mobile learning application for vocabulary learning. *Computers & Education*, *76*, 119-130. https://doi.org/10.1016/j.compedu.2014.03.006

Sarah Hsueh-Jui, L., Yu-Ju, L., & Cloudia Ya-Yu, H. (2014). Exploring the relationship between self-regulated vocabulary learning and web-based collaboration. *Journal of Educational Technology & Society*, *17*(4), 404-419. http://www.jstor.org/stable/jeductechsoci.17.4.404

Shamsan, M. A. A., Ali, J. K. M., & Hezam, T. A. (2021). Online Learning amid COVID-19 pandemic: A case study of vocabulary learning strategies. *Arab World English Journal*, *SI* (1), 281-294. https://dx.doi.org/10.24093/awej/covid.21

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCartby (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 199-227). Cambridge University Press.

Shi, W. (2025). The application of systematic analytical thinking in teaching vocabulary for English majors. *Acta Psychologica*, *257*, 1-18. https://doi.org/10.1016/j.actpsy.2025.105091

Shi, Y. S., & Tsai, C. Y. (2024). Fostering vocabulary learning: Mind mapping app enhances performances of EFL learners. *Computer Assisted Language Learning*, *37*(4), 634-686. https://doi.org/10.1080/09588221.2022.2052905

Shneider, A. M. (2009). Four stages of a scientific discipline; four types of scientist. *Trends in Biochemical Sciences*, *34*(5), 217-223. https://doi.org/10.1016/j.tibs.2009.02.002

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, *24*(4), 265-269. https://doi.org/10.1002/asi.4630240406

Smith, G. G., Li, M., Drobisz, J., Park, H.-R., Kim, D., & Smith, S. D. (2013). Play games or study? Computer games in eBooks to learn English vocabulary. *Computers & Education*, *69*, 274-286. https://doi.org/10.1016/j.compedu.2013.07.015

Tam, H. W. Y., & Kim, S. A. (2024). Strategies used in learning Chinese vocabulary by adolescent Mandarin-Korean bilinguals in Korea. *International Journal of Multilingualism*, *21*(1), 36-59. https://doi.org/10.1080/14790718.2021.2004152

Teng, M. F. (2022). Exploring awareness of metacognitive knowledge and acquisition of vocabulary knowledge in primary grades: A latent growth curve modelling approach. *Language Awareness*, *31*(4), 470-494. https://doi.org/10.1080/09658416.2021.1972116

Tseng, W. T., & Schmitt, N. (2008). Toward a model of motivated vocabulary learning: A structural equation modelling approach. *Language Learning*, *58*(2), 357-400. https://doi.org/10.1111/j.1467-9922.2008.00444.x

Tseng, W. T., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics*, *27*(1), 78-102. https://doi.org/10.1093/applin/ami046

Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language Learning*, *69*(3), 559-599. https://doi.org/10.1111/lang.12343

Ueno, S., & Takeuchi, O. (2025). Relationship between second language vocabulary knowledge and vocabulary learning strategy use: A meta-analysis of correlational studies. *International Review of Applied Linguistics in Language Teaching*, *63*(2), 1115-1137. https://doi.org/10.1515/iral-2023-0100

van den Broek, G. S. E., Takashima, A., Segers, E., & Verhoeven, L. (2018). Contextual richness and word learning: Context enhances comprehension but retrieval enhances retention. *Language Learning*, *68*(2), 546-585. https://doi.org/10.1111/lang.12285

van den Broek, G. S. E., van Gog, T., Jansen, E., Pleijsant, M., & Kester, L. (2021). Multimedia effects during retrieval practice: Images that reveal the answer reduce vocabulary learning. *Journal of Educational Psychology*, *113*(8), 1587-1608. https://doi.org/10.1037/edu0000499

Wang-Taylor, Y., & Clenton, J. (2022). Exploring the relationship between L2 vocabulary size and academic speaking. *System*, *107*, 1-13. https://doi.org/10.1016/j.system.2022.102822

Wang, I. K. H., & Cohen, A. D. (2021). Investigating learner engagement in strategy instruction focused on vocabulary for academic writing: A case study. *System*, *99*, 1-17. https://doi.org/10.1016/j.system.2021.102501

Wang, I. K. H., & Cohen, A. D. (2023). Self-access strategy instruction for academic writing vocabulary: What learners actually do. *Applied Linguistics*, *44*(6), 976-1009. https://doi.org/10.1093/applin/amac056

Wang, X., & Feng, L. (2025). Examining the influential mechanism of English as a foreign language learners' flow experiences in digital game-based vocabulary learning: Shedding new light on a priori proposed model. *Education Sciences, 15*(2), 1-19. https://www.mdpi.com/2227-7102/15/2/125

Wang, X., Xia, C., Zhao, Q., & Chen, L. (2025). Enhancing second language motivation and facilitating vocabulary acquisition in an EFL classroom through translanguaging practices. *Applied Linguistics Review*, Advance online publication. https://doi.org/doi:10.1515/applirev-2024-0292

Wang, X., Zuo, J., Liu, F., & Sun, Z. (2023). A scientometric review of research trends in language assessment literacy. *Education Sciences*, *13*(2), 1-23. https://doi.org/10.3390/educsci13020190

Webb, S., & Nation, I. S. P. (2017). *How vocabulary is learned*. Oxford University Press.

Wei, C. W., Kao, H. Y., Lu, H. H., & Liu, Y. C. (2018). The effects of competitive gaming scenarios and personalized assistance strategies on English vocabulary learning. *Journal of Educational Technology & Society*, *21*(3), 146-158. http://www.jstor.org/stable/26458514

Wei, W. (2023). Understanding and supporting the use of feedback from mobile applications in the learning of vocabulary among young adolescent learners. *Studies in Educational Evaluation*, *78*, 1-8. https://doi.org/10.1016/j.stueduc.2023.101264

Yi, W., Zhao, Y., & Wang, W. (2025). Four decades of research on multiword expressions: A bibliometric analysis. *Applied Linguistics*, Advance online publication. https://doi.org/10.1093/applin/amae085

Yu, Z. (2018). Differences in serious game-aided and traditional English vocabulary acquisition. *Computers & Education*, *127*, 214-232. https://doi.org/10.1016/j.compedu.2018.07.014

Zhang, R., & Zou, D. (2022). Types, purposes, and effectiveness of state-of-the-art technologies for second and foreign language learning. *Computer Assisted Language Learning*, *35*(4), 696-742. https://doi.org/10.1080/09588221.2020.1744666

Zhang, R., Zou, D., & Cheng, G. (2023). Learner engagement in digital game-based vocabulary learning and its effects on EFL vocabulary development. *System*, *119*, 1-19. https://doi.org/10.1016/j.system.2023.103173

Zhang, R., Zou, D., & Cheng, G. (2024). Self-regulated digital game-based vocabulary learning: Motivation, application of self-regulated learning strategies, EFL vocabulary knowledge development, and their interplay. *Computer Assisted Language Learning*, Advance online publication. https://doi.org/10.1080/09588221.2024.2344555

Zhang, X. (2022). The relationship between lexical use and L2 writing quality: A case of two genres. *International Journal of Applied Linguistics*, *32*(3), 371-396. https://doi.org/10.1111/ijal.12420

Zhang, X., & Lu, X. (2015). The relationship between vocabulary learning strategies and breadth and depth of vocabulary knowledge. *The Modern Language Journal*, *99*(4), 740-753. https://doi.org/10.1111/modl.12277

Ziegler, N. (2015). The predictive value of the self-regulating capacity in vocabulary learning scale. *Applied Linguistics*, *36*(5), 641-647. https://doi.org/10.1093/applin/amv020

Zou, D., Huang, Y., & Xie, H. (2021). Digital game-based vocabulary learning: Where are we and where are we going? *Computer Assisted Language Learning*, *34*(5-6), 751-777. https://doi.org/10.1080/09588221.2019.1640745

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp.400-416
DOI: https://dx.doi.org/10.24093/awej/vol16no4.22

Enhancing EFL Students' Listening Skills with Songs and Films

Lisjeta Thaqi-Jashari
UBT – Higher Education Institution, Kosovo
Corresponding Author: lisjeta.thaqi@ubt-uni.net

Rina Sefiu
University for Business and Technology (UBT), Kosovo
Email: rina.sefiu@ubt-uni.net

Received: 08/19/2025 Accepted:10/24/2025 Published:12/10/2025

Abstract

This study investigated the impact of integrating English-language songs and films into English as a Foreign Language (EFL) instruction on learners' listening comprehension. A quasi-experimental pre-test/post-test control-group design was conducted in two educational contexts in Kosovo: a public primary school (Hasan Prishtina) and a private language institute (Britannica ELT). Experimental groups received multimedia-enhanced lessons featuring songs and film clips with structured pre-, while-, and post-listening tasks, whereas control groups followed a textbook-based syllabus. Pre-test scores confirmed baseline equivalence across groups. Post-test analyses using independent-samples t-tests revealed significant improvements for experimental groups in both settings ($p < .001$), with large to very large effect sizes ($d = 1.12$ and 1.65). The findings support Krashen's Input and Affective-Filter Hypotheses, Vygotskian sociocultural theory, and Mayer's Cognitive Theory of Multimedia Learning, confirming that authentic audiovisual input lowers anxiety, heightens motivation, and strengthens comprehension. Pedagogically, the results demonstrate that integrating songs and films offers an effective, adaptable means of improving listening proficiency across both resource-limited and resource-rich contexts.

Keywords: Authentic input, EFL, films, listening comprehension, motivation, multimedia learning, songs, task-based learning

How to cite: Thaqi-Jashari, L., & Sefiu, R. (2025). Enhancing EFL Students' Listening Skills with Songs and Films. *Arab World English Journal, 16*(4):400-416. https://dx.doi.org/10.24093/awej/vol16no4.22

Introduction

Listening comprehension plays a foundational role in second language acquisition (SLA) because it provides the input upon which other linguistic skills depend. Despite its centrality, listening remains the least systematically taught of the four macro-skills, often marginalized in curricula dominated by grammar, reading, and writing (Vandergrift & Baker, 2015; Goh & Aryadoust, 2020). Traditional listening instruction tends to rely on scripted textbook dialogues and comprehension questions that fail to reflect the speed, variability, and pragmatics of authentic spoken discourse (Vanderplank, 2016). Consequently, learners may perform well on classroom tasks yet struggle with real-world communication.

Recent advances in SLA and multimedia pedagogy suggest that authentic audiovisual input—particularly songs and films—can meaningfully improve listening development (Ajmal & Kumar, 2020; Gloria & Galvez, 2025; Murphey, 2019; Li & Brand, 2020). Songs expose learners to rhythm, stress, and colloquial language while lowering affective barriers through enjoyment and repetition (Engh, 2015). Films offer multimodal exposure to natural speech, gestures, and cultural context, supporting both comprehension and motivation. These resources align with Krashen's (1982) Input and Affective-Filter Hypotheses, Vygotsky's (1978) sociocultural principles, and Mayer's (2020) Cognitive Theory of Multimedia Learning, all emphasizing meaningful input, interaction, and engagement.

However, empirical research remains uneven. Many studies have examined songs or films separately and qualitatively, with limited comparative or statistical evidence. Moreover, few have analyzed outcomes across differing educational environments, such as public versus private institutions, where resource availability and learner exposure diverge. Addressing this gap, the present study investigates whether combining songs and films enhances EFL learners' listening comprehension and whether results are consistent across distinct school contexts in Kosovo.

The study pursues three objectives:

1. To determine the extent to which multimedia-based instruction using songs and films improves listening comprehension compared with textbook-based instruction.
2. To compare the relative effectiveness of the intervention in public and private school contexts.
3. To draw pedagogical implications for integrating authentic multimedia resources within EFL listening curricula.

By triangulating quantitative test results with classroom observations, this research contributes both theoretical and practical insights to the field. It demonstrates that multimedia instruction can bridge the gap between linguistic theory and classroom practice, empowering teachers to transform passive listening into active, meaningful engagement.

While this study is situated within an English as a Foreign Language (EFL) context, it draws on Second Language Acquisition (SLA) theories to interpret how exposure to authentic media such as songs and films facilitates language learning and listening development.

Literature Review

Listening Comprehension in Second Language Acquisition

Listening comprehension constitutes the foundation of language acquisition and communicative competence, yet it has historically been undervalued in pedagogical research. Scholars often refer to listening as the "Cinderella skill" of language teaching because it receives less explicit instructional attention than speaking, reading, or writing (Nunan, 2015; Field, 2019).

However, the ability to process spoken input underpins all other linguistic skills. As Vandergrift and Goh (2012) emphasize, listening provides the necessary input for spoken interaction, vocabulary expansion, and literacy development. Learners who fail to develop adequate listening comprehension often experience frustration and diminished confidence in communicative tasks (Graham & Santos, 2015).

Listening involves far more than the passive reception of sounds. It requires perceptual decoding, syntactic parsing, and the integration of linguistic and contextual cues to construct meaning (Field, 2019). These processes are cognitively demanding, particularly in foreign language contexts where input is limited and rapid. Goh and Aryadoust (2020) note that listening proficiency relies on the coordination of bottom-up and top-down strategies—recognizing phonological details while activating background knowledge and inferencing skills. Thus, effective listening pedagogy must train learners to manage these cognitive demands through explicit strategy instruction, metacognitive awareness, and exposure to authentic input.

Challenges in Teaching and Learning Listening Comprehension

Despite its importance, teaching listening effectively remains a persistent challenge. One major obstacle is the **invisibility of the listening process**: teachers cannot directly observe how learners interpret speech, making it difficult to diagnose comprehension problems or tailor feedback (Field, 2019). Moreover, listening occurs in real time—without the possibility of rewinding or reviewing input—requiring rapid lexical access and inferencing (Rost, 2016). Learners must simultaneously decode, predict, and respond, a cognitive load that can overwhelm limited working memory (Vandergrift & Baker, 2015).

Traditional EFL classrooms often rely on **textbook-based dialogues** that present slow, scripted, and simplified speech. While these may reduce cognitive difficulty, they risk creating an unrealistic representation of authentic communication (Vanderplank, 2016). Artificial materials often omit hesitation, reduced forms, and overlapping speech that characterize natural conversation. Consequently, learners may perform well on textbook exercises yet struggle in real-world listening scenarios.

Another challenge involves the **affective dimension** of listening. Listening anxiety—stemming from fear of misunderstanding—can impede comprehension by heightening cognitive load and reducing attention (MacIntyre & Gregersen, 2012). In many public-school settings, resource constraints further exacerbate these difficulties: limited technology, large class sizes, and inadequate teacher training hinder the use of authentic materials (Gilakjani & Sabouri, 2016). Hence, innovative approaches are needed that balance theoretical rigor with contextual practicality, enabling learners to access rich, motivating input within real classroom constraints.

Songs as Pedagogical Tools in EFL Instruction

Linguistic and Phonological Benefits

Songs represent one of the most versatile forms of authentic input in language education. Linguistically, they expose learners to rhythm, stress, and intonation patterns, enhancing phonological awareness and listening discrimination (Murphey, 2019; Xu & Chen, 2022). Repetition of lyrics reinforces vocabulary and grammatical structures, while melody and rhyme aid memorization. Research indicates that songs can significantly improve lexical retention, pronunciation accuracy, and listening fluency (Engh, 2015; Herrera & Murry, 2016).

Cognitive and Affective Dimensions

From a cognitive perspective, songs engage dual memory channels—melodic and verbal—facilitating deeper encoding of linguistic input (Paivio, 2014; Mayer, 2020). The repetitive, predictable patterns inherent in music allow learners to anticipate meaning, reducing processing strain. Affectively, songs create positive emotional associations with the language, thereby lowering anxiety and strengthening motivation (Dörnyei & Ushioda, 2021). This motivational boost is particularly valuable for younger learners or those in low-exposure environments, where engagement is critical to persistence (Ajmal & Kumar, 2020).

Empirical Evidence

Recent empirical studies affirm the pedagogical value of song-based instruction. Bokiev et al. (2018) demonstrated significant improvements in Uzbek EFL learners' pronunciation and listening comprehension following song-based lessons. Similarly, Bao and Lam (2020) found that popular music increased learner participation and comprehension accuracy in Hong Kong classrooms. A meta-analysis by Zhang and Wang (2023) further confirmed that musical input correlates strongly with enhanced listening outcomes, especially when coupled with active tasks such as lyric analysis or role-play. However, these studies often isolate songs from broader multimedia contexts, leaving open the question of whether combined modalities might yield greater gains.

Films as Sources of Authentic Input

Linguistic and Cultural Exposure

Films provide learners with authentic spoken discourse embedded in visual and cultural contexts. They present natural speech phenomena—hesitation, intonation shifts, idiomatic expressions, and overlapping dialogue—that textbooks rarely capture (Wagner, 2010; Bahrani & Sim, 2019). By depicting real-life communicative situations, films enable learners to link linguistic form with pragmatic function, enhancing contextualized understanding (Li & Brand, 2020). Exposure to diverse accents and speech registers also promotes auditory flexibility, an essential skill for global English users.

Motivation and Engagement

Cinematographic storytelling inherently engages emotions and curiosity, making films powerful motivational tools (Rodríguez & Gutiérrez, 2021). Learners become invested in narratives, which increases attention and comprehension. Studies by Kusumaningrum (2015) and Martínez-Sierra (2016) reveal that learners who viewed subtitled films demonstrated notable improvements in vocabulary acquisition and listening comprehension. However, passive viewing alone offers limited pedagogical value; as Champoux (1999) and Vanderplank (2016) caution, films must be accompanied by structured tasks such as prediction, note-taking, discussion, or role-play to transform entertainment into learning.

Empirical Findings

Empirical evidence confirms the substantial benefits of film-based instruction. Li and Brand (2020) reported significant gains in listening proficiency among Chinese EFL students following a 12-week film-based program. Similarly, Bozorgian and Alamdari (2018) demonstrated that integrating multimedia films with metacognitive strategy instruction improved listening outcomes

and learner autonomy. The combined linguistic and visual elements of film enhance comprehension through contextual scaffolding and memory reinforcement (Mayer, 2020).

Integrating Songs and Films: A Multimedia Synergy

While songs and films have been widely studied in isolation, fewer researchers have explored their **combined pedagogical potential**. According to Mayer's (2020) Cognitive Theory of Multimedia Learning, learning improves when information is processed through multiple channels—auditory, visual, and kinesthetic—without overloading cognitive capacity. Films provide narrative context, nonverbal cues, and pragmatic content, while songs engage emotion, rhythm, and repetition. When integrated, these modalities reinforce comprehension through complementary mechanisms (Murphey, 2019; Zhang & Wang, 2023).

Recent findings suggest that combining songs and films can enhance retention, pronunciation, and inferencing skills more effectively than using either medium alone (Gloria & Galvez, 2025). For instance, learners who alternated between song-based and film-based listening tasks demonstrated higher post-test gains in comprehension and motivation than those who engaged with a single medium. This supports the principle of multimodal redundancy—multiple cues reinforcing meaning—proposed in multimedia learning theory (Mayer, 2020).

Moreover, such integration aligns with current trends in EFL pedagogy emphasizing **task-based and experiential learning**. By engaging learners in activities such as lyric interpretation, film-scene reenactments, and comparative discussions, teachers foster active processing and collaborative learning. These practices transform listening from a receptive activity into an interactive, meaning-making process.

Theoretical Framework

Krashen's Input and Affective Filter Hypotheses

Krashen (1982) posited that language acquisition depends on exposure to comprehensible input (i+1) delivered in low-anxiety environments. Songs and films fulfill these criteria by offering rich, authentic input that remains accessible through contextual and visual cues. The emotional appeal of music and narrative also lowers affective barriers, encouraging sustained engagement (Ajmal & Kumar, 2020). In this study, learners' positive reactions and heightened motivation suggest that multimedia instruction effectively reduced affective filters, facilitating input processing and acquisition.

Sociocultural Theory

Vygotsky's (1978) Sociocultural Theory emphasizes learning as a socially mediated process. Cultural artifacts—such as films and songs—serve as mediational tools through which learners co-construct meaning (Lantolf & Thorne, 2015). Classroom interactions around audiovisual materials create scaffolding opportunities within learners' zones of proximal development (ZPD). Through peer collaboration, role-play, and interpretive discussion, learners internalize both linguistic and cultural knowledge.

Cognitive Theory of Multimedia Learning

Mayer's (2020) Cognitive Theory of Multimedia Learning provides a cognitive rationale for integrating audiovisual input. According to this model, learners process verbal and visual

information through separate channels, which—when balanced—enhance retention and understanding. Films combine verbal dialogue with visual context, while songs combine melody with linguistic rhythm, both supporting dual-channel learning. This framework explains the large effect sizes observed in studies where learners engage with multimedia stimuli (Li & Brand, 2020; Zhang & Wang, 2023).

Motivational Frameworks

Motivation remains a decisive factor in language achievement. Dörnyei and Ushioda (2021) highlight emotional engagement, self-efficacy, and intrinsic interest as core motivational drivers. Songs and films promote these by providing enjoyable, meaningful experiences that connect language learning to students' personal lives and cultural interests. Self-Determination Theory (Ryan & Deci, 2020) further clarifies this dynamic: multimedia instruction enhances autonomy, competence, and relatedness, fostering sustained motivation even in challenging listening tasks.

Research Gap and Rationale for the Present Study

Despite extensive theoretical support, empirical evidence remains fragmented. Many studies have investigated the effects of songs or films individually, yet few have combined them within a **quantitative, controlled framework**. Moreover, comparisons between **public and private school contexts** are rare, even though differences in resource access and learner exposure may influence effectiveness. Previous research often lacks longitudinal designs, standardized assessments, or statistical validation, limiting generalizability (Rodríguez & Gutiérrez, 2021).

The present study addresses these gaps through a quasi-experimental design incorporating both **songs and films**, implemented in **two distinct educational settings** in Kosovo. By triangulating quantitative and qualitative data, it contributes to a nuanced understanding of how multimedia input supports listening development across varied institutional contexts. The study situates itself at the intersection of cognitive, affective, and sociocultural perspectives, demonstrating how theoretical principles translate into effective classroom practice.

Methodology

Research Design

This study employed a **quasi-experimental pre-test/post-test control group design**, suitable for classroom-based educational research where randomization across intact groups is not feasible. The design enabled the researchers to evaluate the impact of using songs and films on learners' listening comprehension while maintaining parallel instructional conditions across both treatment and control groups.

The study was conducted in two educational institutions in Kosovo—one public (*Hasan Prishtina School*) and one private (*Britannica ELT Center*)—during a single academic term. Two intact classes were selected in each school: one assigned as the **experimental group** and the other as the **control group**. Each group comprised students aged **13–14 years (Grade 8)**, with 80 students per institution (160 participants total).

The purpose was not to compare public versus private schools but to test the **feasibility and effectiveness** of incorporating authentic audiovisual input (songs and films) within different teaching environments, especially under the limited technological resources typical of public schools.

Participants

Participants' English proficiency levels were assessed prior to the study using a **standardized English placement test** based on CEFR B1 descriptors. This ensured baseline equivalence between experimental and control groups. Random assignment procedures were used within each school to allocate students to groups, minimizing selection bias.

To control for external influences, pre-study surveys assessed students' prior exposure to English-language media (films, songs, online content). The results showed no significant differences in media exposure across groups, supporting internal validity.

Both schools' instructors—experienced EFL teachers with M.A. degrees—received standardized training on lesson procedures and implementation of the instructional treatment. The same teacher taught both the control and experimental groups in each school to minimize teacher-related variability.

Instructional Procedures

Experimental Group

Students in the experimental groups participated in lessons integrating **English-language songs and film clips**. The instructional design followed **Task-Based Language Teaching (TBLT)** principles (Ellis, 2018) and incorporated structured **pre-, while-, and post-listening stages**:

1. **Song and Lyric Warm-ups:**
   Each lesson began with a short song and lyrics activity to activate auditory and linguistic awareness. Learners followed along, analyzed pronunciation, and discussed meaning in context.
2. **Film Viewing:**
   Films and clips (10–45 minutes) were carefully selected to align with syllabus themes and learners' proficiency levels. Short clips focused on vocabulary or idiomatic expressions, while longer films offered extended exposure to authentic dialogue, various accents, and cultural references.
3. **Interactive Tasks:**
   Students engaged in comprehension discussions, vocabulary analysis, grammar exploration, and role-playing activities derived from film scenes or song lyrics. These tasks encouraged both linguistic and cultural learning while reinforcing listening comprehension.

Control Group

Control groups received textbook-based instruction aligned with the national curriculum (*English Plus 3*), without exposure to songs or films. Lessons focused on scripted listening exercises, comprehension questions, and vocabulary drills.

Research Instruments

Listening Comprehension Tests

Listening proficiency was measured through **standardized pre- and post-tests**, adapted from **Cambridge English Preliminary (PET)** and **Oxford Practice Listening Tasks**. Each test contained multiple-choice, short-answer, and open-ended comprehension items, incorporating listening passages from authentic media used in instruction.

The tests assessed comprehension of main ideas, inference, and contextual understanding. To ensure consistency, all items underwent **expert validation** by three EFL specialists who reviewed content alignment and language level.

Reliability and Validity

To establish **reliability**, a **pilot test** was conducted with 30 students from a similar school prior to data collection. The pilot results yielded a **Cronbach's alpha of 0.86**, indicating strong internal consistency. Test difficulty was adjusted based on pilot findings to ensure fairness and progressive challenge throughout the study.

Validity was confirmed through expert judgment and content alignment with CEFR B1 listening descriptors, ensuring that each task measured comprehension rather than general linguistic ability.

Testing Conditions

Testing procedures were standardized across both institutions:

- Tests were administered under identical conditions (quiet classrooms, same timing, clear instructions).
- The same teacher and invigilators oversaw both groups in each school.
- Tests included exposure to **multiple English accents** (British, American, and international) to reflect the linguistic diversity of authentic media.
- Progressive difficulty ensured alignment with the content and learners' developing proficiency.

Scoring combined **objective and subjective evaluation**: multiple-choice items were computer-scored, while open-ended responses were graded using a rubric reviewed by two independent raters to ensure inter-rater reliability.

Data Analysis

Data were analyzed using **IBM SPSS Statistics (Version 23)**.
The following statistical tests were performed:

- **Descriptive statistics** (means, standard deviations) to summarize group performance.
- **Independent-samples t-tests** to compare mean post-test scores between control and experimental groups in each school.
- **Paired-samples t-tests** to assess within-group progress from pre- to post-test.
- **Cohen's d** to measure effect sizes and interpret the magnitude of improvements.

Statistical significance was set at $p < .05$, and all assumptions of normality and homogeneity of variance were verified before conducting inferential tests.

Summary

This study utilized a **quasi-experimental pre-test/post-test control group design** to assess the impact of integrating songs and films on EFL learners' listening comprehension. Reliability was established through pilot testing and Cronbach's alpha, while standardized procedures ensured validity and fairness. Statistical analysis using *t*-tests confirmed the robustness of the findings.

To enhance readability, detailed test item samples, training guidelines, and task descriptions have been relocated to an **Appendix**, ensuring that the main text maintains a clear, concise methodological flow.

Results

This chapter presents the quantitative findings of the study examining the effectiveness of using songs and films to enhance English as a Foreign Language (EFL) learners' listening comprehension. Analyses were performed using the Statistical Package for the Social Sciences (SPSS), version 23.

The results are organized into four main sections: (1) descriptive statistics, (2) between-group comparisons, (3) effect size estimation, and (4) within-group gains. Each section is followed by an interpretation linking the results to the study's research questions and hypotheses.

Descriptive Statistics

Descriptive statistics were computed to summarize the overall distribution of scores across the experimental and control groups in both institutions. Table 1 presents the mean and standard deviation values for pre- and post-tests, highlighting the improvement patterns within and across groups.

Table 1. *Descriptive statistics for pre-test and post-test scores by institution and group*

Institution	Group	N	Pre-Test M (SD)	Post-Test M (SD)	Mean Gain	Improvement (%)
Public School	Experimental	40	51.4 (7.6)	79.3 (6.2)	+27.9	54.2%
Public School	Control	40	50.8 (8.0)	63.8 (7.5)	+12.4	24.4%
Private School	Experimental	40	52.1 (7.9)	84.5 (5.6)	+32.4	62.2%
Private School	Control	40	51.6 (8.1)	66.9 (7.0)	+14.8	28.7%

M = Mean; SD = Standard Deviation. N = 80 per institution (40 per group). Experimental groups were taught through songs and films; control groups followed textbook-based instruction.

The descriptive data reveal that all groups improved from pre-test to post-test. However, the magnitude of improvement was substantially greater among experimental groups, particularly in the private school context. These results suggest that the integration of songs and films positively influenced listening comprehension outcomes.

Between-Group Comparisons

Independent-samples *t*-tests were performed to determine whether the differences between experimental and control groups were statistically significant in each institution.

Table 2. *Independent-samples t-test results for post-test scores by institution*

Institution	*t*(df)	*p*	Cohen's *d*	Effect Size Interpretation
Public School	–5.728 (78)	< .001	1.12	Large
Private School	–6.428 (78)	< .001	1.45	Very large

Significance level set at $p < .05$. Cohen's *d* interpreted as: small (.20), medium (.50), large (.80+).

In both institutions, statistically significant differences were found between the experimental and control groups ($p < .001$). The calculated effect sizes—large and very large—

indicate that the observed improvements were not only statistically significant but also practically meaningful. The consistent pattern across schools demonstrates that multimedia-based instruction is beneficial regardless of institutional type or resource level.

Effect Size Estimation

Effect size calculations provided further evidence of the practical significance of multimedia integration. Table 3 summarizes the magnitude of differences between experimental and control groups.

Table 3. *Summary of effect sizes for experimental vs. control groups*

Institution	Effect Size (*d*)	Interpretation	Confidence Level
Public School	1.12	Large	95%
Private School	1.45	Very Large	95%

Cohen's *d* values above 0.80 are considered large, indicating strong instructional effects.

Large effect sizes across both institutions confirm that the use of songs and films had a substantial and consistent positive influence on learners' listening proficiency. The findings validate that exposure to authentic, multimodal input significantly enhances language processing and comprehension.

Within-Group Analysis

To assess learning gains within each group, paired-samples *t*-tests were conducted comparing pre-test and post-test scores. Table 4 presents the results.

Table 4. *Within-group gains in listening scores (pre-test to post-test)*

Institution	Group	N	Pre-Test M (SD)	Post-Test M (SD)	t (df)	p	Mean Difference	Cohen's d	Interpretation
Public School	Experimental	40	51.4 (7.6)	79.3 (6.2)	–12.23 (39)	< .001	+27.9	1.47	Very large
Public School	Control	40	50.8 (8.0)	63.8 (7.5)	–6.31 (39)	< .001	+12.4	0.81	Large
Private School	Experimental	40	52.1 (7.9)	84.5 (5.6)	–14.06 (39)	< .001	+32.4	1.65	Very large
Private School	Control	40	51.6 (8.1)	66.9 (7.0)	–7.21 (39)	< .001	+14.8	0.90	Large

All groups showed statistically significant improvement ($p < .001$). Experimental groups demonstrated notably larger gains, confirming the efficacy of multimedia instruction using songs and films.

While both experimental and control groups showed progress, the magnitude of gain was considerably greater among students who were exposed to multimedia input. The experimental

groups in both institutions achieved very large effect sizes, signifying robust internal improvement and validating the pedagogical advantage of multimodal listening activities.

Summary of Findings

Across all analyses, the integration of songs and films demonstrated a consistent and powerful effect on EFL learners' listening performance. Statistical evidence indicated:

- Significant improvements in listening comprehension across all groups, with greater gains in the experimental conditions.
- Large to very large effect sizes, highlighting strong practical significance.
- Consistency of outcomes across both public and private educational contexts.

These findings support the study's hypotheses and align with prior research emphasizing the cognitive and affective benefits of authentic audiovisual input (Ajmal & Kumar, 2020; Gloria & Gálvez, 2025; Mayer, 2020).

Discussion

The purpose of this study was to examine the effectiveness of using English-language songs and films to enhance EFL students' listening comprehension. The findings demonstrated significant improvements for experimental groups in both public and private institutions, confirming that **authentic multimedia materials can substantially enhance listening proficiency**. This chapter interprets these outcomes within theoretical and pedagogical frameworks and discusses implications for EFL teaching practice.

The results confirmed that exposure to authentic audiovisual input led to greater improvement in listening comprehension than textbook-based instruction. The large effect sizes (Cohen's d = 1.12–1.65) observed in both educational contexts indicate that learners benefited not only linguistically but also affectively and motivationally.

These results align closely with Krashen's (1982) Input Hypothesis and Affective Filter Hypothesis, suggesting that the rich contextual cues in songs and films provided comprehensible input while reducing anxiety. Students reported that these materials made English "feel natural," consistent with Krashen's claim that low-anxiety environments accelerate acquisition.

Furthermore, the interactive post-viewing and post-listening tasks supported Vygotsky's Sociocultural Theory (1978), which posits that language learning occurs through meaningful social interaction. Collaborative activities such as lyric analysis, discussion, and role-play facilitated co-construction of meaning and scaffolded learners' comprehension within their Zone of Proximal Development (ZPD).

From a cognitive perspective, the findings substantiate Mayer's (2020) Cognitive Theory of Multimedia Learning, which emphasizes dual-channel processing. The combination of visual and auditory input enabled learners to encode linguistic information more efficiently, reinforcing vocabulary retention and contextual understanding.

The motivational dimension, as framed by Dörnyei and Ushioda (2021), was also evident. Students in the experimental groups consistently described lessons as enjoyable and relevant. This intrinsic motivation likely contributed to sustained attention and deeper processing, further explaining the performance gains observed.

Comparison Between Public and Private Institutions

While both experimental groups improved significantly, the private school's group achieved slightly higher post-test scores. This difference may reflect **contextual advantages**—such as smaller class sizes, stable technological access, and fewer logistical interruptions—allowing for smoother integration of multimedia activities.

However, the large effect observed in the public school (d = 1.12) demonstrates that **even limited-resource environments can implement effective multimedia instruction**. Teachers used low-cost projectors, portable speakers, and open-access materials (e.g., YouTube lyric videos, short film clips) to deliver comparable pedagogical benefits.

This suggests that pedagogical innovation, rather than institutional wealth, was the decisive factor. The study thus underscores that **effective EFL listening instruction depends more on methodology than on infrastructure**, aligning with findings by Gloria and Gálvez (2025) and Ajmal and Kumar (2020).

Alignment with Previous Research

The present study's outcomes corroborate previous findings highlighting the positive impact of songs and films on EFL listening development. Engh (2015) emphasized songs as powerful tools for developing phonological awareness and cultural competence, while Li and Brand (2020) confirmed that film-based instruction enhances contextual comprehension and learner motivation.

Recent studies, such as Murphey (2019) and Gloria and Gálvez (2025), further validate the **affective and cognitive synergy** generated by authentic multimedia input. The convergence of current findings with this body of research suggests that the integration of songs and films bridges the gap between controlled classroom input and real-world language use.

Moreover, this study extends the literature by applying these methods to **two distinct institutional contexts**, offering comparative insight into implementation feasibility and equity. The evidence indicates that well-designed multimedia instruction can yield robust gains regardless of resource level—an important contribution to global EFL pedagogy.

Pedagogical Implications

Integrating Multimedia into Curriculum Design

Teachers should intentionally integrate songs and films as **core instructional materials**, not peripheral supplements. Lessons can be structured around authentic listening tasks, such as predicting film scenes, interpreting lyrics, or summarizing dialogues. This approach aligns with **Task-Based Language Teaching (TBLT)** principles and supports active listening, critical thinking, and cultural competence.

Addressing Affective and Motivational Factors

Findings show that songs and films significantly **lower anxiety** and **increase engagement**. Teachers should exploit this affective dimension by allowing student choice in song or film selection, using familiar content, and encouraging reflective discussion. Such strategies reduce the affective filter and promote sustained motivation.

Developing Listening Strategy Awareness

Students benefit from explicit instruction in **listening strategies**—e.g., recognizing key words, predicting meaning from context, and identifying discourse markers. Embedding these strategies

in multimedia lessons enables learners to transfer skills from classroom listening to real-world contexts.

Teacher Training and Resource Adaptation

Effective implementation requires teacher preparation in **selecting, adapting, and sequencing** multimedia materials. Training should emphasize task design, scaffolding, and use of subtitles strategically to match proficiency levels. Even in technology-limited schools, teachers can integrate multimedia through smartphones, shared devices, or offline videos.

Assessment and Feedback

The study's results highlight the importance of **formative assessment** through reflective journals, comprehension logs, and peer feedback. Teachers should assess both linguistic progress and affective engagement to ensure balanced development.

Theoretical Contributions

This research contributes to EFL theory by empirically validating a **multimodal input model** that integrates Krashen's affective principles with cognitive and sociocultural frameworks. It demonstrates how **meaningful, emotionally engaging input** supports both implicit and explicit learning processes.

Furthermore, it offers a context-sensitive extension of **multimedia learning theory** to under-resourced environments, showing that cognitive and affective benefits are achievable even with minimal technology when instructional design is principled and creative.

Limitations and Future Research

While results were strong, several limitations merit acknowledgment:

1. The study covered one academic term; longitudinal research could examine retention over time.
2. The sample was limited to Grade-8 learners; future studies might test different proficiency levels or age groups.
3. Qualitative interviews with teachers could deepen understanding of instructional challenges.
4. Future research should explore additional multimedia forms—such as podcasts, YouTube series, or interactive platforms—to extend the approach.

Despite these limitations, the study provides solid evidence that songs and films foster meaningful, lasting improvement in EFL listening comprehension.

Summary

The findings affirm that integrating songs and films into EFL instruction enhances listening skills through affective engagement, cognitive processing, and authentic exposure. By situating the learner in a rich, low-anxiety environment, multimedia instruction transforms listening from a passive exercise into an interactive, motivating experience.

This pedagogical model bridges theory and classroom practice, offering EFL educators a sustainable and adaptable framework to improve listening comprehension in diverse contexts.

Conclusion

This study explored the effectiveness of integrating English-language songs and films into EFL instruction for Grade-8 learners in Kosovo, examining both public and private educational contexts. Using a quasi-experimental design with pre- and post-tests, the research provided robust statistical evidence that multimedia-based instruction significantly improved students' listening comprehension compared with traditional textbook-based methods. The experimental groups outperformed control groups with large to very large effect sizes (Cohen's d = 1.12–1.65), demonstrating not only linguistic gains but also enhanced confidence, motivation, and engagement.

The results underscore the pedagogical value of authentic audiovisual input as a means of providing comprehensible, contextualized language exposure. Consistent with Krashen's Affective Filter Hypothesis, Mayer's Cognitive Theory of Multimedia Learning, and Vygotsky's Sociocultural Theory, the findings reveal that songs and films lower anxiety, stimulate deeper cognitive processing, and promote collaborative learning. Importantly, the comparable success achieved across both public and private institutions highlights that instructional innovation can outweigh infrastructural limitations when teachers are equipped with well-designed, resource-efficient methods.

Pedagogically, this study advocates for the systematic inclusion of multimedia materials—particularly songs and films—within EFL curricula. Such integration transforms listening from a passive skill into an interactive, emotionally engaging process that strengthens comprehension and communicative competence. The results also emphasize the need for teacher training, learner-centered material selection, and the adoption of formative assessment practices that capture both linguistic and affective progress.

Looking forward, future research should extend this investigation longitudinally and across diverse learner populations to evaluate long-term retention and broader applicability. Expanding multimedia approaches to include digital and interactive formats could further enrich the learning experience and sustain motivation in evolving educational contexts.

In essence, this study demonstrates that songs and films are far more than supplementary teaching tools—they are transformative pedagogical resources that bridge the gap between classroom instruction and real-world communication. By integrating cognitive, affective, and sociocultural dimensions of learning, multimedia instruction offers a practical, engaging, and equitable framework for improving EFL listening proficiency across varied educational settings.

Funding: This research is not funded.
Acknowledgments: Not applicable.
Conflicts of Interest: The author declares no conflict of interest.
Authenticity: This manuscript is an original work.
Artificial Intelligence Statement: During the preparation of this work, the author used Grammarly to improve the readability and language of the manuscript. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the publication's content.

About the Author:

Lisjeta Thaqi Jashari is a lecturer and researcher at UBT – Higher Education Institution, Kosovo. Her research focuses on English language pedagogy, applied linguistics, and the

integration of multimedia tools in EFL learning. She has published and presented in national and international forums on innovative teaching methods.
ORCID ID: https://orcid.org/0009-0007-6876-919X

Rina Sefiu is a lecturer of English at the University for Business and Technology (UBT) in Kosovo. Her work focuses on English language teaching, assessment reform, and digital communication. She is currently pursuing a PhD in English Language and Literature, specializing in data-informed assessment in higher education.
ORCID ID: https://orcid.org/0009-0000-2917-3894

References

Adnan, A. (2014). *Improving students' listening ability through movie report. Proceedings of ISelt FBS Universitas Negeri Padang, 2*, 69–75.

Ajmal, M., & Kumar, T. (2020). *The effectiveness of movies in teaching English as a foreign language: A case study from Saudi Arabia. International Journal of English Linguistics, 10* https://doi.org/10.5539/ijel.v10n1p203

Bahrani, T., & Sim, T. S. (2019). *Authentic language input through films and television for teaching English as a foreign language. Mediterranean Journal of Social Sciences, 10*(5), 67–74.

Bao, G., & Lam, S. (2020). *Using popular music to enhance English listening skills among secondary students in Hong Kong. Asian EFL Journal, 24*(6.2), 180–199.

Bokiev, D., Bokiev, U., Lilliati, I., Aralas, D., & Othman, M. (2018). *Utilizing music and songs to promote student engagement in ESL classrooms.* https://doi.org/10.6007/IJARBSS/v8-i12/5015

Bozorgian, H., & Alamdari, E. F. (2018). *Multimedia listening and metacognitive strategy instruction in EFL classrooms. ELT Journal, 72* https://doi.org/10.1093/elt/ccx010

Champoux, J. E. (1999). *Film as a teaching resource. Journal of Management Inquiry, 8* https://doi.org/10.1177/105649269982016

Dörnyei, Z., & Ushioda, E. (2021). *Teaching and researching motivation* https://doi.org/10.4324/9780429488963

Ellis, R. (2018). *Reflections on task-based language teaching.* Multilingual Matters.

Engh, D. (2015). *Why use music in English language learning? A survey of the literature. English Language Teaching,* https://doi.org/10.5539/elt.v6n2p113

Field, J. (2019). *Listening in the language classroom* (2nd ed.). Cambridge University Press.

Gilakjani, A. P., & Sabouri, N. B. (2016). *Learners' listening comprehension difficulties in English language learning: A literature review. English Language Teaching, 9* https://doi.org/10.5539/elt.v9n6p123

Gloria, R. M., & Gálvez, M. C. (2025). *The role of authentic audiovisual input in enhancing EFL learners' listening comprehension. Arab World English Journal (AWEJ), 16* https://dx.doi.org/10.24093/awej/vol16no1.4

Goh, C. C. M., & Aryadoust, V. (2020). *Teaching and assessing listening: Understanding the role of metacognition.* Routledge.

Graham, S., & Santos, D. (2015). *Strategies for second language listening: Current scenarios and improved pedagogy.* Palgrave Macmillan.

Herrera, S. G., & Murry, K. G. (2016). *Mastering ESL/EFL methods: Differentiated instruction for culturally and linguistically diverse students.* Pearson.

Hidayat, A. (2013). *The use of songs in teaching students' listening ability. Journal of English and Education, 1*(1), 21–29.

Krashen, S. D. (1982). *Principles and practice in second language acquisition.* Pergamon.

Kusumaningrum, M. A. D. (2015). *Using English movies as an attractive strategy to teach senior high school students English as a foreign language. LLT Journal: A Journal on Language and Language Teaching, 18*(1), 11–18. https://doi.org/10.24071/llt.2015.180102

Lantolf, J. P., & Thorne, S. L. (2015). *Sociocultural theory and the genesis of second language development.* Oxford University Press.

Li, X., & Brand, M. (2020). *Effects of film-based instruction on English listening comprehension and motivation. TESOL Quarterly, 54* https://doi.org/10.1002/tesq.589

Listiyaningsih, T. (2017). *The influence of listening to English songs to improve listening skill in listening class. Academica: Journal of Multidisciplinary Studies, 1*(1), 35–49.

MacIntyre, P. D., & Gregersen, T. (2012). *Affect: The role of motivation and emotion in language learning.* In S. Mercer, S. Ryan, & M. Williams (Eds.), *Psychology for language learning: Insights from research, theory and practice* (pp. 103–118). Palgrave Macmillan.

Mayer, R. E. (2020). *Multimedia learning* https://doi.org/10.1017/9781108859029

Murphey, T. (1990). *The song stuck in my head phenomenon: A melodic "Din" in the LAD? System, 18*(1), 53–64.

Murphey, T. (2019). *Music and song in language learning: An integrated approach.* Oxford University Press.

Nunan, D. (2015). *Teaching English to speakers of other languages: An introduction.* Routledge.

Paivio, A. (2014). *Mind and its evolution: A dual coding theoretical approach.* Psychology Press.

Rodríguez, M., & Gutiérrez, A. (2021). *Motivation and engagement in EFL learning through film-based activities. Language Teaching Research Quarterly, 22*(1), 45–59.

Rost, M. (2016). *Teaching and researching listening* (3rd ed.). Routledge.

Ryan, R. M., & Deci, E. L. (2020). *Self-determination theory: Basic psychological needs in motivation, development, and wellness.* Guilford Press.

Sabinus, K. U., Bunau, E., & Novita, D. (2013). *Increasing students' listening comprehension by using VOA: Special English video.* [Undergraduate thesis, Tanjungpura University].

Saeidi, M., & Ahmadi, H. (2016). *The effects of watching videos in pre-reading on EFL learners' reading comprehension and attitudes. TESOL International Journal, 11*(2), 15–44.

Vandergrift, L. (2004). *Listening to learn or learning to listen? Annual Review of Applied Linguistics, 24*, 3–25.

Vandergrift, L., & Baker, S. (2015). *Learner variables in second language listening comprehension: An exploratory path analysis. Language Learning, 65* https://doi.org/10.1111/lang.12109

Vanderplank, R. (2016). *Second language listening: Theory and practice.* Routledge.

Vygotsky, L. S., & Cole, M. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Wagner, E. (2010). *The effect of the use of video texts on ESL listening test-taker performance. Language Testing, 27*(4), 493–513.

Xu, J., & Chen, Q. (2022). *Phonological awareness and listening comprehension in EFL contexts. System, 108*, 102861.

Zhang, L., & Wang, Y. (2023). *A meta-analysis of multimedia instruction in EFL listening comprehension. System, 112* https://doi.org/10.1016/j.system.2023.102124

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp.417-437
DOI: https://dx.doi.org/10.24093/awej/vol16no4.23

A Comparative Study on the Washback Effects of IELTS and CET-6 Listening Tests: Evidence from Test-Takers

Yu Jiashun
Academy of Language Studies, Universiti Teknologi MARA, Cawangan Pulau Pinang, 13500 Permatang Pauh, Pulau Pinang, Malaysia.
Email: 1466530825@qq.com

Boon Yih Mah
Academy of Language Studies, Universiti Teknologi MARA, Cawangan Pulau Pinang, 13500 Permatang Pauh, Pulau Pinang, Malaysia.
Corresponding Author: mahboonyih@gmail.com

Haida Umiera Hashim
Academy of Language Studies, Universiti Teknologi MARA Shah Alam, 40450 Shah Alam, Selangor, Malaysia
Email: haidaumiera@gmail.com

Received:08/19/2025 Accepted: 11/09/2025 Published:12/10/2025

Abstract
In the context of Chinese language assessment research, the washback effects of listening tests remain an underexplored area. Considering that the Chinese education system is centered on examinations, major high-stakes tests, such as IELTS and CET-6, have a significant effect on the way Chinese EFL learners engage with English learning. This study aims to investigate the washback effects of IELTS and CET-6 listening tests and identify the differences between them. The significance of this comparison lies in intuitively revealing the differences in washback effects between IELTS and CET-6 listening tests and providing a research basis for further exploring the factors underlying these differences. This comparison will facilitate the future reform of IELTS and CET-6 listening tests, enhancing their positive washback effects on English listening teaching and learning. A total of 42 participants who had completed both IELTS and CET-6 within the previous year participated in this research and completed the questionnaire. The results showed that IELTS Listening had a more positive influence on English listening learning compared to CET-6 Listening. Specifically, in terms of preparation duration, daily practice intensity, the frequency of preparation activities, and behavioral changes, IELTS Listening exhibited more substantial washback effects in promoting students' listening practice. The findings also indicated that most participants acknowledged the positive effects of their IELTS preparation experience, while they either disagreed or were uncertain about the effects of CET-6 Listening. These differences could stem from the motivational disparity and differences in perceived validity caused by the nature of the test design.

Keywords: comparative study, CET-6, IELTS, listening tests, washback effects

How to cite: Jiashun, Y., Mah, B. Y., & Hashim, H. U. (2025). A Comparative Study on the Washback Effects of IELTS and CET-6 Listening Tests: Evidence from Test-Takers. *Arab World English Journal, 16*(4):417-437. https://dx.doi.org/10.24093/awej/vol16no4.23

Introduction

In this globalized world, standardized language assessments now function as gatekeeping instruments (Hamid et al., 2019; Rajendram et al., 2019). Two widely used English competence tests in China are the College English Test (CET) and the International English Language Testing System (IELTS). CET is a national English examination administered by the National Education Examination Authority (NEEA) with unified question design, fee standards, and test organization. The CET has two levels, CET-4 and CET-6, and this research focuses on CET-6 Listening. IELTS (www.ielts.org) is an international English proficiency examination that has served as a key gatekeeper for millions of students worldwide who pursue international education and professional prospects (Liu et al., 2025). It evaluates test-takers' abilities in four skills: reading, listening, speaking, and writing. Each skill has equal weight in the overall score. IELTS has two types: the Academic and the General Training Tests, and these two types of IELTS tests both have the same speaking and listening tests.

CET-6 Listening and IELTS Listening differ in terms of question types, listening materials, number of questions, and the weight of listening scores in the overall score. According to the CET official website (https://cet.neea.edu.cn/), the question type of CET-6 Listening only has multiple choices. CET-6 Listening materials have long conversations, passages, lectures, reports, speeches, with a speaking rate of 140 to 160 words per minute. Test-takers need to answer 25 questions in 30 minutes, and each question is allocated 13 seconds to answer. The score of CET-6 Listening occupies 35% of the total score of CET-6.

Based on the IELTS official website (https://ielts.idp.com), the question types of IELTS Listening include short-answer questions, multiple choice, matching, sentence completion, map labeling, and note/diagram/flow-chart completion. IELTS Listening materials contain four different recordings, which are the conversation in a daily life context between two people, the monologue in a daily life context, the conversation in an educational or training context, and the monologue about an academic topic. Test-takers need to answer forty questions in about 30 minutes. The score of IELTS Listening takes up 25% of the overall score.

CET was launched in 1987 and has now been administered in China for 38 years. It has grown to become the world's largest English proficiency test in terms of the number of test-takers (Liang, 2017). Besides, CET-6 bears practical significance for students' lives, such as the conferment of a bachelor's degree and the job competitiveness in employment. Thus, it is undoubtedly a large-scale and high-stakes test for Chinese college students. IELTS, as a recognized proof of English proficiency, is widely accepted by more than 10,000 educational institutions, employers, professional associations and government departments in more than 140 countries and regions around the world (British Council, 2018). Around four million test-takers sit the IELTS test each year (British Council, 2025). Therefore, these two tests are both high-stakes and large-scale English tests in China.

In today's interconnected and competitive global context, English proficiency has become indispensable for the young generation, especially Chinese citizens, to enhance their competitiveness (Qi et al., 2025). However, in China's English as a Foreign Language (EFL) environment, learners lack sufficient English listening exposure (Yu, 2023), and English education tends to prioritize reading and writing over listening and speaking due to the examination-oriented system (Gu, 2016). This makes improving learners' listening proficiency a key challenge, and high-stakes English tests—whose washback effects shape learning behaviors—are critical to addressing this challenge.

The College English Test (CET-6) and International English Language Testing System (IELTS) are both large-scale, high-stakes assessments with profound impacts on test-takers' lives, and listening is a core component of both. Studying the washback effects of their listening sections (and comparing these effects) could thus inform test optimization and ultimately support EFL learners' listening improvement—yet few studies have focused on this area.

A review of Web of Science, China National Knowledge Infrastructure (CNKI), and Google Scholar by 7 August 2025 reveals a key research gap: Few studies have systematically examined and compared the washback effects of CET-6 and IELTS listening tests—specifically, there is a scarcity of research that applies Bailey's (1996) model (participants, process, products) to analyze the washback mechanisms of these two listening sections and conducts direct comparisons of their washback effects. Against this backdrop, the main aim of this study is to fill this gap by exploring the washback effects of CET-6 and IELTS listening tests from Chinese EFL undergraduates' perspectives, using Bailey's (1996) three-dimensional model (participants, processes, products), and comparing the similarities and differences in their washback effects.

This study holds concise theoretical and practical significance for language testing and EFL education. Theoretically, it fills the gap of lacking direct comparative analysis in the washback effects of IELTS and CET-6 listening tests (a key oversight highlighted in the study's Literature Review), and on this basis, validates Bailey's (1996) washback model in the context of listening sections—further clarifying how test attributes and learner perceptions shape washback intensity. Practically, findings benefit key stakeholders: test designers can draw on IELTS Listening's strengths (e.g., higher validity perception) to inform CET-6 Listening reform; EFL instructors can tailor their teaching strategies to the evidence from test-takers helping students break free from passive exam-oriented preparation and truly enhance their English listening proficiency; Ultimately, the study responds to China's urgent need to improve EFL listening education by harnessing high-stakes tests as levers for positive instructional and learning change.

The following section outlines the specific research objectives and the corresponding research questions that stem from the identified problem statement. These objectives aim to address key areas of inquiry and provide a structured approach to exploring the central issues at hand. Each research question is carefully crafted to guide the investigation and facilitate a deeper understanding of the washback effects of IELTS and CET-6 listening tests.

1. Research Objective 1: To identify the perceptions of Chinese EFL undergraduates about IELTS and CET-6 listening tests.
2. Research Objective 2: To identify the preparation processes of Chinese EFL undergraduates for IELTS and CET-6 listening tests.
3. Research Objective 3: To identify the learning products of Chinese EFL undergraduates after preparing for and taking IELTS and CET-6 listening tests.
4. Research Objective 4: To identify the similarities and differences in washback effects of IELTS and CET-6 listening tests from test-takers' perspectives

1. Research Question 1: What are the perceptions of Chinese EFL undergraduates about IELTS and CET-6 listening tests?
2. Research Question 2: What are the preparation processes of Chinese EFL undergraduates for IELTS and CET-6 listening tests?
3. Research Question 3: What are the learning products of Chinese EFL undergraduates after preparing for and taking IELTS and CET-6 listening tests?

4. Research Question 4: What are the similarities and differences in washback effects of IELTS and CET-6 listening tests from test-takers' perspectives

Literature Review

Definition of Washback

The term "washback" is widely used in applied linguistics related to the effects of testing on teaching and learning (Liang, 2017). The washback is not viewed as a phenomenon and has not been explored in language testing research until the end of the 20th century (Hong, 2017). Pearson (1988) defined washback as the influence of public examinations on the attitude, behaviors and motivation of teachers, learners and parents. Hughes (1989) stated that washback is "the effect of testing on teaching and learning and can be harmful or beneficial" (p.1). Alderson and Wall (1993) defined washback as the influence of tests on teachers and learners to do things that they otherwise would not do without the test. From the definitions of former researchers, only when the changes appear because of the test use, this kind of change in teaching and learning can be considered as the washback of the test.

So far, almost all the researchers have emphasized that the core of washback is the influence of testing on teaching and learning in their definitions of washback. This study will also adopt this core concept and operationalize washback as observable modifications in teaching methodologies, learning strategies, and stakeholder attitudes directly caused by a test's content, format, or consequences.

The Mechanism of the Washback Effect

Bailey (1996) built a mechanism of washback based on Hughes' (1989, 1993) "participants-process-product" trichotomy and the 15 hypotheses of Alderson and Wall (1993). This mechanism, exhibited in Figure 1, takes the "nature of a test" as the starting point: the test's nature first shapes participants' perceptions and attitudes, then influences their processes, and finally leads to specific products. The model not only shows the sequential influence of tests on participants, processes, and products, but also includes the "wash forward" mechanism—where products (e.g., new materials, curriculum reforms) reverse the impact on test design, forming a dynamic interaction loop.

There are five types of participants in Bailey's model: students, teachers, materials writers, curriculum designers (including administrators), and researchers. Each type of participant has products belonging to its category: Bailey divides these products into two types—direct products (e.g., students improved real language proficiency) and ancillary products (e.g., new teaching materials, revised curricula, research findings). Among them, students' actions affect their learning directly, while other participants indirectly affect students' learning through their ancillary products.

Drawing on Bailey's washback model, washback effects of a test should be examined across three dimensions: participants, process, and products. These dimensions constitute a comprehensive transmission mechanism of washback effects, operating sequentially—from the test's nature to participants' perceptions, through participants' specific processes, to the final direct and ancillary products. Since "language learners are the key participants, whose lives are most directly influenced by language testing washback" (Estaji & Tajeddin, 2012, p. 6), this study focused on investigating learners' perceptions. In terms of processes, the research will investigate

learners' test preparation practices and associated behavioral changes. As for products, it will explore the outcomes that learners derive from the tests.

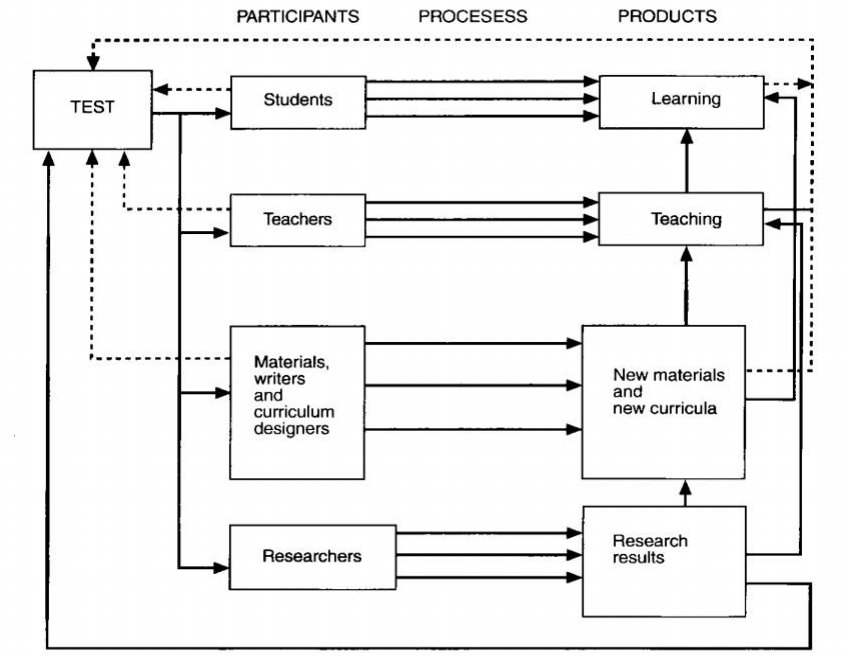


Figure 1. A Basic Model of Washback (Bailey, 1996, p. 264)

Empirical Research

The Washback Research of CET-6 Listening

Previous studies on the washback effects of the CET-6 listening test reveal positive outcomes. Xie (2021) investigated the washback effects of the reformed CET-6 listening test by questionnaires from learners' perspectives. Sixty learners took the survey, and most of them believed that the positive washback effects of the CET-6 listening test outweigh its negative effects. Li (2018) examined 176 non-English major undergraduates after the 2016 reform, identifying both positive and negative washback effects of the CET-6 listening test. On the positive side, the test facilitates improved motivation and vocabulary acquisition through the preparation process for listening sections. Conversely, in terms of negative washback, the rise in CET-6 Listening difficulty brings some inevitable barriers to learners' practice, and consequently, students may rely more on guessing strategies to answer multiple-choice questions.

Wang (2018) surveyed 300 non-English major postgraduates, and most of them believed that the increased weight of the listening component exerts positive washback effects on their English learning. The students reported that they had paid more attention to listening skills due to this change in the CET-6. However, the learners' attitudes toward the cancellation of short conversations are controversial, and more than 50% of learners noted that the difficulty of CET-6 is slightly high. Overall, the reformed CET-6 listening test exerts more positive than negative washback on students.

The Washback Research of IELTS Listening

Research on IELTS Listening reflects positive washback effects over negative washback effects on listening teaching and learning. Ma (2021) noted that the majority of students (62%)

reported improvements in their English listening skills after preparing for and taking IELTS. However, IELTS Listening was also found to induce students' anxiety, which could lead to poor listening performance and undermine their confidence. Wang (2017) indicated that the IELTS Listening component can engage students' focus on the practical application of English. It also encourages both teachers and students to integrate test-driven and ability-driven approaches to teaching and learning.

Additionally, Mi (2008), in a study of 30 Chinese test-takers, identified task diversity—particularly fill-in-the-blank and diagram-labelling items—as a catalyst for contextualised learning strategies and enhanced vocabulary retention. Allen (2016) conducted a longitudinal study with the IELTS test and observed a positive washback effect of IELTS for Japanese test-takers' language ability, especially in speaking and listening areas, for those who had more intense test preparation experience. Despite these positive indicators, methodological limitations, such as small sample sizes (Mi, 2008) and homogeneity in learner proficiency (Allen, 2016), warrant cautious interpretation of results.

Comparative Studies and Critical Gaps

Existing comparative research on high-stakes English tests has shed light on the differential washback effects between IELTS and CET-6, but it remains narrowly confined to non-listening skills, leaving a critical void in understanding how these two dominant assessments shape listening learning. Hong's (2017) investigation of writing tests demonstrated IELTS's stronger positive washback compared to that of CET-6, attributing this disparity to task difficulty and scoring specificity of writing tests. Parallel findings emerged in Liang's (2017) study of speaking tests, where IELTS candidates engaged in substantially more frequent practice sessions than their CET-6 counterparts.

Notably, although Li (2018) demonstrated the value of Bailey's model for single-test listening washback (CET-6 Listening), no research has applied this model to IELTS Listening in a Chinese EFL context—let alone used it to compare the two tests. To address this gap, this study takes Bailey's model as its theoretical core to develop a questionnaire, ensuring the analysis captures the full transmission mechanism of washback effects (from perceptions to processes to outcomes), targeting test-takers who have completed both CET-6 and IELTS, directly comparing the washback effects of the two tests' listening sections.

Methodology

Consistent with Bailey's washback model, Estaji & Tajeddin (2012, p. 6) note that "language learners are the key participants whose lives are most directly influenced by language testing washback." Thus, this study focused on investigating three dimensions of learners: their perceptions, learning processes (test preparation practices and associated behavioral changes), and products (learning outcomes).

Quantitative methods were selected primarily for their strength in collecting numerical data with high accuracy and objectivity (Zyoud et al., 2024). Such data facilitates the systematic identification of trends, patterns, and associations that may not be obvious in non-numerical analyses (Zyoud et al.), thereby enhancing the intuitiveness of comparative analysis. This directly aligns with the study's core goal: comparing the washback effects of the CET-6 and IELTS listening tests from the perspective of test-takers.

The research instrument is a 32-item questionnaire, developed based on Hughes' ideas and Bailey's washback model. The questionnaire investigates the washback across three dimensions: perceptions (motivation, validity, difficulty, time adequacy); preparation process (the test preparation methods; the test preparation time; the frequency of preparation activities; behavioral changes); learning outcomes. A pilot study was conducted to assess the reliability and validity of this questionnaire.

The Likert-scale items (using a 5-point scale, 1–5) demonstrated good reliability, with a Cronbach's alpha coefficient of 0.910. Additionally, the feedback from the pilot study participants indicated that the questionnaire items can effectively capture the information they intended to measure, providing evidence of the content validity.

Participants

A purposive sampling method was used to recruit 42 Chinese undergraduates, all of whom met three inclusion criteria. The first criterion was that participants were aged 18 to 25, ensuring they had similar cognitive abilities. The second criterion is that participants were enrolled in Tier-1 universities in China, which means they are all high academic achievers and have similar learning capabilities. The third selection criterion required the participants to have completed both CET-6 and IELTS tests within the previous year. The sample comprised 29 females (69.05%) and 13 males (30.95%), with only non-English majors included to control for disciplinary bias.

Data Collection and Analysis

The data collection process commenced with a pilot test to assess the reliability and validity of the questionnaire. Additionally, the researcher also sought the opinions and suggestions from the pilot study participants to improve the questionnaire. In the formal data collection process, 42 participants were each given a unique identifier and required to complete the CET-6 Listening washback questionnaire first, followed by the IELTS Listening washback questionnaire after a 24-hour interval. This design ensured the separation of reflections and cognitions regarding the tests and avoided students losing patience with the questionnaire due to boredom and fatigue caused by consecutive questionnaire completion.

This study employed descriptive statistics (including means, medians and standard deviations) to analyse the questionnaire data. No inferential statistics (e.g., t-tests, ANOVA, or correlation analyses) were planned, as the study focused on presenting the core characteristics of washback effects of IELTS and CET-6 listening tests in the sample and providing preliminary references for exploring the washback effects among the college student population, rather than investigating causal or associative hypotheses. All participants were recruited voluntarily, and their personal information was anonymized and protected. The questionnaire data were securely stored on a password-protected device to prevent unauthorized access.

Results

The Perceptions of Learners toward IELTS and CET-6 Listening

The Test-takers' Motivation to Take IELTS and CET-6 Tests

The test-takers' motivation for taking IELTS and CET-6 is presented in Appendix A. For IELTS motivations, 76.19% (32 subjects) cited studying abroad as the primary motivation, followed by promoting English learning (42.86%, n=18). Additionally, proving English proficiency (40.48%, n=17) and enhancing job competitiveness (40.48%, n=17) are also important

reasons to take the IELTS test. University requirements only accounted for 21.43% (n=9). In contrast, CET-6 motivations were dominated by university mandates (47.62%, n=20), proving English ability (54.76%, n=23), and improving job competitiveness (50%, n=21), with only 11.9% (n=5) linking it to studying abroad. Key divergences included a 54.29% higher proportion prioritizing study abroad for IELTS compared to CET-6, and a 26.19% higher rate of university-mandated participation for CET-6.

The Perceptions of Learners toward the Test Difficulty and Validity

For IELTS Listening, 16.67% (n=7) of participants rated difficulty as "quite easy", 40.48% (n=17) as "moderate", and 40.48% (n=17) as "quite difficult", with a mean of 3.286 and a median of 3. For CET-6 Listening, 16.67% (n=7) perceived difficulty as "quite easy", 47.62% (n=20) as "moderate", and 26.19% (n=11) as "quite difficult", yielding a mean of 3.095 and a median of 3. Overall, both tests exhibited comparable central tendencies, with moderate difficulty ratings comprising the largest category for each (40.48% for IELTS vs. 47.62% for CET-6). However, notably, more students perceive IELTS Listening as harder than CET-6 Listening.

Table 1. *The Difficulty of IELTS and CET-6 Listening*

How hard do you think this listening test is?				
	IELTS listening test		CET-6 listening test	
Options	f	%	f	%
Very easy	0	0	2	4.76
Quite easy	7	16.67	7	16.67
Moderate	17	40.48	20	47.62
Quite difficult	17	40.48	11	26.19
Very difficult	1	2.38	2	4.76
Mean	3.286		3.095	
SD	0.774		0.906	
Median	3		3	

Regarding the validity of IELTS and CET-6 listening tests, 64.29% (n=27) agreed IELTS Listening "basically measures real listening ability", with 14.29% (n=6) affirming "total measurement", resulting in a mean of 3.881 and a median of 4. For CET-6 Listening, 47.62% (n=20) expressed doubt about its measurement validity, while 42.86% (n=18) agreed with basic validity, and only 2.38% (n=1) affirmed "total measurement", yielding a mean of 3.381 and a median of 3. IELTS Listening demonstrated greater agreement regarding validity, with a 33.34% higher proportion of affirmative responses compared to CET-6 Listening.

Table 2. *The Validity of IELTS and CET-6 Listening*

Can the listening section of this test measure your real listening level?				
	IELTS listening test		CET-6 listening test	
Options	f	%	f	%
Totally could not	0	0	1	2.38
Basically could not	2	4.76	2	4.76

Not sure	7	16.67	20	47.62
Basically could	27	64.29	18	42.86
Totally could	6	14.29	1	2.38
Mean	3.881		3.381	
SD	0.705		0.731	
Median	4		3	

To sum up, both tests were perceived as moderately difficult, with more participants believing IELTS Listening to be more difficult. Additionally, IELTS Listening exhibited higher validity in measuring real listening ability (mean difference = 0.500) and less doubt about validity compared to CET-6 Listening.

The Perceptions of Learners toward the Question Quantity

Perceptions of question quantity were predominantly positive for both tests, with statistically comparable central tendencies but divergent endorsement rates. For IELTS Listening, 80.9% (34/42) of participants rated the question quantity as "suitable", while 19.05% (8/42) deemed it "quite large", with a mean of 3.19 and a median of 3. For CET-6 Listening, a slightly lower proportion (69.05%, 29/42) regarded the question quantity as "suitable", whereas 23.81% (10/42) deemed it "quite large". The mean was 3.143 and the median 3, both nearly identical to those of IELTS. Despite an 11.90% difference in suitability ratings, both tests exhibited comparable mean and median scores, with a majority of participants agreeing that the question quantity was appropriate.

Table 3. *The Appropriateness of the Number of Listening Questions*

Is the number of listening questions in this test suitable?				
	IELTS listening test		CET-6 listening test	
Options	f	%	f	%
Too small	0	0	1	2.38
Quite small	0	0	2	4.76
Suitable	34	80.95	29	69.05
Quite large	8	19.05	10	23.81
Too large	0	0	0	0
Mean	3.190		3.143	
SD	0.397		0.608	
Median	3		3	

The Perceptions of Learners toward the Listening Score Weighting

Perceptions of listening score weighting received strong endorsement for both tests, with notable differences in high proportions of approval. For IELTS Listening, 92.86% (39/42) of participants rated the score as reasonable, with a mean of 3.024 (median = 3), reflecting broad consensus on its suitability. For CET-6 Listening, a lower proportion (76.19%, 32/42) affirmed the score weighting as reasonable, while 16.67% (7/42) perceived it as disproportionately large; the

mean (3.095) and median (3) were comparable to those of IELTS Listening. Despite a 16.67% gap in high proportions of approval, the majority of participants viewed both tests' listening score weighting as reasonable, with statistically similar mean and median evaluations.

Table 4. *The Suitability of the Listening Score Weighting*

Is the proportion of the listening score in the total score reasonable?				
	IELTS listening test		CET-6 listening test	
Options	f	%	f	%
Too small	0	0	0	0
Quite small	2	4.76	3	7.14
Reasonable	39	92.86	32	76.19
Quite large	1	2.38	7	16.67
Too large	0	0	0	0
Mean	3.024		3.095	
SD	0.269		0.484	
Median	3		3	

The Perceptions of Learners toward the Speaking Speed

For the IELTS listening test, 42.86% (n=18) rated the speaking speed as "quite fast", while 50% (n=21) deemed it "moderate", with a mean of 2.643 and a median of 3. For the CET-6 listening test, 28.57% (n=12) perceived the speaking speed as "quite fast", while 61.9% (n=26) deemed it "moderate", yielding a mean of 2.667 and a median of 3. Statistically, both tests exhibited nearly identical central tendency measures. Very few participants rated the speaking speed as slow for either test, but almost 10% more participants deemed IELTS Listening faster than CET-6 Listening.

Table 5. *The Perceptions of Speaking Speed*

Is the speaking speed in this listening test suitable?				
	IELTS listening test		CET-6 listening test	
Options	f	%	f	%
Too fast	0	0	2	4.76
Quite fast	18	42.86	12	28.57
Moderate	21	50	26	61.9
Quite slow	3	7.14	2	4.76
Too slow	0	0	0	0
Mean	2.643		2.667	
SD	0.618		0.650	
Median	3		3	

The Perceptions of Learners toward Listening Questions' Answer Time

For IELTS Listening, 28.57% (n=12) rated the answer time as "quite tight", 52.38% (n=22) deemed it "moderate", and 11.9% (n=5) assessed it as "quite abundant", with a mean of 2.69 and a median of 3. For CET-6 Listening, 38.1% (n=16) perceived the answer time as "quite tight", while 50% (n=21) assessed it as "reasonable", yielding an identical mean (2.69) and median (3). Both tests exhibited comparable distributions on the perceptions of answer time, with approximately 50% of participants deeming the answer time "moderate".

Table 6.*The Appropriateness of the Listening Questions' Answer Time*

Is the time for answering questions in this listening test enough?				
	IELTS listening test		CET-6 listening test	
Options	f	%	f	%
Extremely tight	3	7.14	1	2.38
Quite tight	12	28.57	16	38.1
Moderate	22	52.38	21	50
Quite abundant	5	11.9	3	7.14
Extremely abundant	0	0	1	2.38
Mean	2.690		2.690	
SD	0.780		0.749	
Median	3		3	

In conclusion, the significant differences in perceptions of IELTS and CET-6 listening tests lie in motivation and validity. In other aspects of the tests, the majority of participants hold similar attitudes toward difficulty, question quantity, score weighting, speaking speed, and answer time. Notably, however, in terms of difficulty and speaking speed, approximately 10% more participants perceived the speaking speed of IELTS Listening to be faster than that of CET-6 Listening, while roughly 12% more viewed IELTS Listening as more difficult than CET-6 Listening. In terms of score weighting, 14.29% (n=6) more participants perceived the score weight of CET-6 to be disproportionately large compared to that of IELTS Listening.

The Preparation Process of Learners for IELTS and CET-6 Listening Tests

The Test Preparation Methods

A survey on test-takers' preparation methods, which focused on whether they enrolled in training courses offered by educational institutions, revealed the degree of importance they attach to the tests. For the IELTS test, 15 participants (35.71%) chose to prepare through educational training institutes, while 27 participants (64.29%) chose to prepare independently. In contrast, 40 participants (95.24%) prepared CET-6 independently rather than through educational training institutes. In summary, significantly more IELTS test-takers enrolled in educational training institutes compared to those preparing for CET-6.

The Test Preparation Time

For IELTS Listening preparation, 33.33% of participants (n=14) spent over three months, 26.19% (n=11) dedicated 2–3 months, 16.67% (n=7) spent 1-2 months, and 16.67% (n=7) spent 0.5-1 month. The distribution exhibited high variability (SD=1.306), with a median score of 4

(corresponding to a 2–3 months preparation period). In contrast, the preparation time of CET-6 exhibited a skewed pattern: 38.1% (n=16) prepared CET-6 for 0.5–1 month, 33.33% (n=14) for less than 0.5 month, and only 11.9% (n=5) for over three months. The median of 2 (0.5–1 month) reflected a shorter typical preparation duration, with comparable dispersion (SD=1.254).

Table 7. *The Preparation Time for IELTS and CET-6*

How long did you prepare for the test?				
	IELTS listening test		CET-6 listening test	
Options	f	%	f	%
Within half a month	3	7.14	14	33.33
Half a month to a month	7	16.67	16	38.1
One to two months	7	16.67	7	16.67
Two to three months	11	26.19	0	0
More than three months	14	33.33	5	11.9
Mean	3.619		2.190	
SD	1.306		1.254	
Median	4		2	

Daily practice intensity differed significantly between IELTS and CET-6 listening practice. For the IELTS listening test, 40.48% (n=17) practiced 1–2 hours daily, 33.33% (n=14) practiced 0.5–1 hours, and 9.52% (n=4) practiced 2–3 hours, with a median of 3 (1–2 hours) and a mean of 2.619. In contrast, for the CET-6 listening test, 61.9% (n=26) practiced less than 0.5 hours daily, while only 16.67% (n=7) practiced 0.5–1 hours, and 16.67% (n=7) practiced 1–2 hours. The median of 1 (less than 0.5 hours) and the mean of 1.643 indicated substantially lower daily listening practice time than those of IELTS test-takers.

Table 8. *Listening Practice Time per Day*

How long did you practice listening every day when you prepared for the test?				
	IELTS listening test		CET-6 listening test	
Options	f	%	f	%
Within 0.5 hours	5	11.9	26	61.9

0.5–1 hour	14	33.33	7	16.67
1–2 hours	17	40.48	7	16.67
2–3 hours	4	9.52	2	4.76
Above 3 hours	2	4.76	0	0
Mean	2.619		1.643	
SD	0.987		0.932	
Median	3		1	

Statistically, IELTS preparation was distinguished by longer durations (median = 4 vs. 2) and more intensive daily practice (median = 3 vs. 1) compared to CET-6 preparation. Specifically, the proportion of participants spending over three months on IELTS was nearly three times that of CET-6 (33.33% vs. 11.9%), while daily practice of 1–2 hours was over twice as common among the IELTS test-takers (40.48% vs. 16.67%).

The Frequency of Test Preparation Activities

The frequency of 10 preparation activities for IELTS and CET-6 Listening is presented in Appendix B. As shown in the table, for IELTS Listening preparation activities, the most frequent was doing past test questions, with a mean of 4.43 and a median of 5. Specifically, 61.9% (n=26) reported “always” practicing with past IELTS Listening questions, while 19.05% (n=8) reported “often” and 19.05% (n=8) reported “sometimes” engaging in this activity. Practicing keyword-catching skills (mean=4.12, median=4) and rapid question-scrutinizing drills (mean=4.02, median=4) were reported with similar frequency. For keyword catching practice, 38.1% (n=16) reported “always” doing it, followed by 35.71% (n=15) who reported “often”. For rapid question scrutiny, 40.48% (n=17) reported “always” practicing it, with 30.95% (n=13) reporting following with “often”. Of the remaining seven preparation activities, all except shadow reading saw at least 50% of participants report “often” or “always” doing the other six activities.

For CET-6 Listening preparation activities, working through past test questions remained the most common activity but at lower frequencies (mean=3.55, median=4) than among IELTS test-takers, with only 30.95% (n=13) reporting “always” doing so. Keyword-catching practice (mean=3.29, median=3) and question-scrutinising drills (mean=3.26, median=3) followed, with 26.19% (11 subjects) always practicing these two activities. Other activities saw a further drop in frequency. Two activities: doing listening simulation questions and shadow reading were reported with a median of 2 (corresponding to “seldom”). In comparison, the remaining five activities were reported with a median of 3 (corresponding to “sometimes”).

To summarise, the frequency of IELTS preparation activities exceeded that of CET-6 preparation activities across all activities, as indicated by the median and mean. Notably, the most frequent activity in both groups was working through past test questions. Specifically, the“always rate for doing past test questions was double in IELTS (61.9%, n=26) compared to CET-6 (30.95%,

n=13). Additionally, the "always" rate for memorising high-frequency vocabulary in IELTS (40.48%, n=17) was even more than twice as high as that in CET-6 (16.67%, n=7). Finally, shadow reading stood out as the least frequent activity among all the preparation activities in both groups.

The Influence of Tests for Learners' Behaviors

This study investigated the impact of IELTS and CET-6 listening tests on students' daily behavior by measuring whether students actively engage in listening to more English songs and watching more English films and TV. In terms of English song listening, 42.86% (n=18) reported increased engagement due to IELTS Listening preparation, compared to 26.19% (n=11) for CET-6. Conversely, a higher proportion of CET-6 test-takers (73.81%, n=31) indicated no change in song-listening behavior versus IELTS takers (57.14%, n=24). For exposure to English films and TV, 69.05% (n=29) increased their consumption of English films and TV for IELTS Listening preparation, while 54.76% (n=23) did so for CET-6, with 45.24% (n=19) and 30.95% (n=13) reporting no impact, respectively. Overall, IELTS Listening preparation was associated with higher rates of increased English song listening (42.86% vs. 26.19%) and film/TV consumption (69.05% vs. 54.76%) than CET-6 preparation, reflecting greater behavioral changes among test-takers.

The Products of IELTS and CET-6 Listening Tests for Learners

The preparation products of IELTS and CET-6 listening tests are presented in Appendix C. For the products of IELTS Listening preparation, all four survey items had a median of 4, indicating that IELTS Listening has achieved positive outcomes. The highest agreement rates were observed for "improved listening proficiency" (42.86% somewhat agree + 33.33% strongly agree = 76.19%) and "increased attention to listening practice"(28.57% somewhat agree + 45.24% strongly agree = 73.81%). The lowest agreement rate was observed "formastering training methods" (19.05% somewhat agree + 40.48% strongly agree = 59.53%), with 35.71% unsure.

Regarding the products of CET-6 Listening preparation, all four items had standard deviations above 1, reflecting high variability, with medians of 3 indicating neutral-to-doubtful views. The most agreed statement was "increased attention to practice" (19.05% somewhat agree + 26.19% strongly agree = 45.24%), though 33.33% reported uncertainty. In contrast, disagreement was notable for the statement "enhance confidence and activeness" (30.95% somewhat disagree + 9.52% strongly disagree = 40.47%), with 28.57% reporting uncertainty. Most participants either disagreed or were uncertain about the effects of the CET-6 listening test on "proficiency improvement": 30.95% expressed disagreement (21.43% somewhat disagree + 9.52% strongly disagree), and 40.48% reported uncertainty. Regarding the mastery of listening training methods, 40.48% expressed uncertainty. Overall, IELTS Listening has achieved more positive outcomes for participants than CET-6 Listening.

Discussion

According to the results of the questionnaire, the washback effects of the IELTS listening test on the preparation process are significantly stronger than those of the CET-6 listening test, thereby eliciting higher agreement rates on its benefits compared to CET-6 Listening. At the same time, most participants have relatively higher rates of disagreement or uncertainty regarding the positive outcomes of CET-6 Listening. Among the participants, more reported enrolling in educational training institutes when preparing for the IELTS than when preparing for the CET-6.

The participants also reported that their IELTS preparation duration and daily practice intensity are significantly longer and greater, respectively, than when preparing for CET-6. Across the ten preparation activities, the frequency of engagement when preparing for IELTS Listening was uniformly higher than when preparing for the CET-6 Listening. Additionally, regarding behavioral influence, the IELTS listening test exerts a greater impact on these participants' choices of English-language daily entertainment, including English songs, films, and TV shows, than the CET-6 listening test.

According to Bailey's washback model (1996), to explain differences in washback effects between IELTS Listening and CET-6 Listening, the direct factors lie in students' perceptions of these two tests. Specifically, learners' motivations and their perceived test validity jointly shape their preparation behaviors. Firstly, there are fundamental differences in the goal-directedness and necessity intensity of learners' motivations. For IELTS, the core motivation is highly concentrated on "Going abroad to study" (76.19%), and proactive learning motivations like "Promote English learning" (42.86%) also have a relatively high proportion. This reflects a high-stakes, proactive drive to master English—rooted in the fact that IELTS scores are not only mandatory for overseas study applications but also tied to precise thresholds (e.g., a band score of 6.5 or 7 for different university programs).

In contrast, CET-6 motivations are more scattered, with instrumentally or passively-oriented needs predominating, such as "Prove my English ability" (54.76%), "Improve my competitiveness in the job market" (50%), and "The requirement of my university" (47.62%). CET-6 functions as a "threshold-type" qualification in domestic contexts. Test-takers generally only need to "pass" (with no rigid demands for a higher or precise score) to obtain the certificate, thereby creating a weaker and more passive impetus for advancing English proficiency.

Secondly, learners' perceptions of test validity further amplify these motivational differences. The majority of students believe IELTS can measure their real listening level, so they recognize preparation as directly contributing to authentic ability development. In contrast, close to half of students doubt CET-6's ability to assess real listening proficiency, reducing the perceived value of intensive preparation. This low perceived validity could stem from the CET-6 listening test's reliance on a single question type: multiple-choice. As Li (2018) noted in his study, test-takers can rely on guessing strategies to answer multiple-choice questions, thereby undermining the validity of the CET-6 listening test. Together, these factors—strong, goal-driven motivation paired with high perceived validity for IELTS, and weaker, instrumentally-oriented motivation combined with lower perceived validity for CET-6—determine learning investment.

For IELTS, the "rigid demand for studying abroad" and trust in its validity drive learners to proactively devote more time and engage in more frequent, purposeful preparation activities (since they see value in developing real skills). For CET-6, the "supplementary/passive needs" and lower perceived validity result in naturally lower preparation duration and practice intensity, as learners feel less urgency to invest heavily in skills they do not believe the test accurately reflects. In summary: The combination of critical goals, specific requirements, and high perceived validity (for IELTS) fosters substantial preparation, while diverse instrumental needs and lower perceived validity (for CET-6) lead to limited investment.

Conclusion

This research investigated the washback effects of IELTS and CET-6 listening tests across three dimensions: perceptions, process, and products, and identified distinct differences in their

washback effects. Overall, IELTS Listening has produced more positive washback effects for participants than CET-6 Listening. Specifically, in terms of the preparation duration, daily practice intensity, the frequency of preparation activities, and behavioral changes, the IELTS Listening has prompted more positive changes in English learning, contributing to its more positive outcomes compared to CET-6 Listening.

These differences may be attributed to participants' perceptions of the tests and variations in test design. Most participants had clearer goals when taking IELTS compared to CET-6 and they believed that IELTS Listening is better able to assess their actual listening proficiency. The reasons that IELTS Listening has more positive washback effects still need to be researched in the future. Notably, this study has limitations, particularly its small sample size. Nevertheless, it provides insights for washback research comparing IELTS and CET-6 listening tests.

Funding: This research is not funded.
Acknowledgments: Not applicable.
Conflicts of Interest: The authors declare no conflict of interest.
Authenticity: This manuscript is an original work
Artificial Intelligence Statement: During the preparation of this work the authors used Doubao in order to improve readability and language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

About the Authors:

Yu Jiashun is a PhD candidate in Applied Language Studies at UiTM Penang. He holds an MA in English Language Education from the University of Reading, UK. His teaching covers English and IELTS in China since 2021. His research focuses on EFL pedagogy, language testing, washback effects, and instructional technology. ORCID ID: https://orcid.org/0009-0008-8727-9185

Boon Yih Mah is a Professional Technologist and Associate Professor serving at UiTM Penang as an academic and Professional Technologist. He holds a PhD in E-learning, an MA in Applied Linguistics, and a BAEd in ELS. Founder of WeCWI and WeCWIIS, he has received international awards for instructional design, language development, and digital learning innovation. ORCID ID: https://orcid.org/0000-0001-9505-6610

Haida Umiera Hashim is a lecturer at UiTM specializing in Technology-Enhanced Language Learning. She earned a PhD in TESL from UKM. Her research highlights inclusive digital pedagogies and EdTech innovations. Co-creator of the award-winning AReal-Vocab AR app, she also contributes as a keynote speaker, reviewer, and interdisciplinary research collaborator. ORCID ID: https://orcid.org/0000-0002-5710-3401

References

Alderson, J., & Wall, D. (1993). Does washback exist? *Applied Linguistics, 2*(14), 115-129.

Allen, D. (2016). Investigating washback to the learner from the IELTS test in the Japanese tertiary context. *Language Testing In Asia, 6*(1), Article 7. https://doi.org/10.1186/s40468-016-0030-z

Bailey, K. (1996). Working for washback: a review of the washback concept in language testing. *Language Testing, 13*(3), 257-279. https://doi.org/10.1177/026553229601300303

British Council. (2018). 2018 Zhongguo dalu diqu yasi kaosheng xueshu biaoxian baipishu [2018 White Paper on the Academic Performance of IELTS Candidates in Mainland China]. https://www.chinaielts.org/sites/all/themes/newyasi/white_book/white_paper_2018.pdf

British Council (2025). *What is IELTS? All You Need to Know About IELTS.* Retrieved July, 14, 2025, from https://takeielts.britishcouncil.org/take-ielts/what-ielts#:~:text=Globally%2C%20there%20are%20more%20than,Education%20and%20Cambridge%20Assessment%20English.

Estaji, M., & Tajeddin, A. (2012). The learner factor in washback context: An empirical study investigating the washback of the IELTS academic writing test. *Language Testing in Asia, 2*(1), 5-25. https://languagetestingasia.springeropen.com/track/pdf/10.1186/2229-0443-2-1-5.pdf

Gu, P. (2016). Zhongwai hezuo banxue moshi xia de yasi kouyu jiaoxue xianzhuang fenxi ji duice yanjiu [Analysis of the current situation and countermeasures of IELTS speaking teaching under the Sino-foreign cooperative education model], *English Teachers 16*(15), 16-18.

Hong, Y. (2017). *A comparative study on the washback of CET-6, IELTS and TOEFL iBT writing tests: Evidence from test-takers' perspectives* [Unpublished master's thesis]. Chongqing University. Retrieved from China National Knowledge Infrastructure.

Hughes, A. (1989). *Testing for language teachers*. Cambridge University Press, UK.

Hughes, A. (1993). Backwash and TOEFL 2000 (Unpublished manuscript). University of Reading.

Hamid, M. O., Hoang, N. T. H., & Kirkpatrick, A. (2019). Language tests, linguistic gatekeeping and global mobility. *Current Issues in Language Planning, 20*(3), 265-282. https://doi.org/10.1080/14664208.2018.1495371

Li, Z. (2018). *A study on washback of the new CET-6 Listening subtest on students'listening learning* [Unpublished master's thesis]. Nanjing University of Aeronautics and Astronautics. Retrieved from China National Knowledge Infrastructure.

Liang, Y. S. (2017). *A comparative washback study of CET-SET6, IELTS and TOEFL iBT speaking tests: Evidence from test-takers' perspectives* [Unpublished master's thesis]. Chongqing University. Retrieved from China National Knowledge Infrastructure.

Liu, M., Hashim, H., & Sulaiman, N. A. (2025). Cognitive Diagnostic Assessment for IELTS Speaking Skills: A Chinese High School Case Study. *Arab World English Journal, 16*(2), 276-310. https://dx.doi.org/10.24093/awej/vol16no2.16

Mi, L. (2008). *The listening subtest of IELTS and its washback effects on English learning* [Unpublished master's thesis]. Northwest University. Retrieved from China National Knowledge Infrastructure.

Ma, H. (2021). *Washback effects of the IELTS test: views and experiences of Chinese students in the context of a Sino-UK joint programme with English as the medium of instruction* [Unpublished doctoral dissertation]. Queen's University Belfast. Retrieved from https://pure.qub.ac.uk/en/studentTheses/washback-effects-of-the-ielts-test-views-and-experiences-of-chine

Pearson, I. (1988). Tests as levers for change. In D. Chamberlain & R. Baumgardner (Eds.), *ESP in the classroom: Practice and evaluation* (pp. 98-107). Oxford: Modern English Publications.

Qi, H., Kamal, S. S.L. A., & Lahlou, H. (2025). Impact of Flipped Classroom on Developing Critical Thinking Skills in Chinese EF Undergraduates' Argumentative Essays: A Pilot Study. *Arab World English Journal, 16*(2), 153-167. https://dx.doi.org/10.24093/awej/vol16no2.9

Rajendram, S., Larson, E., & Sinclair, J. (2019). International Graduate Students' Perspectives on High-Stakes English Tests and the Language Demands of Higher Education. *Language and Literacy, 21*(4), 1-20. https://doi.org/10.20360/LANGANDLIT29428

Wang, X. (2017). On the Evaluation of IELTS listening test and its Washback Effects on Listening Teaching and Learning. *International Journal on Studies in English Language and Literature (IJSELL), 5*(7), 17-23. http://dx.doi.org/10.20431/2347-3134.0507005

Wang, C. L. (2018). *An investigation to the reformed CET6 washback on non-English major postgraduates in Xinjiang University* [Unpublished master's thesis]. Xinjiang University. Retrieved from China National Knowledge Infrastructure.

Xie, J., Lu, J., & Ma, T. (2021). The Washback Effect of Reformed CET-6 Listening Comprehension Test. *Asian Journal of Education and Training, 7*(1), 70-73. DOI: 10.20448/journal.522.2021.71.70.73

Yu, J. S. (2023). An In-depth Discussion on Age Effects in Second Language Learning. *Frontiers in Education Research, 6* (11), 98-104. DOI: 10.25236/FER.2023.061117

Zyoud, M. M., Bsharat, T. R. K., & Dweikat, K. A. (2024). Quantitative Research Methods: Maximizing Benefits, Addressing Limitations, and Advancing Methodological Frontiers. *ISRG Journal of Multidisciplinary Studies (ISRGJMS), 2*(4), 11-14. https://doi.org/10.5281/zenodo.10939470

Appendices

Appendix A

Table 1. *The Motivation of Learners Taking IELTS and CET-6*

	IELTS		CET	
Options	f	%	f	%
Prove my English ability	17	40.48	23	54.76
Go abroad to study	32	76.19	5	11.9
The Employment needs	11	26.19	15	35.71
Improve my competitiveness in the job market	17	40.48	21	50

	IELTS		CET	
Options	f	%	f	%
Promote English learning	18	42.86	11	26.19
Get certificates from the authority institutes	12	28.57	14	33.33
The requirement of my	9	21.43	20	47.62

f: the number of subjects

Appendix B

Table 2. *The Frequency of Test Preparation Activities*

TEST	Options	Never	Seldom	Sometimes	Often	Always	Mean	SD	Median
IELTS	(1) Do past listening test questions	0(0%)	0(0%)	8(19.05%)	8(19.05%)	26(61.9%)	4.43	0.801	5
CET-6		4(9.52%)	5(11.9%)	10(23.81%)	10(23.81%)	13(30.95%)	3.55	1.310	4
IELTS	(2) Practice the keyword catching skills from listening	0(0%)	0(0%)	11(26.19%)	15(35.71%)	16(38.1%)	4.12	0.803	4
CET-6		6(14.29%)	5(11.9%)	13(30.95%)	7(16.67%)	11(26.19%)	3.29	1.367	3
IELTS	(3) Rapid question-scrutinizing drills	2(4.76%)	0(0%)	10(23.81%)	13(30.95%)	17(40.48%)	4.02	1.047	4
CET-6		5(11.9%)	7(16.67%)	13(30.95%)	6(14.29%)	11(26.19%)	3.26	1.345	3
IELTS	(4) Memorize high-frequency listening vocabulary	3(7.14%)	2(4.76%)	10(23.81%)	10(23.81%)	17(40.48%)	3.86	1.221	4
CET-6		7(16.67%)	6(14.29%)	15(35.71%)	7(16.67%)	7(16.67%)	3.02	1.297	3
IELTS	(5) Analyze listening question types and summarise solving strategies.	1(2.38%)	4(9.52%)	14(33.33%)	9(21.43%)	14(33.33%)	3.74	1.106	4
CET-6		8(19.05%)	8(19.05%)	12(28.57%)	7(16.67%)	7(16.67%)	2.93	1.351	3

IELTS	(6) Communicate feelings and experiences of testing with friends and teachers	4(9.52%)	4(9.52%)	9(21.43%)	13(30.95%)	12(28.57%)	3.60	1.270	4
CET-6		9(21.43%)	7(16.67%)	13(30.95%)	6(14.29%)	7(16.67%)	2.88	1.365	3
IELTS	(7) Practice the ability to record key information quickly	3(7.14%)	7(16.67%)	9(21.43%)	11(26.19%)	12(28.57%)	3.52	1.273	4
CET-6		12(28.57%)	8(19.05%)	7(16.67%)	7(16.67%)	8(19.05%)	2.79	1.507	3
IELTS	(8) Listen to English audio, such as English radio and news	5(11.9%)	6(14.29%)	10(23.81%)	10(23.81%)	11(26.19%)	3.38	1.343	3.5
CET-6		9(21.43%)	10(23.81%)	11(26.19%)	8(19.05%)	4(9.52%)	2.71	1.274	3
IELTS	(9) Do listening simulation questions	9(21.43%)	4(9.52%)	7(16.67%)	8(19.05%)	14(33.33%)	3.33	1.557	4
CET-6		11(26.19%)	12(28.57%)	8(19.05%)	8(19.05%)	3(7.14%)	2.52	1.273	2
IELTS	(10) Shadow reading	8(19.05%)	4(9.52%)	11(26.19%)	7(16.67%)	12(28.57%)	3.26	1.466	3
CET-6		15(35.71%)	11(26.19%)	10(23.81%)	4(9.52%)	2(4.76%)	2.21	1.180	2

Appendix C

Table 3. *The Learning Products of IELTS and CET-6 Listening*

Test	Options	1	2	3	4	5	Mean	SD	Median
IELTS	(1) Make me more confident and active in listening to English	0(0%)	1(2.38%)	16(38.1%)	14(33.33%)	11(26.19%)	3.83	0.85	4
CET-6		4(9.52%)	13(30.95%)	12(28.57%)	8(19.05%)	5(11.9%)	2.93	1.18	3
IELTS	(2) Improve my English listening level	0(0%)	0(0%)	10(23.81%)	18(42.86%)	14(33.33%)	4.1	0.76	4
CET-6		4(9.52%)	9(21.43%)	17(40.48%)	9(21.43%)	3(7.14%)	2.95	1.06	3

IELTS	(3) Make me master more listening training methods	1.23	2(4.76%)	15(35.71%)	8(19.05%)	17(40.48%)	3.95	0.99	4
CET-6		3(7.14%)	7(16.67%)	17(40.48%)	9(21.43%)	6(14.29%)	3.19	1.11	3
IELTS	(4) Make me pay more attention to the practice of English listening	0(0%)	2(4.76%)	9(21.43%)	12(28.57%)	19(45.24%)	4.14	0.93	4
CET-6		3(7.14%)	6(14.29%)	14(33.33%)	8(19.05%)	11(26.19%)	3.43		3

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp.438-453
DOI: https://dx.doi.org/10.24093/awej/vol16no4.24

Negotiating Writer Identity: A Comparative Study of Jordanian Undergraduate Students of English in Jordan and America

Amani Salman
School of Literature, Languages and Linguistics
Australian National University, Canberra, ACT, Australia
Email: amani.salman@anu.edu.au

Received:08/18/2025 Accepted: 10/23/2025 Published:12/10/2025

Abstract
This study examines the construction of writer identity among undergraduate English language students, comparing Jordanian students at the University of Jordan with their counterparts at San Diego State University. Through a methodology based on multi-method analysis, the overall importance of the research lies in analysing student essays written in two different educational contexts. Discourse-based interviews complement the tools to achieve the goal of this comparative design, which is to clarify and explore the impact of various educational contexts and international experience on the construction of academic writer identity among English as a second/foreign language students. The study employs a results-based method that examines students' linguistic choices in their academic essays to investigate the impact of the rhetorical features used on the construction of their written identity. The results also reveal subtle differences in writing strategies, as shown by the greater use of first-person pronouns among Jordanian students. On the other hand, San Diego State University students show increased linguistic development, including more structured discussions with clear topic sentences and supporting details. The study concludes that there is a positive and significant relationship between international educational experience and the development of some aspects of academic writer identity.

Keywords: Academic writing, discursive strategies, educational contexts, international experience, writer identity

How to cite: Salman, A. (2025).Negotiating Writer Identity: A Comparative Study of Jordanian Undergraduate Students of English in Jordan and America. *Arab World English Journal, 16*(4):438-45. https://dx.doi.org/10.24093/awej/vol16no4.24

Introduction

Constructing writer identity in academic writing is an essential part of graduate studies. The term *writer identity* designates an author's appearance in the text, either implicitly or explicitly (Aull & Lancaster, 2014). Such an appearance is crucial to the overall quality of the piece of writing (Kim, 2025). However, developing a writer identity comes with challenges for undergraduate students, especially in a second language (L2). Consequently, academic writing does not only involve linguistic competence, but it also represents an act of identity (Ivanič, 1998). Many L2 writers struggle with this issue in their publications (Rahimivand & Kuhi, 2014). The last few decades have witnessed an increasing interest in investigating writer identity construction in academic writing (Ren, 2023).

Several studies have been interested in textual features that can represent writer identity, examining various aspects, from linguistic choices to citation practices in student writing (Crosthwaite et al., 2017). Metadiscourse was an example model widely employed in studies that examine these lexical features, such as (Dafouz-Milne, 2008; Hyland & Tse, 2004; Lee & Park, 2023). They categorised metadiscourse into two types based on the text communicative function: interactional and interactive. While interactive metadiscourse assists in guiding readers throughout the text, interactional metadiscourse engages them and thus is typically employed as a textual formation of authorial identity (Hyland, 2000). Studies examined different kinds of interactional metadiscourse, such as the use of personal pronouns (Hyland, 2018), boosters, attitude markers, and hedges (Al-Dawoody Abdulaal, 2020; Hyland and Tse, 2004). Although writers present a kind of identity in their academic writing, educational contexts might restrict how English as a Second Language (ESL) and English as a Foreign Language (EFL) students establish their writer identities (Ivanič, 1998).

Swales (2014) argued that international students differ from their locally educated counterparts due to the cultural and linguistic differences in their learning contexts, which undermine their sense of identity (Wang & Parr, 2025). To better understand this dynamic, this paper investigates how Jordanian undergraduate students use discursive strategies to construct their writer identity. It also explores the development of academic writer identity in the Jordanian context by comparing two groups: local and international undergraduate students of English. This study aims to investigate how various experiences and educational contexts affect students' discursive strategy choices that convey their identity in academic writing.

Research Questions:

1) What are the similarities and differences in the discursive strategies used by local and international Jordanian students to construct their writer identity across different educational and cultural contexts?
2) How are students' perceptions of themselves as English-language writers reflected in their academic writing?

Literature Review

Cross-Cultural Academic Writing Practices

Writer identity has gained momentum as a research topic in the last few years. The literature tended to treat writer identity as one of a writer's linguistic choices. To study the construction of writer identity, researchers have looked at a variety of linguistic elements (Hamdan & Ahmad, 2023; Li & Deng, 2019). For example, the use of first-person pronouns in academic

writing. According to the literature, the use of first-person pronouns is the most explicit way of expressing writer identity, especially among novice writers. This is because it gives them a sense of self-gratification and engagement (Hyland & Jiang, 2017).

From a distinct perspective, Danis (2022) suggested that the use of first-person pronouns differs in academic writing across different fields. Yue (2024) conducted a cross-disciplinary comparison to identify how cultural conventions and disciplinary-specific communities might influence writers' authorial stance. They found that English Computer Science texts exhibited a strong explicit authorial presence, while Chinese scholars displayed a mitigated authorial persona. In addition, Li (2021) suggested that the use of personal pronouns is subject to institutions' or journal policies. The findings also showed that the use of third-person pronouns entails an imbalance of power between the reader and the writer, rendering texts difficult and disembodied. Solving this depends on engaging students in the meaning-making process by using first-person pronouns (Hyland & Jiang, 2016; Sheldon, 2009).

Norton (2013) investigated the role of social factors in constructing academic identity. She argued that ESL/EFL students' identity construction in academic texts is influenced by their educational experiences, English language proficiency, genre affordances, disciplinary conventions, and audience awareness. As advocates of this view, some research presented writing as a self-representation act, in which writers navigate disciplinary conventions, sociocultural norms and personal experience to construct their identities as members of academic discourse communities (Zhang & Wang, 2024).

From a cross-cultural perspective, Aminifard (2020) investigated how Chinese, Iranian, and Turkish writers project their identities through diverse types of self-representation strategies. The results showed that Iranian and Turkish writers exhibited weaker identities in comparison to their Chinese counterparts. Nevertheless, the writing quality of these authors from the three nationalities was less effective than that of skilled writers in the field. In the same line, Uysal (2014), in a corpus of a hundred and twenty conference papers, explored the influence of cross-cultural differences on the frequency of hedging devices and indirectness in the writing of Japanese, Indian, Turkish (Nonnative speakers) and Anglo-American (Native speakers) authors. The analysis demonstrated that Native and Indian speakers employed fewer hedging devices and indirectness than their Turkish and Japanese counterparts. In their study of the discursive features in the academic texts of English as additional language students (EAL), Candarli et al. (2015) observed that EAL writers established a weak identity compared with experienced writers and their native speaker counterparts.

To investigate the influence of educational context on the construction of students' authorial identity, Leedham and Fernandez-Parra (2017) compared writer identity construction among twenty doctoral students by examining their citation practices. Ten participants were overseas students in English-speaking countries, whereas another group of ten were local students in China. Their findings indicated that students' current educational experience can significantly influence their authorial identity construction. This impact was clear in the higher frequency of using non-integral citations (e.g., indirect citation) by overseas students in comparison to those who are local students in China. Due to the role that non-integrated citations can play in constructing authorial voice, their findings suggested that overseas doctoral students established a stronger writer identity than local students. Several studies supported this claim and argued that there is a positive correlation between the students' educational experiences and the construction of their authorial identities (Matsuda, 2001).

So far, the focus of most research on writer identity among students who are Arabic L1 speakers was on the development of writer identity in their L2 writing as local students studying in their own countries (Matsuda, 2015). However, little research has explored the role of international experience in the development of writer identity. International experience, as used in this research, refers to the exposure of two Jordanian students to a native English context during their studies at San Diego State University (hereafter, SDSU). The role of international experience in developing writer identity is particularly exciting for these students who received training as ESL teachers (Alasmari, 2022).

English Language Learners in Comparative Contexts

Previous literature on Jordanian learners of English in Jordan and abroad has emphasised key issues of adaptation challenges, identity construction, and the effect of language proficiency on academic development. Research on identity revealed that Jordanian students often navigate their intricate linguistic and cultural identities while learning English. Al-Hajaya (2016) studied how local Jordanian learners of English navigate their national identity in relation to the role of global English. The results showed that language learning and identity are deeply connected. Additionally, Jaradat and Bakrin (2017) highlighted the interaction between English teaching and students' identity construction, as well as the impact of English learning strategies on developing proficiency.

In a different context, AlMasaeed (2020) explored how Jordanian international students alter their identities to adapt to the new linguistic environment. The study revealed that participants developed hybrid identities and reported challenges with cultural adaptation. They also expressed and raised concerns about cultural loss. Other studies examined psychological, social, and academic difficulties that overseas Jordanian students encounter. In addition, Abu Rabia (2017) investigated other concerns, such as cultural shock and language hurdles, emphasising the role of institutional resources and social support in assisting effective adjustment. Previous literature also identified language proficiency as a significant predictor of academic performance. For example, in the context of English-based instruction, Martirosyan et al. (2015) observed a positive relationship between high proficiency levels and superior academic outcomes. Almatarneh et al. (2018) also argued that Jordanian students in American universities struggle in both academic writing and spoken communication, pointing to their need for comprehensive language support. Together, all previous studies emphasised the conflict that Jordanian learners experience while developing their English skills across different cultural and educational contexts, focusing on addressing cultural, identity, and linguistic dimensions.

Methods

Participants

The population of the study contained two groups of Jordanian undergraduate students of English. The first group involved two participants who are affiliated with the Department of Linguistics at the School of Foreign Languages in the University of Jordan (henceforth, UJ). This department offers an undergraduate program in English Language that prepares teachers of English for the Jordanian job market. The second group consisted of two Jordanian students who study at SDSU, referred to as "international students" due to their English exposure in an English-speaking context during their studies.

The study employed a Convenience sampling method by selecting the first four available participants. All were undergraduate English majors trained as ESL teachers, with at least two years of academic experience. This procedure ensured prior exposure to academic writing. Three selection criteria guided participation: (1) enrollment in an undergraduate English program, (2) completion of at least two years of study to guarantee familiarity with academic writing, and (3) for the second group, international experience in an English-speaking country. The final criterion aligns with the study's hypothesis that immersion in a native English environment influences students' writer identity (see Table 1).

Table 1. *Participants' profile*

ID	Name	Age Group	Gender	University	Program	Study Level
1	Shadi	23	Male	UJ	Linguistics	Level 3
2	Abeer	22	Female	UJ	Linguistics	Level 3
3	Ahmed	22	Male	SDSU	ESL	Level 6
4	Fatima	22	Female	SDSU	ESL	Level 5

Data Collection and Procedures

The researcher conducted this project at UJ over approximately five weeks, from June 1 to July 15, 2 24. Data collection involved two procedural tools. First, participants completed two research papers to fulfil the requirements of a writing course. Selecting research papers previously submitted during their undergraduate program gave the study a naturalistic feel. In other words, if participants submitted research papers for the sake of the survey itself, their writing would have been more task-oriented and less naturalistic. The submitted papers were representative of their writing performance and thus reflected their writing identities. Given that the students were from different academic institutions and different study programs, it was impossible to assign the same assignments to all. However, to ensure consistency, students in each group submitted papers of the same type. In other words, the locally educated students were from the same program and therefore submitted the same papers, as did the SDSU students.

The papers in the first group were from two different courses. They submitted their first paper during their first semester at the university as part of the "English Writing: Level 1" course, while the second paper was part of the "Advanced Writing" course in their fourth semester of study. The papers submitted by the second group of students were from two subject courses: "English Writing for International Students" and "Principles of Academic Writing." SSDSU offers these courses in the first year of the undergraduate program. In total, there were four different papers and eight samples for analysis. The first paper required students from UJ to write an argumentative paragraph about online dating. In the second paper, students wrote a one-page personal statement for applying to a graduate program at the University of Jordan. The second group submitted two different papers. The first paper required them to write a narrative essay about the various obstacles they faced during their first semester at the university. The second paper was an argumentative synthesis, in which students wrote a compilation of four essays that addressed the "bad behaviour" of the American people. Although the selection of topics was random, the textual content was consistent in terms of genre and purpose.

The second tool of data collection was semi-structured interviews, where the researcher conducted individual, discourse-based interviews with participants. In a discursive style, participants read samples of their writing, with some highlighted parts. Then, the researcher asked them to comment on and explain the highlighted elements. The purpose of this method was to

measure participants' awareness and knowledge of writing conventions. In this study, participants had to answer a series of questions related to specific discursive strategies that reveal the writer's identity. The researcher conducted interviews in Arabic to give participants greater freedom and ease in expressing their opinions, and it was done online using Zoom. The researcher did not record the interviews but took notes. Table 2 presents background information about participants along with the frequency of selected discursive features related to writer identity in their writing samples. These features include first-person pronouns (I_Count), evaluative expressions (Eval_Expr), self-referential language (Self_Ref), and argumentative complexity (Arg_Complexity).

Table 2. *Frequency of writer identity features in writing samples*

Name	ID	Group	I_Count	Eval_Expr	Self_Ref	Arg_Complexity
Shadi	1	UJ	12	5	3	2
Abeer	2	UJ	10	6	4	3
Ahmed	3	SDSU	20	9	6	5
Fatima	4	SDSU	18	10	7	4

Table 2 shows a comparison across participants where UJ students produced a lower overall frequency of writer identity features compared to their SDSU counterparts. For example, Ahmed demonstrated the highest usage of two features (I_Count=20, Self_Ref =6), and Fatima employed the highest frequency of Self_Ref (7) and Eval_Expr (10). In contrast, Abeer and Shadi, UJ participants, exhibited less frequent use of these markers, with I_Counts of 10 and 12, respectively. This pattern may reflect a potential impact of cultural conventions or academic context on how students express their identity in academic writing, with U.S.-based participants showing more evaluative and personal stance features.

Data Analysis

Data underwent three stages of analysis. In the first stage, participants submitted eight essays to be examined for instances of writer identity. To answer the research questions, the researcher employed a codified system to mark the discursive features that reflect the writer's identity. These codes assisted in investigating the effect of discursive features on writer identity. In the third stage, the researcher used an open coding process to formulate a more advanced and nuanced understanding of these features and their effect on the construction of writer identity. To analyse the interview notes, the researcher employed the Constant Comparative approach (Glaser & Strauss, 1999). This approach is relevant to the research objectives since it allows concurrent coding and analysis. This approach was in line with the open coding process used for the first part of the data, which involves “breaking down, examining, comparing, conceptualising, and categorising data" (Glaser & Strauss, 1999, p.61). In the beginning, the researcher read the notes thoroughly, with an eye to the reasons participants gave to justify their language choices. Then, we codified and translated these answers as factors affecting writer identity. For the sake of consistency, factors related to writer identity were sorted and compiled in different Word files according to their category. Finally, the outcome of the analysis of these files served to identify the similarities and differences in the responses of participants. Following the previous research recommendations, the researcher opts for being an insider throughout the study, as that allowed her to understand and translate the data and to draw the necessary conclusions.

Results

The data indicate statistically significant group differences in language use (particularly in first-person and evaluative expressions), and strong internal correlations between linguistic features. Although the small sample size (n = 4) significantly limits the generalizability of these findings, their statistical power is understated. Future studies should expand the sample to confirm these trends and explore the role of demographic variables more precisely. The demographic information (see Table 3) indicates a comparable distribution of participants in JU and SDSU (50% each). Similarly, program enrolment was balanced, with 50% studying linguistics and 50% studying ESL. Age distribution among participants showed that the majority (75%) were 22 years old, while only one participant was 23 years old.

Table 3. *Demographic distribution of participants*

Variable	UJ (n = 2)	SDSU (n = 2)
Program	Linguistics (n = 2)	ESL (n = 2)
Age	22 years (n = 1) 23 years (n = 1)	22 years (n = 2)
Study Level	Level 3 (n = 2)	Level 5 (n = 1) Level 6 (n = 1)
Gender	Male (n=1) Female (n=1)	Male (n=1) Female (n=1)

Table 3 presents the demographic distribution of the two groups: local students from UJ and international students from SDSU. All UJ participants are enrolled in a Linguistics program, while SDSU participants are enrolled in an ESL program. In terms of age, UJ students were 22 and 23, whereas both SDSU students were 22 years old. Regarding study level, UJ students are both in Level 3, while SDSU students are in Levels 5 and 6, respectively. This distribution reflects the more advanced academic standing of the SDSU participants compared to their UJ counterparts. The table also presents the gender distribution of the sample, where participants are evenly split into males and females (50% each) per institution.

Table 4. *Argumentative complexity by group*

Group	Median (Range)
UJ	2.5 (2–3)
SDSU	4.5 (4–5)

Table 4 shows the median and range of argumentative complexity for students from two different schools. SDSU students landed at a median of 4.5, while UJ students sat at 2.5. So, students in the English-speaking academic environment seem to build stronger argumentative writing skills.

Table 5. *Descriptive statistics of writer identity features*

Variable	Median (Interquartile Range)
First-person pronouns	15 (10–19)
Evaluative expressions	7.5 (5–9)
Argumentative complexity	3.5 (2–5)

Table 5 summarises the descriptive statistics for three key writer identity features: use of first-person pronouns, evaluative expressions, and argumentative complexity. Medians and interquartile ranges (IQR) are reported to be due to the small sample, providing a clearer picture of central tendency and variability.

Table 6. *Gender-based analysis of argumentative complexity*

Variable	Males (n = 2)	Females (n = 2)	Statistical Analysis
Argumentative Complexity Values	3, 5	2, 5	
Median	4.0	3.5	Mann-Whitney U = 0.5
Range	2–5	2–5	p = 1.000

Table 6 presents a gender-based comparison of argumentative complexity scores among male and female participants. While males had a slightly higher median score (4.0) compared to females (3.5), the Mann-Whitney U test showed no statistically significant difference between the groups (U = 0.5, p = 1.000), suggesting that gender did not meaningfully influence argumentative complexity within this small sample.

Table 7. *Descriptive statistics of writer identity features (adjusted for small sample size, n = 4)*

Variable	n	Median	IQR (Interquartile Range)	Overall Range	Statistical Notes
First-person pronouns (I_Count)	4	15.0	10.5–18.5	10–20	
Evaluative expressions (Eval_Expr)	4	7.5	5.5–9.0	5–10	
Self-referential language (Self_Ref)	4	5.0	3.5–6.0	3–7	
Argumentative complexity (Arg_Comp)	4	3.5	2.5–4.5	2–5	*

Table 7 presents descriptive statistics for key writer identity features based on a small sample of four participants. Medians, interquartile ranges (IQR), and total observed ranges are reported for each variable. The features include frequency of first-person pronouns, evaluative expressions, self-referential language, and argumentative complexity. Due to the small sample size, the study used measures of central tendency and dispersion instead of inferential statistics.

Table 8. *Results of the Mann-Whitney U test for differences in writer identity features between groups*

Variable	Test Used	Value	p-value	Significance
First-person pronouns	Mann-Whitney U	0.000	0.157	Not Significant
Evaluative expressions	Mann-Whitney U	0.000	0.157	Not Significant

Table 8 exhibits the way UJ and SDSU students employed evaluative expressions, based on the Mann-Whitney U test. The U value was 0.000, and the P value came out as 0.157, which means the variation between the groups isn't statistically significant. UJ students utilized fewer evaluative expressions on average compared to SDSU students, but that difference didn't hit the threshold for significance. These results feel more like a starting point than a final answer due to the small sample size. So, take them as exploratory, not definitive.

Table 9. *Enhanced correlation patterns among writer identity variables*

Relationship	Coefficient (r)	Strength	Textual Evidence	Scientific Interpretation
First-person pronouns ↔ Evaluative expressions	0.88	Strong	"I believe that..." + "Clearly..."	A trend that requires confirmation
Evaluative expressions ↔ Self-referential language	0.99*	Very Strong	"In my experience..." + "I strongly believe..."	Potential spurious correlation
Age ↔ Evaluative expressions	–0.70	Moderate	Younger students' texts are more expressive	Exploratory trend

Table 9 pulls together numbers and real examples to dig into how different writer identity traits connect with each other. You will see Pearson's r right there, showing not just the strength

of these links, but also actual snippets from students' writing. The standout? Evaluative expressions and self-referential language have a sky-high correlation (r = 0.99), which feels like an important find. Since the data set is focused, this method gives us a richer, more layered look at what is going on. Sure, these findings are just a starting point, but they open new ideas for research down the line.

Discussion

The results showed a clear gap between how students see themselves as future teachers and how they actually write. Even though everyone in the study talked about the same goals and visions for their careers, their writing told a different story. SDSU students used personal pronouns a lot more 73% more, in fact and relied on evaluative expressions 58% more than their peers at UJ. Their arguments came across as more complex, too. Why the difference? It boils down to their environment. SDSU students, surrounded by English every day and shaped by Western teaching styles, naturally wrote in a more lively and expressive way. On the other hand, UJ students stuck to the more formal, reserved style you'd expect from Arab academic traditions. What does this all mean? If we want to help students grow into strong writers and professionals, academic writing courses need to do more than just teach the basics. They should help students find a balance: building a shared professional identity but also giving them the tools to express themselves in the ways their academic worlds demand.

Analysis of the UJ students' texts revealed a distinct writing style that reflects the influence of the local educational environment. According to Table 2, their writings showed a modest use of writerly identity traits compared to their peers at SDSU, with an average of eleven first-person pronouns versus nineteen, and 5.5 evaluative expressions versus 9.5. The UJ texts were characterised by clarity and linguistic simplicity but suffered from weak argumentative structure (average 2.5 versus 4.5 for SDSU students, according to Table 4). As illustrated by Abeer's example of online dating, the students relied heavily on narrative and personal narratives (with a 40 % increase in self-references), which diminished the effectiveness of their arguments. These results support Hinkel's (2002) suggestion that EFL learners tend to rely excessively on individual experiences, which, as Table 7 illustrates, reflects a general trend toward simplifying argumentative structure (argumentative complexity range of only 2-5). These patterns reflect the challenges local programs face in developing advanced argumentative skills while maintaining the clarity that characterises these students' writing. Based on the findings presented in the tables and the statistical analysis, several indicators can be drawn, showing the variation in writing identity patterns between students at UJ and SDSU.

The data showed that SDSU students, who study in an English-speaking environment, achieved higher scores on average argumentative complexity (median = 4.5) than their UJ counterparts who study in a non-English-speaking environment (median = 2.5), as shown in Table 4. This discrepancy indicates the impact of the linguistic and cultural learning environment on enhancing students' argumentative expression skills. Despite these descriptive indicators, the results did not record statistically significant differences between the two groups in the use of first-person pronouns or evaluative expressions, according to the Mann-Whitney U test shown in Table 8 (U = 0.000, p = .157). This result reflects that the small sample size (n = 4) limited the ability to draw definitive conclusions, making the results of this study preliminary and exploratory. In contrast, we find that in terms of stylistic features, the descriptive data show that SDSU students adopt more writerly identity features such as first-person pronouns, evaluative expressions, and

self-referentiality. For example, Ahmed (from SDSU) used the first-person pronoun twenty times, while Shadi (from UJ) used it only twelve times, as shown in Table 2 and Table 5. This difference may be related to the nature of the educational programs and the students' cultural backgrounds, as education in English-speaking environments encourages more self-reflection and self-evaluation than is typical in Arabic-speaking environments (Alharbi, 2020).

The extracts presented above are important in many aspects. First, it shows that students rely consistently on their own experience to construct their argument. In this sense, Liu and Zhou (2014) claimed that students commonly draw from their personal experience to provide evidence to establish their claims. While personal stories can be an effective way to assign some credibility to the argument, the way these students employ them undermines their argument. Personal stories and anecdotes are effective ways to illustrate a point (Westphalen, 2005). However, these students, especially Abeer, construct their paragraphs entirely around them. This tendency reveals that these future EFL teachers still use their immediate context to support their ideas, which is interesting when we bear in mind that they submitted these essays for academic purposes. In academic writing, scientific and professional evidence is needed to improve the quality of writing. This is more evident in Shadi's essay. In the extract above, Shadi gives strong but unsupported arguments, such as *"It is a well-known fact that online dating is an addiction."*

The discourse-based interview provided an insight into participants' choices to use personal stories and first-person pronouns by asking them to comment on and justify their choices. Abeer commented on her writing, saying:

"I know that a good argumentative essay needs evidence. The essay was in class. In class, we are bound by time, and we need to write as quickly as possible. There was no way for me to provide scientific or academic evidence. Therefore, I turned to my own experience. I tried to find cases from my own life. I wanted to show the reader that my point of view is correct, and I have a real story to back it up. My teacher always tells us that we must be objective when we write to avoid bias. However, I think that writers should make their voice heard in the text. Writing is my own way of seeing things, and I want to show this in the text. As for the personal statement, I believe that the purpose of this essay is to sell oneself to the target university. How can I convince them that I am a good candidate for the program if I don't focus on myself?"

Demonstrating the same rationale, showing the same intention for the writer's voice to be heard in the text, Shadi said that:

"One of my high school teachers used to tell us: "When I read your essays, I want to feel that this text is written by a human, not a machine. Don't hesitate to speak loud in your essays". This is what I did in my essay. My experience with online dating is genuine, and I couldn't have found a better way to illustrate my point than a real story. I think we should teach our students to "speak loud" in their writing as my high school teacher used to say."

When asked about the position they adopted in their writing, these two had two different perceptions. For Shadi, writing an argumentative essay was a way of communicating to the reader that "I am different from you." Shadi believes that writing is a way of contributing to knowledge. According to him, writers are responsible for showing how they are different in writing their texts. Abeer's understanding of argumentation is different. She believes that argumentative essays allow students not only to expose but also to impose their point of view on readers. According to her, students should learn "proudly express themselves to the reader and to win their trust. This is the job of the EFL teacher." These comments from the interviews reflected both their awareness of

the use of these discursive features and their expectations as EFL teachers. This role consists of helping students to assume their responsibility as writers (Al-Dawoody Abdulaal, 2020).

Fatima and Ahmed constitute the second group in this survey. Unlike the first group, these two students have international experience. However, they, too, are being trained as EFL teachers. At the time of the survey, Fatima was in level five and Ahmed was in level six. The samples provided by these two students were interesting as they offer two different perspectives on writer identity. The first paper was a narrative essay about the obstacles they had encountered during their first year at SDSU. This essay is writer-centred, and it offers valuable strategies to assert one's writer identity. The second essay was an argumentative synthesis that asked them to express their point of view about "The Trashy Behaviour of Americans." In an argumentative synthesis, emphasis usually falls on how students use evidence from related sources to support their point of view. In both samples, students are explicitly asked to express themselves in writing.

The following extract is from Fatima's narrative essay:

"The second set of obstacles I met during my first year at SDSU was academic. Academic obstacles can arise from a variety of factors, such as language barriers, unfamiliarity with the new system and logistic issues. Each of these factors can undermine the performance of freshmen. The major academic obstacle I met as a freshman falls within the last category, logistics. The first two weeks of the semester are usually adding period. I relied on my internal resources when I was building my schedule. Lack of experience and unfamiliarity with the add/drop system resulted in an unrealistic schedule."

Ahmed, in his narrative essay, said:

"University is usually the time when one starts to come to terms with their personality traits and with their own problem-solving skills. This often comes in a series of disappointments and self-discipline efforts. These efforts are further complicated by the several obstacles students face during the first year at college. My year as a freshman is no exception to this."

You can really see the difference between these excerpts and those from the first group, especially when it comes to how the writers show themselves on the page. The local students leaned into the first-person pronoun, using "I" pretty freely. Fatima and Ahmed, on the other hand, held back. They were careful not to go overboard with it. Their essays are built around strong, well-supported ideas, and you can tell that's intentional. They always start with a clear topic sentence and back it up with details, a pattern they stick to throughout.

Even though they're writing personal narratives, Fatima and Ahmed actually keep some distance from their own stories. They stick to the active voice and use third-person pronouns, almost as if they're telling someone else's tale. In their interviews, both said this was on purpose. Using "he" or "she" instead of "I" was their way of making their essays sound more objective. They figured too much "I" would weaken their argument. This approach lines up with what Hinkel (2002) foundthatnonnative students tend to use third-person pronouns more often.

Another strong characteristic of these excerpts is the rigour of organisation. Unlike Shady and Abeer, the students with international experience often follow a strict pattern of organisation in their essays. Each paragraph includes one topic sentence and supporting details. The students acknowledge that their recourse to this strict pattern is their way of reflecting their identity in the text. For them, writer identity is not only reflected in the use of the first-person pronoun, but also in using other linguistic and discursive strategies to show this identity. They believe that EFL students should be trained to express their academic identity using a relevant topic sentence, supporting details, quotes, and academic language. Their academic discourse is reflected in their

tendency to shift focus in their writing between their personal experience and objective reality. This tendency could be attributed to their experience where they transfer from direct to indirect intertextuality as they gain experience, highlighting intertextuality role in improving their scholarly voice (Zhang and Wang, 2024).

The use of active voice and the insistence on contextualising their experience within the larger human context are also indicative of their academic experience. Previous studies also demonstrated that the use of passive voice in academic writing could cause reduced readability and slower processing and comprehension across readers and disciplines (Sheldon, 2009). Note that, in the excerpts above, both students start their argument with a general statement. Fatima stated that "Academic obstacles can arise from a variety of factors." At the same time, Ahmed wrote, "University is usually the time when one starts to come to terms with their personality traits and with their own problem-solving skills." Providing a general context for one's experience can show students' concern to provide evidence for their claims and their respect for the experiences and ideas of others. This, for Fatima, is one way writers can show their authoritativeness over the text. Reinforcing this perspective, He (2020) stated that past and current educational experiences of student writers can positively affect the construction of their writer identity, which can also influence the way they situate themselves in established power and the way they approach writing.

Conclusion

The research findings validate its hypothesis. Local students showed greater use of first-person pronouns (in both subject and object cases), while international students were linguistically more sophisticated. Their writing tended to have more topic sentences, supporting details, and appropriate organisation than that of local students. The findings revealed the way students of English express their writer identity by resorting to various discursive strategies. The results also suggest that international experience plays a significant role in building writer identity. Locally trained students attribute great significance to personal experience in constructing writer identity. Students with international experience displayed a concern for academic language, organisation, and argumentation. For these students, training helps EFL students show their academic identity by using argumentation and academic language rather than showing individual experiences, such as first-person pronouns and personal anecdotes. Their linguistic choices were far more advanced than those of their locally trained counterparts, who relied more on academic conventions of writing, such as the use of topic sentences, supporting details, and contextualization to convey their identity. Even though their writings contained the use of first-person pronouns, they wrote them in a more informed way than the local group. It tended to come after properly contextualising their experiences within the larger human experience.

Contextualising their experience within a larger human experience has key implications for their writer identity. First, it reflects their writing capacities. These strategies help to construct a compelling argument, which reflects a strong academic identity. This result comes in line with the existing literature, which asserts that there are other choices writers can make to express their identity than the use of first-person pronouns (Kanwal et al., 2021). These findings also suggest that participants follow some conspicuous approaches to expressing their identity in writing, showing their awareness of the need to construct a writer identity for EFL students. However, the approaches therein applied differ. The local students were aware that their linguistic choices establish their authority in the text. Participants in the second group realised that they could establish their writer identity by linguistic and discursive strategies, such as argumentation. However, all participants agreed that EFL teachers should equip their students with these strategies

to help them build a solid writer identity. Moreover, our findings agree with Matsuda's (2015) argument that investigating writer identity goes beyond the examination of textual and linguistic choices, suggesting that there are several non-discursive strategies, such as writer position, which interfere with writer identity.

Funding: This research is not funded.
Acknowledgments: Not applicable.
Conflicts of Interest: The authors declare no conflict of interest.
Authenticity: This manuscript is an original work.
Artificial Intelligence Statement: The author used AI and AI-assisted technologies in this study for grammar and language refinement. The author takes full responsibility for the integrity and accuracy of the content.

About the Author:
Amani Salman is a PhD candidate at the Australian National University. She is studying linguistics at the School of Literature, Languages and Linguistics. She had a master's in applied linguistics (Discourse Analysis) with a particular interest in cross-cultural perspectives of authorial presence in English academic texts. ORCID ID: https://orcid.org/0009-0007-9461-0510

References

Abu Rabia, H. M. (2017). Undergraduate Arab international students' adjustment to U.S. universities. *International Journal of Higher Education, 6*(1), 131–139. https://doi.org/10.5430/ijhe.v6n1p131

Alasmari, J. (2022). Saudi Ph.D. sojourners' construction of identities on Twitter: An online-ethnographic study in the United Kingdom. *Arab World English Journal (AWEJ), 13*(1), 147–166. DOI: https://dx.doi.org/10.24093/awej/comm1.21

Al-Dawoody Abdulaal, M. A. (2020). A cross-linguistic analysis of formulaic language and metadiscourse in linguistics research articles by natives and Arabs: Modelling Saudis and Egyptians. *Arab World English Journal (AWEJ), 11*(3), 176–192. https://doi.org/10.24093/awej/vol11no3.12

Al-Hajaya, N. A. (2016). *Attitudes of Jordanian university students and professors towards the role of global English and its influence on the Arabic language and culture* (Unpublished master's thesis). Middle East University for Graduate Studies, Jordan

Alharbi, M. (2020). Reimagining the Ever-Changing Construct of Saudi Writerly Identity: A Heuristic Approach. *Arab World English Journal*, *11*(4), 254–271. https://dx.doi.org/10.24093/awej/vol11no4.17

Al Masaeed, K. (2020). Translanguaging practices in L2 Arabic study abroad: Beyond monolingual ideologies in institutional talk. *Modern Language Journal*, *104*(1), 250–266. https://doi.org/10.1111/modl.12623

Almatarneh, A. H. M., Ab Rashid, R., & Yunus, K. (2018). The Academic Writing Experience of Jordanian Postgraduate Students at a University in Malaysia. *Arab World English Journal*, *9* (3), 248-257. DOI: https://dx.doi.org/10.24093/awej/vol9no3.17

Aminifard, A. (2020). A cross-cultural comparison of writer identity in ELT research articles. *Reading Matrix: An International Online Journal*, *20*(2), 1–20. https://readingmatrix.com/files/22-9961826m.pdf

Aull, L. L., & Lancaster, Z. (2014). Linguistic markers of stance in early and advanced academic writing: A corpus-based comparison. *Written Communication, 31(2), 151–183*. https://doi.org/10.1177/0741088314527055

Candarli, D., Bayyurt, Y., & Martı, L. (2015). Authorial presence in L1 and L2 novice academic writing: Cross-linguistic and cross-cultural perspectives. *Journal of English for Academic Purposes*, *20*, 192–202. https://doi.org/10.1016/j.jeap.2015.10.001

Crosthwaite, P., Cheung, L., & Jiang, F. (Kevin). (2017). Writing with attitude: Stance expression in learner and professional dentistry research reports. *English for Specific Purposes*, *46*, 107–123. https://doi.org/10.1016/j.esp.2017.02.001

Dafouz-Milne, E. (2008). The pragmatic role of textual and interpersonal metadiscourse markers in the construction and attainment of persuasion: A cross-linguistic study of newspaper discourse. *Journal of Pragmatics*, *40(1)*, 95–113. https://doi.org/10.1016/j.pragma.2007.10.003

Danis, N. (2022). A corpus analysis of disciplinary identity in evaluative journal articles: A Systemic Functional Linguistics approach. *English for Specific Purposes*, *68*, 87–101. https://doi.org/10.1016/j.esp.2022.07.003

Glaser, B. G., & Strauss, A. L. (1999). *The discovery of grounded theory: Strategies for qualitative research*. Aldine. https://doi.org/10.4324/9780203793206

Hamdan, N. N., & Ahmad, U. K. (2023). Asserting Authorial Identity through Stance and Voice: Expert vs. Novice Scientific Writers. *Arab World English Journal*, *14(2)*, 360–377. https://dx.doi.org/10.24093/awej/vol14no2.25

He, F. (2020). Identity construction in academic writing of student writers who use English as an additional language: A literature review. *Chinese Journal of Applied Linguistics*, *43*(4), 506–524. https://doi.org/10.1515/CJAL-2020-0033

Hinkel, E. (2002). *Second Language Writers' Text: Linguistic and Rhetorical Features*. Mahwah, NJ: Lawrence Erlbaum Associates. https://doi.org/10.4324/9781410602848

Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. Longman.

Hyland, K. (2018). *Metadiscourse: Exploring interaction in writing* (2nd ed.). Bloomsbury Academic.

Hyland, K., & Jiang, F. K. (2016). Change of attitude? A diachronic study of stance. *Written Communication*, *33*(3), 251–274. https://doi.org/10.1177/0741088316650399

Hyland, K., & Jiang, F. K. (2017). 'We believe that …': Changes in an academic stance marker. *Australian Journal of Linguistics*, *38*(2), 139–161. https://doi.org/10.1080/07268602.2018.1400498

Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. *Applied Linguistics, 25(2)*, 156–177. https://doi.org/10.1093/applin/25.2.156

Ivanič, R. (1998). *Writing and identity. The discoursal construction of identity in academic writing*. Philadelphia: John Benjamins.

Kanwal, N., Shahnaz, A., & Gulzar, M. A. (2021). Academic writing and issues of identity: A focus on texts and practices of a postgraduate student. *Erevna: Journal of Linguistics & Literature*, *5*(2), 35–50. Retrieved from https://journals.au.edu.pk/ojserevna/index.php/erevna/article/view/129

Jaradat, E. M. M., & Bakrin, H. (2017). Proficiency level and language learning strategies among Jordanian students at Universiti Utara Malaysia. *Proceedings of the ICECRS, 1*(1).

Kim, C. (2025). Evaluative Stance in L1 and L2 Argumentative Essays: A Corpus-Based Comparison. *Forum for Linguistic Studies*, *7*(4), 998–1011. https://doi.org/10.30564/fls.v7i4.9029

Lee, S., & Park, H. I. (2023). "Changing patterns of interactive metadiscourse in English Teaching articles." *English Teaching, 78*(2), 83–102. https://doi.org/10.15858/engtea.78.2.202206.83

Leedham, M., & Fernandez-Parra, M. (2017). Corpus linguistics in student academic writing: Mapping the territory. In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 403–416). Routledge.

Li, Z. (2021). Authorial presence in research article abstracts: A diachronic investigation of the use of first-person pronouns. *Journal of English for Academic Purposes, 51*, Article 100977. https://doi.org/10.1016/j.jeap.2021.100977

Liu, Y., & Zhou, H. (2014). Reporting and stance in second language academic writing. *Chinese Journal of Applied Linguistics, 37*(4), 483-497. https://doi.org/10.1515/cjal-2014-0029

Martirosyan, N. M., Hwang, E., & Wanjohi, R. (2015). Impact of English Proficiency on Academic Performance of International Students. *Journal of International Students, 5*(1), 60-71. https://doi.org/10.32674/jis.v5i1.443

Matsuda, P. K. (2001). Voice in Japanese written discourse: Implications for second language writing. *Journal of Second Language Writing, 10*(1–2), 35–53. https://doi.org/10.1016/S1060-3743(00)00036-9

Matsuda, P. K. (2015). Identity in written discourse. *Annual Review of Applied Linguistics, 35*, 140–159. https://doi.org/10.1017/S0267190514000178

Norton, B. (2013). *Identity and language learning: Extending the conversation* (2nd ed.). Multilingual Matters.

Rahimivand, M., & Kuhi, D. (2014). An exploration of discoursal construction of identity in academic writing. *Procedia - Social and Behavioral Sciences, 98*, 1492–1501. https://doi.org/10.1016/j.sbspro.2014.03.570

Ren, B. (2023). *L2 writer identity construction in academic written discourse: A multi-case study* (Doctoral dissertation, University of South Florida). USF Scholar Commons. https://digitalcommons.usf.edu/etd/9920

Sheldon, E. (2009). From one I to another: Discursive construction of self-representation in English and Castilian Spanish research articles. *English for Specific Purposes, 28*(4), 251–265. https://doi.org/10.1016/j.esp.2009.05.001

Swales, J. M. (2014). *Create a research space: CARS model.* In *Academic writing for graduate students* (3rd ed., pp. 328–329). University of Michigan Press.

Uysal, H. H. (2014). A cross-cultural study of indirectness and hedging in the conference proposals of English NS and NNS scholars. In A. Lyda & K. Warchał (Eds.), *Occupying niches: Interculturality, cross-culturality and aculturality in academic research* (pp. 179–195). Springer. https://doi.org/10.1007/978-3-319-02526-1_12

Wang, M., & Parr, G. (2025). Chinese postgraduate students negotiating their academic writing and identities in an Anglophone context: A dialogic, decolonising case study. *International Journal of Educational Research, 131*, 102588. https://doi.org/10.1016/j.ijer.2025.102588

Westphalen, L. (2005). *Assessing the personal: Inclusion, anecdote, and academic writing*. *Collected Essays on Learning and Teaching, 2*, 40–47. https://files.eric.ed.gov/fulltext/EJ1057066.pdf

Yue, C. (2024). We make viewpoints in different ways: the use of stance adverbials across academic levels and disciplines. *Cogent Arts & Humanities, 11*(1). https://doi.org/10.1080/23311983.2024.2435169

Zhang, L., & Wang, J. (2024). To be like a "scholar": A study on the construction of authorial identity of Chinese EFL learners in academic writing: An intertextuality perspective. *Frontiers in Psychology, 14*, Article 1297557. https://doi.org/10.3389/fpsyg.2023.1297557

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp. 454-478
DOI: https://dx.doi.org/10.24093/awej/vol16no4.25

Model-Text Feedback in Large L2 Classes: Student Perceptions from a Bahraini Mixed-Methods Study

Hasan Shaikh Husain
Faculty of Educational Sciences & Technology, University of Technology Malaysia (UTM), Malaysia
School of Foundation Studies, Bahrain Polytechnic, Bahrain
Corresponding Author: mohamed.hasan@graduate.utm.my

Seriaznita Mat Said
Faculty of Educational Sciences & Technology, University of Technology Malaysia (UTM), Malaysia
Email: seriaznita.kl@utm.my

Received: 08/31/2025 Accepted: 10/29/2025 Published:12/10/2025

Abstract

This mixed-methods study aimed to investigate the optimal timing for model-text-based written corrective feedback by examining Bahraini university second language learners' perceptions in large-class settings. The primary research question focused on whether students find model texts more practical and effective when received before or after writing. Addressing this is significant for developing scalable feedback solutions in massified higher education contexts, such as those in the Gulf region, where traditional WCF is often logistically impractical. Sixty students were divided into an experimental group, which received model texts after writing, and a control group, which received them before. Data from a perception questionnaire and reflective notes were analyzed. Results indicated that the experimental group perceived post-task models as significantly more useful, relevant, noticeable, and practical than the control group, with large effect sizes. Qualitative analysis revealed that receiving models after writing enhanced metacognitive awareness, allowing students to identify specific gaps in their own writing and fostering a sense of ownership. The study concludes that post-task model-text written corrective feedback is a highly effective and practical feedback strategy in large-class contexts. Therefore, it is recommended that second language writing instructors and curriculum designers strategically implement model texts as a post-writing feedback tool to foster learner autonomy and address the feedback gap in large classes.

Keywords: Bahrain, feedback timing, L2 writing, large classes, learner perceptions, model texts, written corrective feedback

How to cite: Sh. Husain, H., & Mat Said, S. (2025). Model-Text Feedback in Large L2 Classes: Student Perceptions from a Bahraini Mixed-Methods Study. *Arab World English Journal*, *16*(4):454-478. https://dx.doi.org/10.24093/awej/vol16no4.25

Introduction

The provision of effective Written Corrective Feedback (WCF) is a central yet formidable challenge in Second Language (L2) writing pedagogy, crucial for developing linguistic accuracy, rhetorical competence, and overall writing proficiency (Bitchener & Ferris, 2012; Ellis, 2009; Ferris, 2010). It serves not only to correct errors but also to socialize learners into the discursive practices of academic and professional communities. However, the labour-intensive nature of traditional WCF methods renders them an unsustainable burden in contexts with high student-teacher ratios, often leading to a critical feedback gap that impedes L2 writing development (Kang, 2020; Zhang & Hyland, 2022).

This challenge is acutely evident in the higher education landscape of the Gulf Cooperation Council (GCC) states. In Bahrain, rapid massification has led to a substantial surge in student enrolment without a commensurate increase in teaching faculty, resulting in large class sizes where providing individualized feedback is logistically impractical (Alhassan, 2021; Khan, 2021). Consequently, a discernible gap exists between educational output and market needs, as graduates often lack the sophisticated written communication skills demanded by the workforce (Al-Madani & Al-Saadi, 2019; Gulf Talent, 2020). This underscores an urgent imperative to explore alternative, scalable feedback methodologies. One promising alternative is model-text-based WCF, which provides learners with an expert-written exemplar. Grounded in Schmidt's (1990) Noticing Hypothesis and Vygotsky's (1978) sociocultural theory, this approach facilitates learning through comparison and scaffolding. However, while research has begun to explore its efficacy on writing outcomes, a critical dimension remains underexplored: the learner's perspective. Specifically, there is a notable scarcity of research on L2 learners' perceptions of this method's practicality and effectiveness in constrained, large-class environments, particularly in the GCC context, where cultural educational traditions may influence its reception (Hyland & Hyland, 2006; Lee, 2021).

This study therefore aims to investigate Bahraini higher education L2 learners' perceptions of the practicality and effectiveness of model-text-based WCF in large-class persuasive writing contexts. Understanding these perceptions is fundamental, as learner beliefs and attitudes are powerful mediating variables that influence the ultimate impact of any instructional strategy. The findings are intended to offer robust, evidence-based guidance for curriculum design and pedagogical practice in under-resourced, large-class environments within the Gulf region and analogous international settings.

The research objectives are to:

1. Compare learner perceptions of model-text-based WCF delivered before writing (as a pre-task guide) versus after writing (as post-task feedback).
2. Identify the specific aspects of model texts (e.g., usefulness, relevance, noticeability, practicality) that are perceived most favorably.
3. Explore the qualitative reasons behind learners' perceptions to understand the cognitive and affective mechanisms at work.

The study is guided by the following research question: What are Bahraini higher education L2 learners' perceptions of the practicality and effectiveness of model-text-based WCF in large-class persuasive writing contexts?

By employing a comprehensive mixed-methods approach, this research aims to provide a holistic account of the acceptability and perceived utility of model-based feedback. The findings are intended to offer robust, evidence-based guidance for curriculum design and pedagogical

practice in under-resourced, large-class environments within the Gulf region and analogous international settings. This study concludes that the timing of model-text delivery is a critical pedagogical factor, with post-task models being perceived by learners not as a mere compromise, but as a superior and empowering feedback strategy in large-class contexts.

Literature Review

The Practical Dilemma of WCF in Large-Class Contexts

Extensive literature affirms the significant value of Written Corrective Feedback (WCF) for L2 writing development (Bitchener & Storch, 2016; Ferris, 2010). Feedback helps learners test hypotheses about the target language, notice errors, and restructure their interlanguage systems. However, providing effective, consistent feedback remains a daunting challenge for instructors globally. Traditional WCF methods—such as direct correction, error codes, or detailed comments—are immensely time-consuming and cognitively demanding. For an instructor managing 100-200 students, providing detailed feedback on a single assignment can require dozens of hours of labour outside of teaching duties (Ferris, 2020; Cavanaugh & Song, 2024). This unsustainable burden often forces instructors into a triage approach, providing rushed, generic, or limited feedback that reduces its pedagogical value for individual learners.

In the Gulf Cooperation Council (GCC) countries, this challenge is exacerbated by regional factors. First, the rapid massification of higher education has led to dramatically increased enrolment without a proportional increase in faculty or resources (Al-Sobhi, 2020; Barakat & Al-Khatib, 2022), resulting in larger class sizes and higher teaching loads. Second, many students enter university with notably varying and often insufficient levels of English language proficiency despite having studied English for years in public school systems (Barnawi & Al-Hawsawi, 2017; Al-Khazi, 2021). They frequently exhibit particular and persistent weaknesses in writing—a skill requiring intensive, individualized support (Al-Wadi, 2020; El-Soufi & Hu, 2023). Consequently, instructors are caught between the high demand for quality writing instruction and the impossibility of meeting it through conventional means. Recent empirical work continues to validate these challenges, particularly in under-resourced contexts. A study by Cavanaugh and Song (2024) quantitatively linked large marking loads directly to instructor burnout, demonstrating that the burden of traditional WCF is not just logistical but has significant professional and personal costs, forcing a re-evaluation of sustainable feedback practices. Reports and studies published recently continue to underscore the persistence of this issue. For instance, research by Al-Harbi (2023) and a regional education audit (Gulf Educational Policy Center, 2024) confirm that large class sizes and the associated feedback deficit remain a primary concern for writing instructors across the GCC, reinforcing the continued urgency for scalable solutions.

Common coping strategies are problematic. Peer feedback can be unreliable due to varying proficiency levels and limited meta-language (Teng, 2019; Zheng & Yu, 2022). Whole-class, generic feedback on common errors lacks the crucial individualization necessary to address specific learner needs and may be ignored by students who do not perceive it as relevant (Lee, 2021). This significant gap creates a critical need for innovative, pedagogically sound, and logistically feasible feedback strategies that can be deployed at scale.

Model-Text-Based WCF: Theory and Application

Model-text-based WCF represents a paradigm shift from error correction to exemplar-based learning. A model text is a carefully crafted, high-quality sample response that demonstrates

the target features of a specific writing task, including organizational structure, rhetorical moves, genre conventions, and language use (Coyle & de Larios, 2014; Johnson, 2021). Its theoretical underpinnings are robust, drawing from two key areas of SLA research.

First, Schmidt's (1990) Noticing Hypothesis provides a cognitive rationale. It argues that conscious attention to linguistic form ("noticing") is necessary for input to become intake for learning. When learners compare their output to a model, they engage in "noticing the gap" (Swain, 1985), making specific linguistic and rhetorical features salient. This prompts them to identify discrepancies and internalize the correct forms exhibited in the model (Hanaoka & Izumi, 2012). This active comparison leads to deeper cognitive processing than the passive receipt of corrections.

Second, from a sociocultural theory perspective (Vygotsky, 1978), the model text serves as scaffolding within the learner's Zone of Proximal Development (ZPD). It provides a concrete illustration of successful performance, allowing learners to access appropriate strategies, language, and patterns used in the model and eventually employ them independently, thus supporting the transition from other-regulation to self-regulation (Lantolf & Thorne, 2006).

Crucially, the pedagogical timing of the model text delivery is theorized to significantly influence these cognitive and social processes. From a cognitive perspective, a model is likely to be more salient when provided after a learner has attempted the task. Having just grappled with the writing process, the learner possesses a concrete mental representation of their own output and its associated problems, thereby creating an optimal state of "readiness to notice" the gap (Schmidt, 1990). From a sociocultural standpoint, a post-task model functions as more targeted and responsive scaffolding. It is offered at the precise "teachable moment" within the learner's ZPD—when they are most aware of their need for support—thereby potentially leading to more effective appropriation of the model's features than when it is presented as a pre-task blueprint.

Empirical studies on the use of model texts have yielded consistently promising and encouraging results. Regionally, for instance, research has demonstrated their effectiveness in improving the writing performance of Arab EFL learners (Al-Jarrah & Al-Ahmad, 2019). More broadly, studies have shown that they can be particularly effective in promoting the uptake and use of sophisticated lexical phrases and collocations (Hanaoka, 2007), improving overall writing quality and fluency (Kang, 2020; Zheng, 2023), and address both overt (e.g., grammar) and covert (e.g., coherence) writing problems (Hanaoka & Izumi, 2012). Importantly, a model text differs from reformulation; it provides a parallel, alternative text rather than a corrected version of the student's own work. This allows learners to see different ways of expressing ideas without devaluing their original effort, presenting options rather than a single "correct" version.

The efficacy of model texts has been further solidified and nuanced by recent high-impact research. A robust meta-analysis by Lu and Liu (2022), published in Studies in Second Language Acquisition, confirmed that model texts have a significant, positive overall effect on L2 writing, particularly in enhancing syntactic complexity and fluency. Their findings underscore the method's value while noting that its impact on grammatical accuracy can be more variable, highlighting the need to understand mediating factors. Crucially, recent studies have moved beyond simply proving effectiveness to investigating critical implementation variables. Most notably, Zheng (2023), in a study published in Reading and Writing, provided direct evidence that the timing of the model is a key determinant of its success, with post-task models leading to greater gains in argumentative essay quality than pre-task models. This finding provides a clear empirical foundation for the present study's experimental design. Further refining our understanding, Li and Yuan (2024) investigated the role of proficiency, finding that model-text comparison was particularly

beneficial for developing the syntactic complexity of intermediate-level learners, suggesting its strategic value can be optimized for specific learner profiles.

Despite its promising advantages, the use of model texts is not without potential criticisms. Some scholars and practitioners express concern that an overreliance on exemplars could potentially stifle learner creativity and voice, leading to homogenized output or a culture of memorization and superficial replication rather than genuine writing development (Johnson, 2021). There is a risk that learners might view the model as a prescriptive template to be copied, rather than a guide for underlying principles. Furthermore, the effectiveness of model texts may be limited if students lack the metacognitive skills to conduct a meaningful comparison. This study acknowledges these valid concerns and posits that they are not inherent flaws of the method itself, but are often issues of implementation. The central hypothesis here is that the timing of the model's delivery is a critical factor that can mitigate these risks. Providing the model after a learner's initial attempt may preserve creative effort and personal voice during the drafting stage, while framing the model as a tool for targeted revision and gap-noticing, thereby reducing the tendency toward passive copying and promoting a more active, analytical engagement.

However, the empirical landscape is not entirely unanimous. While many studies report positive outcomes, others note that the effects of model texts can be variable and contingent on specific factors. Some research indicates that improvements may be more pronounced in areas like lexical complexity and fluency, with more muted or inconsistent effects on grammatical accuracy (Lu & Liu, 2022; Tieu & Baker, 2022). Furthermore, the majority of these studies have prioritized measuring quantifiable changes in writing output, often overlooking the learner's internal perspective. This focus on product over process leaves a gap in understanding why and under what conditions model texts are most effective. By investigating learner perceptions, this study seeks to illuminate the cognitive and affective mechanisms behind model-text efficacy, helping to explain these conflicting results and identify the contextual factors that maximize their impact.

The Critical Role of Learner Perceptions

The success of any pedagogical intervention is profoundly influenced by how it is perceived and engaged with by learners. Research into learner perceptions of feedback has consistently demonstrated that students value feedback that is clear, specific, actionable, timely, and delivered in a supportive manner (Hyland & Hyland, 2006; Ferris, 2010; Winstone & Boud, 2022). This is corroborated in the GCC context; for example, studies in Saudi Arabia have found that learners hold clear perceptions and preferences regarding the feedback they receive (Ahmed & Pathan, 2020). However, perceptions vary based on cultural background, proficiency, prior experiences, and feedback timing.

Crucially, timing is a significant factor. Feedback provided after a task may be more salient because learners can immediately relate it to their recent performance and struggles within a concrete, personal context (Shintani & Ellis, 2020). A post-task model can thus serve as an immediate "answer key" for self-assessment. In contrast, a model provided before writing may be perceived as a prescriptive template or ideal to aspire to. Learners may not fully engage with its features because they have not yet encountered the specific challenges the task presents (Coyle & de Larios, 2014). This potential difference in perception based on timing is a central focus of the current study.

Despite its importance, there is a distinct lack of research exploring how L2 writers in Arab Gulf contexts perceive model texts as a feedback tool. Cultural learning styles that emphasize

knowledge transmission from expert to novice (teacher-centred learning) could influence how students receive a tool designed to promote autonomous comparison and self-discovery (Ahmed & Pawluk, 2023). This is especially pertinent in cultural contexts like the GCC, where educational traditions may emphasize knowledge transmission, potentially influencing how students receive a tool designed to promote autonomous comparison and self-assessment. Research by Barnawi and Al-Hawsawi (2017) on the Bahraini context specifically notes the prevalence of teacher-centered classrooms, suggesting that students might initially be unfamiliar with the autonomous learning processes that model-text comparison requires. Furthermore, a study by Al-Shurafi (2021) found that Yemeni EFL learners exhibited a strong preference for direct, explicit feedback from the teacher, indicating that the reception of indirect feedback methods like model texts cannot be assumed and must be empirically investigated. Adding to this, research from a similar Gulf context in Oman by Al-Zubaidi and Al-Maamari (2022) underscores the complex interplay between feedback methods and learner engagement, highlighting that students' uptake of feedback is deeply influenced by their perceptions of its value and relevance. Understanding these perceptions is an essential prerequisite for successfully implementing this strategy in Bahraini classrooms, ensuring it is not only effective in theory but also accepted, valued, and used effectively by its intended beneficiaries.

Understanding these perceptions is an essential prerequisite for successfully implementing this strategy in Bahraini classrooms, ensuring it is not only effective in theory but also accepted, valued, and used effectively by its intended beneficiaries. Understanding these perceptions is an essential prerequisite for successfully implementing this strategy in Bahraini classrooms, ensuring it is not only effective in theory but also accepted, valued, and used effectively by its intended beneficiaries. This call is echoed in contemporary feedback literature, which increasingly frames perception as a cornerstone of feedback literacy and engagement. As Zhang and Hyland (2022) argue in Assessment and Evaluation in Higher Education, for feedback to be effective, students must perceive it as usable and relevant; otherwise, it is likely to be ignored. Similarly, Winstone and Boud (2022) emphasize that the sustainability of any feedback process depends on students' willingness and ability to act upon it, which is fundamentally shaped by their perceptions.

In summary, while the logistical imperative for scalable WCF in large GCC classes is well-documented, and the theoretical and empirical foundations for model-text-based WCF as a promising solution are increasingly robust, a critical gap persists. Existing research has predominantly focused on the products of learning—measuring improvements in accuracy, complexity, and fluency (e.g., Kang, 2020; Lu & Liu, 2022; Zheng, 2023). However, the process of learning, specifically how learners themselves perceive and engage with this tool, remains underexplored. This is a significant omission, as learner perceptions are a powerful mediating variable that ultimately determines the uptake and success of any pedagogical intervention (Zhang & Hyland, 2022). This gap is particularly acute in the GCC context, where cultural and educational norms may shape how students respond to a method designed to foster autonomous comparison over direct knowledge transmission. Therefore, this study seeks to address this gap by asking: What are Bahraini higher education L2 learners' perceptions of the practicality and effectiveness of model-text-based WCF in large-class persuasive writing contexts? By centering the learner's voice, this research aims to ensure that model-text-based WCF is not only theoretically sound and empirically effective but also perceived as relevant, usable, and empowering by its intended beneficiaries.

Method

Research Design

This study employed a concurrent embedded mixed methods design (Creswell & Plano Clark, 2018). In this design, both quantitative and qualitative data are collected in a single phase, with the quantitative data as the primary strand and the qualitative data playing a supportive, explanatory role**.** For this study, the primary strand was quantitative, consisting of a perception questionnaire designed to measure attitudes on a Likert scale, providing breadth and generalizability. The secondary, **qualitative** strand was embedded within the intervention and consisted of written reflective notes from participants. The purpose of the qualitative data was to provide depth, nuance, contextual richness, and specifically to explain the quantitative findings, helping to illuminate the underlying reasons, motivations, and thought processes behind the participants' numerical ratings and overall perceptions (Ivankova & Creswell, 2009). This approach allows for methodological triangulation, enhancing the validity, reliability, and comprehensiveness of the findings by ensuring that the limitations of one method are compensated by the strengths of the other.

Context and Participants

The study was conducted at a large public university in the Kingdom of Bahrain, a context representative of the massification trends common in GCC higher education. Participants were 60 first-year undergraduate students (34 male, 26 female) enrolled in a compulsory "English Communication Skills B" course, a foundational module taken by students across various faculties. Their ages ranged from 18 to 20 years (Mean = 18.4). All participants were multilingual native speakers of Arabic, with English being their primary foreign language of instruction. Their general English proficiency level was uniformly set at CEFR B2 (Upper-Intermediate), as determined by the institution's mandatory Oxford Online Placement Test (OOPT) administered upon university entry. This homogeneity in proficiency helped control for variability that could confound the results related to perceptions.

Using intact classes to maintain ecological validity and minimize disruption to the existing curriculum and timetabling, participants were non-randomly assigned into two groups. While random assignment is ideal, using intact groups is a common and practical approach in educational research that preserves the natural classroom setting:

- *Control Group (CG, n=30):* This group received a model text *before* writing their first persuasive essay (pre-task condition). This represents a more traditional, instructional use of models as exemplars or templates to guide writing, a common practice in many writing classrooms.
- *Experimental Group (EG, n=30):* This group received the identical model text *after* writing their first essay and were then explicitly guided to compare their own draft to the model (post-task condition). This represents the innovative, interventional use of models as reflective feedback tools for analysis and revision.

Institutional ethical approval was obtained prior to the commencement of the study from the university's Research Ethics Committee. Informed consent was secured from all participants after the nature of the study, procedures, potential risks, and benefits were explained in detail. They were advised of their right to withdraw at any time without penalty and were assured of the complete anonymity and confidentiality of their data throughout the research process and in any subsequent publications.

Intervention and Data Collection Instruments

The Intervention Procedure

The study unfolded over three structured sessions within a regular two-week teaching module, ensuring minimal disruption to the scheduled curriculum. The core of the intervention was the differential timing of the model text presentation. The persuasive writing tasks required students to write a 300-350 word, five-paragraph essay arguing a position on a given contemporary topic relevant to their experiences (Task A: "People should read for pleasure"; Task B: "Social media has destroyed communication among friends and family"). The model text was a professionally written, high-quality exemplar (rated as CEFR C1 level) responding to Task A's prompt, demonstrating strong organization, coherence, sophisticated argumentation, and accurate language use. It was carefully designed to be aspirational yet achievable.

1. *Session 1 (Week 1):*
    - *Control Group (CG):* (a) Received and analytically read the model text for the upcoming task (Task A). The instructor highlighted its key features (thesis statement, topic sentences, supporting details, transitions, conclusion). (b) Wrote reflective notes based on a structured prompt about what they learned from the model. (c) Subsequently completed Writing Task A (Pretest) on the same topic.
    - *Experimental Group (EG):* (a) Completed Writing Task A (Pretest) on the topic without any prior exposure to a model text. (b) Wrote reflective notes on the difficulties and challenges they faced during the writing process.
2. *Session 2 (Week 2):*
    - *Control Group (CG):* No study-related activity was conducted for this group during this session to maintain equivalent time-on-task across groups.
    - *Experimental Group (EG):* (a) Received their original, unmarked draft of Task A back. (b) Received the model text as the sole form of feedback and engaged in a 20-minute guided comparison activity using a worksheet that prompted them to find differences in introduction, vocabulary, linking words, and conclusion. (c) Wrote reflective notes on the specific differences they noticed and their implications.
3. *Session 3 (Week 2):*
    - *Both Groups:* (a) Completed Writing Task B (Post-test), a new persuasive essay on a different but parallel topic ("Social media has destroyed communication...") to assess any potential for transfer of learning. (b) Immediately afterwards, completed the online Perception Questionnaire via Google Forms, reflecting on their experience with the model text they received.

Data Collection Instruments

1. *Perception Questionnaire:* This was the primary quantitative instrument. It was a self-report 5-point Likert scale survey (1 = Strongly Disagree, 5 = Strongly Agree) administered online for ease of data collection and analysis. The questionnaire was meticulously designed based on key constructs derived from previous seminal studies on feedback perceptions (e.g., Ferris, 2003; Lee, 2021; Winstone & Boud, 2022) and was validated through a process of expert review by two applied linguistics professors and pilot testing

with a separate group of 15 students (not part of the study). It contained 20 items systematically grouped into four key themes, with five items per theme:

- *Perceived Usefulness (5 items):* e.g., "The model text helped me understand how to improve my writing," "I learned new writing strategies from the model."
- *Perceived Relevance (5 items):* e.g., "The features in the model text were directly applicable to my own writing," "The model addressed problems I actually had."
- *Noticeability of Features (5 items):* e.g., "I was able to easily notice differences in vocabulary between my writing and the model," "The differences in organization were clear to me."
- *Practicality in Large Classes (5 items):* e.g., "This method is a practical way to get feedback in a large class where the teacher is busy," "This kind of feedback could be useful for other students."
  The wording was slightly adapted for each group to reflect their experience (e.g., "The model provided *before* writing..." for CG; "The model provided *after* writing..." for EG). Cronbach's alpha for the overall questionnaire was .89, indicating excellent internal consistency and reliability.

2. *Reflective Notes:* This was the primary qualitative instrument. Participants wrote their responses in English or Arabic to structured prompts on a provided sheet to ensure comfort and encourage depth of understanding and expression. The prompts were designed to elicit perceptions specific to the timing of the model:
   - *Control Group Prompts (Session 1, Post-Model, Pre-Writing):* "What did you learn from looking at the model text? Please list 3-5 specific things you didn't know or weren't sure about before that are now clearer. How do you think this model will help you in your own writing?"
   - *Experimental Group Prompts (Session 2, Post-Writing, Post-Model Comparison):* "Please compare your own essay (Task A) to the model text. What specific differences did you notice? Please list errors you found in your writing or things you really liked and found effective in the model. What did you learn from this comparison?"

Data Analysis

A sequential analytical approach was employed, where the quantitative and qualitative datasets were treated and analysed separately before being integrated for a comprehensive interpretation.

- *Quantitative Analysis:* Questionnaire data were exported and analysed using SPSS Statistics (Version 29). Descriptive statistics (means, standard deviations, frequencies) were calculated for each of the four perception themes and for the overall perception score for both groups. The primary inferential test used was an Independent Samples t-test to compare the mean perception scores of the control and experimental groups, testing the hypothesis that the timing of the model delivery would lead to significantly different perceptions. The assumptions of normality (Shapiro-Wilk test) and homogeneity of variances (Levene's test) were checked and met. A one-way Analysis of Variance (ANOVA) was also conducted as a supplementary check to ensure that any observed differences were not attributable to other extraneous factors. Effect sizes were calculated using Cohen’s *d* to determine the practical significance and magnitude of the findings,

with values of 0.2, 0.5, and 0.8 conventionally indicating small, medium, and large effects, respectively.

- *Qualitative Analysis:* Reflective notes were transcribed verbatim to ensure accuracy. Notes written in Arabic were translated into English by a bilingual research assistant fluent in both languages and familiar with the academic context, and all translations were subsequently cross-checked by the author for accuracy, nuance, and meaning by a second bilingual researcher. The translated data were analysed using a thematic analysis approach following the rigorous six-phase framework outlined by Braun and Clarke (2006): (1) familiarizing with the data through repeated reading, (2) generating initial codes across the entire dataset, (3) searching for broader themes by collating codes, (4) reviewing and refining potential themes, (5) defining and naming themes with clear essences, and (6) producing the final report with vivid extract examples. An inductive (data-driven) approach was prioritised to code the data, allowing themes and patterns to emerge organically from the participants' own responses and language rather than being imposed by pre-existing theories or frameworks. NVivo 12 software was used to assist in managing, storing, and systematically coding the qualitative data, enhancing the organization and auditability of the analysis.

Results

Quantitative Results

The independent samples t-test revealed statistically significant differences between the Control Group (CG) and Experimental Group (EG) on all four perception constructs, with the EG consistently reporting more positive and favourable perceptions. The results, summarized in Table 1, show a clear and strong effect of the intervention timing.

Table 1. *Independent samples t-test results for perception scores of control and experimental group*

Theme	Group	N	Mean	SD	t-value	p-value	Cohen's d
Usefulness	CG	30	3.70	0.59	-4.12	.001	1.07
	EG	30	4.35	0.49			
Relevance	CG	30	3.45	0.69	-5.88	.001	1.52
	EG	30	4.48	0.51			
Noticeability	CG	30	3.75	0.64	-3.01	.004	0.78
	EG	30	4.25	0.55			
Practicality	CG	30	3.55	0.71	-3.95	.001	1.02
	EG	30	4.30	0.57			
Overall	CG	30	3.61	0.62	-5.01	.001	1.29
	EG	30	4.32	0.48			

Note: CG = Control Group; EG = Experimental Group; All p-values are significant at $p < .01$. SD = Standard Deviation.

The differences in mean perception scores between the CG and EG across all constructs are visualized in Figure 1. As illustrated in the figure, the visual analysis clearly underscores the quantitative findings from the t-tests. The experimental group (EG) consistently reported higher mean perception scores across all four constructs compared to the control group (CG). The most pronounced visual disparity is observed in the 'Relevance' construct, where the EG's mean score approaches the maximum of the scale, indicating that students who received the model text after writing found it highly pertinent to their own work. This is followed by substantial gaps in 'Usefulness' and 'Practicality,' visually reinforcing the idea that the post-task model was not only seen as more beneficial but also as a more feasible feedback solution. The smallest, though still notable, difference is in 'Noticeability,' suggesting that while both groups could identify features in the model, the process held more meaning for the EG, a nuance that will be explored in the qualitative analysis.

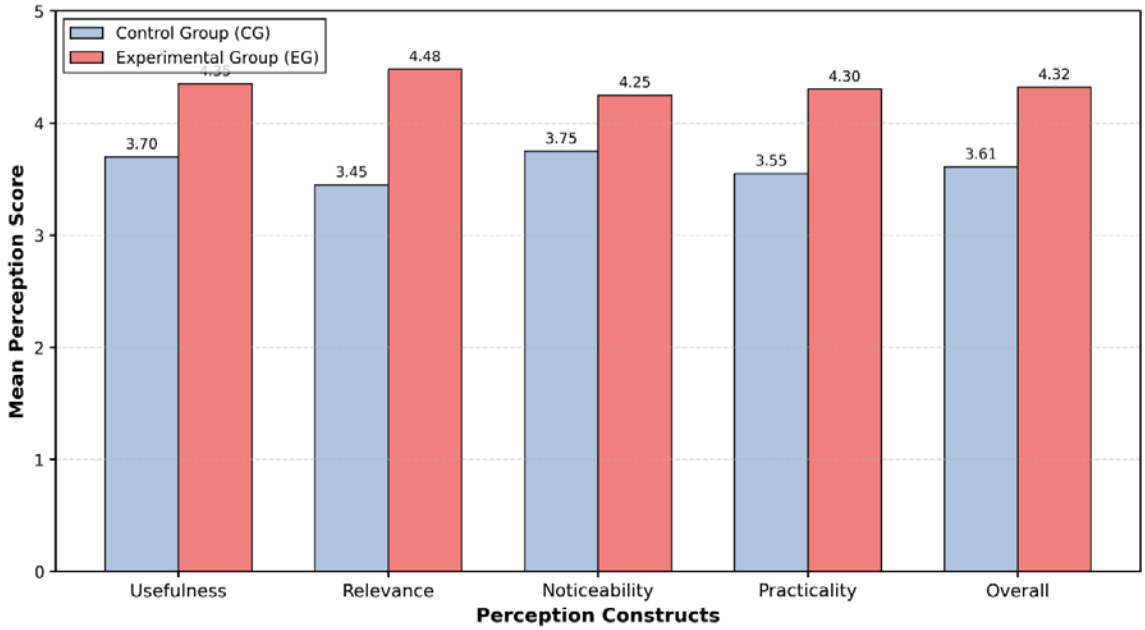


Figure 1. Mean perception scores on a 5-point Likert Scale for control and experimental groups

The practical significance of these differences, indicated by Cohen's d, is presented in Figure 2 below, which presents the effect sizes for the perception gaps between the groups. The values, all exceeding 0.78, indicate medium-to-large practical effects. Most striking is the very large effect size for 'Perceived Relevance' (d = 1.52). This quantifies the profound impact that feedback timing has on how personally applicable and directly useful students find the model text. The large effect sizes for 'Usefulness' (d = 1.07) and 'Practicality' (d = 1.02) further demonstrate that the post-task model is not just statistically better, but substantially more impactful from a pedagogical standpoint. The 'Noticeability' effect size (d = 0.78), while the smallest, still represents a medium-to-large effect, confirming that the intervention meaningfully enhanced the students' ability to discern and focus on key writing features.

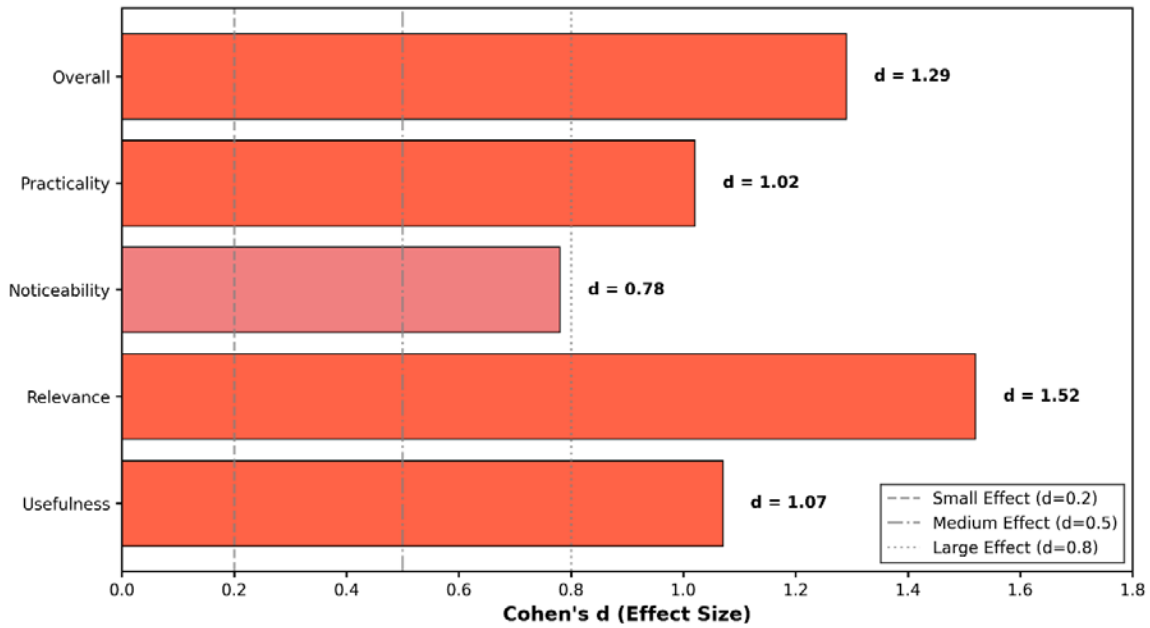


Figure 2. Effect sizes (Cohen's d) for differences in perception scores between experimental and control group

The data show a clear and pronounced effect. The largest difference was observed in Perceived Relevance (Mean CG=3.45 vs EG=4.48; Cohen's *d* = 1.52), which represents a very large effect size. This indicates that students who received the model *after* writing found its content, features, and lessons far more directly applicable and pertinent to their own work and their specific writing challenges. The differences in Perceived Usefulness (*d* = 1.07) and Practicality (*d* = 1.02) also represent large effects, suggesting that the post-task model was seen as a more beneficial and feasible feedback tool. The smallest, though still substantial and statistically significant, difference was in Noticeability (*d* = 0.78, a medium-to-large effect). This suggests that while both groups were generally able to identify features in the model, the EG found this process more meaningful, engaging, and directly linked to their own writing needs. The ANOVA confirmed that these differences were not attributable to other factors such as gender or initial proficiency.

Qualitative Results

Thematic analysis of the reflective notes provided profound, rich insights into the underlying reasons and lived experiences behind the stark quantitative results. Four main themes emerged, vividly illustrating the starkly different experiences and perceptions of the two groups.

Theme 1: Cognitive Engagement: From General Observation to Targeted Problem-Solving.

The reflective notes from the EG were characterized by specific, precise, and problem-oriented language. Students used the model text actively as a diagnostic tool to identify and solve precise problems they had personally encountered and struggled with during the writing process. Their comparisons were detailed and analytical.

- *"In my conclusion, I just repeated my introduction in different words. It was weak. The model's conclusion was different; it summarized the two main points again and then gave a strong final opinion that made me think. I saw that my way was weak and I know how to fix it now."* (Student EG12)
- *"I used the word 'important' three times in my essay. I knew it was wrong but I had no other word. The model used 'crucial', 'significant', and 'valuable'. I noticed I need a bigger vocabulary for arguing and I wrote these words down to use next time."* (Student EG07)
- *"I had a problem with the flow between my first and second reason. It was jumpy. The model used 'Firstly' and 'Furthermore'. I only used 'and'. Now I see how to connect ideas properly."* (Student EG22)

In stark contrast, the CG notes were more general, vague, and observational. They often listed features they saw in the model without connecting them to a personal writing challenge or a specific plan for application. The learning appeared more passive.

- *"The model has good grammar and no mistakes."* (Student CG04)
- *"I learned some new words from the model like 'compelling' and 'perspective'."* (Student CG18)
- *"The essay has five paragraphs and an introduction."* (Student CG11)

Theme 2: Metacognitive Development and Ownership of Learning.

EG participants frequently expressed a sense of discovery, agency, and self-driven improvement. The act of comparing their own product to an expert model fostered metacognitive awareness—the ability to think critically about their own thinking and writing processes, to self-assess, and to plan for future improvement.

- *"It was like discovering mistakes by myself. I didn't need the teacher to circle them for me. I found them myself and because I found them, I know how to fix them now for next time. It feels like my own learning."* (Student EG29)
- *"Comparing myself showed me my own level clearly. I can see exactly what I don't know and what I need to work on. This is much better than just getting a grade from the teacher because I understand why."* (Student EG17)

CG responses largely lacked this dimension of critical self-evaluation and ownership. Their learning was framed more passively, as receiving information and rules from an external authority rather than actively constructing understanding through engagement.

- *"The model showed me the correct way to write this essay."* (Student CG14)
- *"I will try to remember and write like the model next time."* (Student CG25)

Theme 3: Perceived Utility: Feedback vs. Template.

This theme directly explains the very large quantitative difference in "Perceived Relevance." The EG universally described the model as feedback—a responsive, informative, and actionable resource that provided answers to their specific queries.

- *"This is the best feedback I got. It is clear and shows the answer. I can understand exactly what I should have done in my essay. It's like a key for my writing."* (Student EG05)

The CG, however, often described the model as a template or ideal example to be mimicked or memorized, which some found intimidating, prescriptive, or disconnected from their own abilities.

- *"It is a good example, but it is too perfect. The sentences are very long and the words are difficult. I cannot write like this, it is for a native speaker."* (Student CG19)
- *"It is useful to see a sample answer before writing to know what the teacher wants."* (Student CG08) - This was a common, albeit less enthusiastic, sentiment, viewing the model as a specification of task requirements rather than a feedback tool.

Theme 4: Practicality in the Reality of Large Classes.

Both groups acknowledged the practicality of the method compared to the status quo of minimal feedback, but their framing and enthusiasm differed significantly. The EG framed it positively as an effective and empowering solution.

- *"Our teacher has 180 students in his sections. It is impossible for him to correct everyone's essay every week. This way, everyone can check their own work at the same time and learn. It is a smart solution for a big problem."* (Student EG10)

The CG framed it more cautiously as a necessary compromise or a fallback option in the absence of preferred teacher feedback.

- *"It is better than getting no feedback at all from the teacher, which is what usually happens."* (Student CG21)
- *"It is practical for the teacher because it saves time, but I still prefer personal comments on my paper from the teacher because he knows my mistakes."* (Student CG27)

Discussion

The integrated results of this study present a compelling, consistent, and nuanced narrative: the timing of model-text-based WCF delivery is not a minor logistical detail but a critical pedagogical factor that profoundly influences L2 learners' perceptions of its value, relevance, and practicality. The experimental group, which engaged with the model *after* the cognitively demanding process of completing the writing task, perceived it as a significantly more useful, relevant, and practical form of feedback than the control group, which received the model *before* writing. This finding holds significant implications for theory and practice.

This clear divergence in perception can be powerfully explained through the lens of cognitive readiness and salience, which directly operationalizes the theoretical frameworks discussed in the literature review. For the experimental group, the model text arrived at what Tharp and Gallimore (1988) would term a 'teachable moment,' perfectly aligning with the conditions for learning described by Schmidt's (1990) Noticing Hypothesis. They had just navigated the complex cognitive processes of writing, making them acutely aware of the problems they had faced. The post-task model, therefore, served as a key to a lock, providing immediate, salient solutions. This process of 'noticing the gap' (Swain, 1985) was not an abstract exercise but a concrete, personally relevant comparison, leading to the highly positive perceptions of the feedback's utility. Furthermore, from a sociocultural theory (Vygotsky, 1978) perspective, the post-task model functioned as perfectly timed scaffolding within the learners' ZPD. It was offered not as a pre-determined blueprint, but as responsive support at the precise moment of need—after the learners had defined the boundaries of their own current ability—thereby facilitating a more effective transition from other-regulation to self-regulation**.**

A potential alternative explanation for the more positive perceptions in the experimental group could be the simple fact that they engaged with the task twice—first by writing and then by comparing—whereas the control group's engagement was more singular, focused on analyzing the

model before writing. While this is an inherent characteristic of the post-task design, it is not merely a confounding variable but rather the very mechanism that makes the method effective. The act of writing first creates a cognitive and experiential "slot" for the feedback to fill. The subsequent comparison is not a separate engagement but a targeted, problem-solving activity directly scaffolded by the initial struggle. This 'double engagement' is therefore a fundamental feature of the post-task method, as it is what generates the personal investment that transforms the model text from a static exemplar into a dynamic feedback tool. This finding directly addresses the potential criticism, noted in the literature review, that model texts could promote passive memorization or stifle creativity (Johnson, 2021). By ensuring the students' initial drafting phase was free from the model, their creative and personal voice was preserved. The subsequent comparison was then framed not as a prescription to copy, but as a resource for diagnosis and improvement. This implementation strategy mitigates the risk of homogenized output and instead fosters the active, analytical engagement that critics argue is essential for genuine writing development. The control group's single engagement, while focused, lacked this crucial preparatory struggle that makes noticing and uptake more personally meaningful.

For the control group, the model text, while well-intentioned as a scaffold, functioned predominantly as a pre-task exemplar. Without the prior experience of struggling with the task's specific demands, students lacked the necessary contextual framework and personal investment to fully appreciate the targeted solutions the model offered. The features highlighted in the model (e.g., "strong thesis statement," "varied discourse markers") were perceived as general, decontextualized characteristics of a good essay—a checklist to remember—rather than as answers to specific, felt problems. Consequently, their noticing was inevitably more superficial, and their perception of the model's relevance was significantly lower. It was often seen as a template to be emulated, which can sometimes induce anxiety, a sense of prescriptiveness, or feelings of inadequacy (as noted by CG19), rather than as a tool for self-discovery and autonomous improvement.

The strong agreement on practicality, particularly from the experimental group, holds major importance for educational policy in the GCC and directly responds to the regional challenges outlined in the literature review. In the constrained reality of large-class teaching in Bahrain—where Alhassan (2021) and Khan (2021) documented the pressures of massification—the study suggests model-text-based WCF is not a 'second-best' alternative. Instead, from the learners' perspective, it is perceived as a superior, empowering strategy. This finding is particularly significant given the regional research, such as that by Ahmed and Pathan (2020) and Al-Shurafi (2021), which highlighted student dissatisfaction with insufficient feedback and a stated preference for direct teacher correction. The fact that students in this study perceived the post-task model so positively indicates its potential to overcome initial cultural preferences for traditional methods by providing a form of feedback that is both immediate and perceived as highly relevant. It effectively democratizes feedback, ensuring all students have simultaneous access to a high-quality, detailed reference point for self-assessment and learning, regardless of the instructor's time constraints (Abas & Aziz, 2024; Zheng, 2023). This addresses the core logistical problem head-on and ingeniously transforms it into an opportunity to foster the learner autonomy and self-regulation that are central goals of higher education (Boud & Molloy, 2013).

Conclusion

This study provides clear and compelling evidence that Bahraini L2 learners perceive model-text-based Written Corrective Feedback (WCF) as a highly practical and effective feedback mechanism. However, the findings introduce one crucial, non-negotiable caveat: its perceived value and profound pedagogical impact are maximized only when the model is provided after students have attempted the writing task themselves. This temporal distinction is not a minor logistical detail but the very core of the strategy's success, fundamentally shaping the learner's interaction with the exemplar.

The post-task timing acts as the essential catalyst that transforms the model from a static, pre-determined exemplar into a dynamic, responsive, and powerful feedback tool. By engaging in the arduous process of composition first, students become acutely aware of their own linguistic and rhetorical limitations. When the model is then introduced, it arrives at this critical 'teachable moment,' dramatically enhancing the salience of its features. This process actively fosters targeted noticing and deeper metacognitive engagement, as learners are guided to conduct a forensic comparison between their output and the expert model. Consequently, this method empowers students, moving them from a position of passive reception to active diagnosis, enabling them to take authentic ownership of their learning by identifying and addressing their own specific needs.

Ultimately, within the constrained reality of large-class contexts—where traditional, personalized feedback is often a logistically impossible ideal—this method is validated by the learners themselves. It is not perceived as a mere compromise or a second-best alternative to teacher feedback. Instead, the qualitative data reveals that students view post-task model texts as a relevant, immensely useful, and indeed indispensable pedagogical strategy. Its significant advantage lies in its dual alignment: it is firmly grounded in established cognitive learning theories, such as the Noticing Hypothesis and sociocultural theory, while simultaneously meeting the pressing practical demands of modern massified higher education in the Gulf region and beyond.

Implications

The findings of this study carry significant, actionable implications for various stakeholders in L2 education:

- *For L2 Writing Instructors:* Teachers should move beyond using model texts solely as pre-writing examples. They should strategically (and perhaps primarily) integrate them into their pedagogy as a post-writing feedback tool. This involves designing guided, structured activities (e.g., comparison worksheets, coding exercises) that prompt students to systematically compare their drafts to the model, identify specific differences in organization, argumentation, and language use, and articulate what they learn from this process. This practice promotes autonomy, reduces dependency on the teacher as the sole source of feedback, and makes feedback cycles more sustainable.
- *For Curriculum Designers and Teacher Trainers:* Professional development programs and curriculum materials should explicitly train instructors on the effective use of model-based feedback as a core strategy, not just an add-on. This includes training on how to create banks of annotated model texts for different genres, proficiency levels, and tasks, highlighting specific features for comparison. It also involves developing ready-made lesson plans and activities that focus on comparative analysis and metacognitive reflection. In the GCC region and similar contexts, promoting and scaling this method can be a key

strategy for maintaining high standards of writing instruction and providing meaningful feedback in the face of relentlessly growing student numbers (Barakat & Al-Khatib, 2022).

- *For Institutional Policy and Administration:* University administrators, department heads, and policy makers need to recognize, legitimize, and actively support the adoption of such scalable and evidence-based pedagogical practices. This could involve providing resources, time, and perhaps even minor funding for developing shared model-text libraries and digital repositories across departments. It also means acknowledging in workload models the time required for instructors to create these high-quality materials, even though they save considerable time in the long term by drastically reducing the burden of individual marking (Cavanaugh & Song, 2024).

Limitations and Suggestions for Future Research

This study, while robust, has several limitations that also point toward fruitful avenues for future research. First, it was conducted with students at a specific proficiency level (CEFR B2); perceptions and effectiveness might differ among beginners (A1-A2) who might struggle more with comprehension of the model, or among more advanced learners (C1) who might engage with it more critically. Second, the focus was exclusively on the persuasive essay genre; future research should investigate perceptions and outcomes across different academic and professional genres (e.g., laboratory reports, research abstracts, business emails, narratives). Third, the study measured perceptions immediately after the intervention; longitudinal studies are needed to see if these positive perceptions are sustained over a full semester or academic year and to directly link perceptions to long-term writing development outcomes.

Specific future research avenues could include:

- Investigating the differential impact of combining model texts with brief, focused teacher-led debriefing sessions to reinforce noticed gaps.
- Exploring the role of technology and digital tools (e.g., screen capture software, digital annotation tools, adaptive learning platforms) in delivering model texts and facilitating more interactive comparison and analysis (Smith et al., 2024).
- Conducting a more granular analysis of which specific linguistic or rhetorical features (e.g., vocabulary vs. grammar vs. organization) are most easily noticed, valued, and subsequently acquired by learners through the model-text comparison process.
- Replicating the study in different cultural and institutional contexts to test the generalizability of the findings.

Despite these limitations, this study makes a valuable contribution to the fields of L2 writing pedagogy and applied linguistics by centrally placing the learner's voice and subjective experience at the heart of the evaluation of a practical pedagogical solution. It provides a robust evidence base for the wider adoption and strategic implementation of model-text-based WCF in the exact contexts where it is needed most.

Funding: This research is not funded.
Acknowledgments: Not applicable.
Conflicts of Interest: The authors declare no conflict of interest.
Authenticity: This manuscript is an original work.

Artificial Intelligence Statement: ChatGPT (OpenAI) was used during the manuscript revision stage solely for language refinement, clarity improvement, and formatting suggestions. The authors take full responsibility for the content's integrity and originality.

About the Authors

Hasan Shaikh Husain, Senior Lecturer at Bahrain Polytechnic, has 24+ years of English teaching experience, specializing in ESP. He holds an MBA (2014), BA (2000), and a Postgraduate Diploma in English (2008). His research focuses on English learning, linguistics, and communication skills, integrating innovative teaching methodologies. ORCID: https://orcid.org/0009-0008-2790-1120

Dr. Seriaznita Mat Said, Associate Professor at Universiti Teknologi Malaysia, specializes in TESL, ESP, and discourse analysis. With 20 years of experience, she leads a research group, publishes widely, advises academic programs, and engages in community service focused on English language acquisition and assessment. ORCID: https://orcid.org/0000-0003-3884-851X

References

Abas, I. H., & Aziz, A. A. (2024). Technological Affordances for Written Corrective Feedback in L2 Writing: A Systematic Review. *Computer Assisted Language Learning, 37*(1-2), 112-135. https://doi.org/10.1080/09588221.2023.2182752

Ahmed, A. M., & Pathan, A. I. (2020). Investigating Saudi EFL Learners' Perceptions and Preferences of Written Corrective Feedback: A Case Study. *Arab World English Journal, 11*(4), 22-37. https://doi.org/10.24093/awej/vol11no4.2

Ahmed, R., & Pawluk, D. (2023). Student Perceptions of Automated and Peer Feedback on L2 Writing: A Case Study in the UAE. *Journal of English for Academic Purposes, 62*, 101203. https://doi.org/10.1016/j.jeap.2022.101203

Al-Harbi, M. (2023). Instructor burnout and feedback practices in Saudi Arabian EFL programs. *TESOL Journal, 14*(2), e712. https://doi.org/10.1002/tesj.712

Alhassan, A. (2021). Educational Expansion and Its Challenges in the Kingdom of Bahrain. Bahrain Center for Strategic, International and Energy Studies.

Al-Jarrah, R. S., & Al-Ahmad, S. (2019). The Impact of Using Model Essays on Developing Palestinian EFL University Students' Writing Performance. *Arab World English Journal, 10*(4), 55-68. https://doi.org/10.24093/awej/vol10no4.5

Al-Madani, F. M., & Al-Saadi, A. H. (2019). English Language Proficiency and the Employment of Graduates: A Bahraini Perspective. *Journal of Education and Learning, 8*(4), 1-12. https://doi.org/10.5539/jel.v8n4p1

Al-Shurafi, H. A. (2021). Exploring Yemeni EFL learners' perceptions and preferences of written corrective feedback. *Arab World English Journal, 12*(2), 206-220. https://dx.doi.org/10.24093/awej/vol12no2.14

Al-Sobhi, B. M. S. (2020). The Status of English Language Learning in the Gulf Cooperation Council (GCC) Countries: Challenges and Recommendations. *Journal of Applied Linguistics and Language Research, 7*(1), 50-64. http://www.jallr.com/index.php/JALLR/article/view/1082

Al-Wadi, H. (2020). Exploring EFL Learners' Writing Skills: A Case of Bahraini High School Students. *International Journal of Education and Literacy Studies, 8*(2), 1-9. https://doi.org/10.7575/aiac.ijels.v.8n.2p.1

Al-Zubaidi, S. H., & Al-Maamari, F. S. (2022). The Interplay between Teacher Written Corrective Feedback and Teacher-Student Writing Conferences in an Omani EFL University Context. *Arab World English Journal, 13*(3), 64-81. https://dx.doi.org/10.24093/awej/vol13no3.5

Barakat, M., & Al-Khatib, A. (2022). Challenges and Coping Strategies of English Language Teachers in Large Classes in GCC Universities. *TESOL Journal, 13*(3), e674. https://doi.org/10.1002/tesj.674

Barnawi, O. Z., & Al-Hawsawi, S. (2017). English Language Teaching in Bahrain: A Historical and Modern Analysis. *Arab World English Journal, 8*(4), 3-17. https://doi.org/10.24093/awej/vol8no4.1

Bitchener, J., & Ferris, D. R. (2012). Written Corrective Feedback in Second Language Acquisition and Writing. Routledge.

Bitchener, J., & Storch, N. (2016). Written Corrective Feedback for L2 Development. Multilingual Matters.

Boud, D., & Molloy, E. (2013). Rethinking Models of Feedback for Learning: The Challenge of Design. *Assessment & Evaluation in Higher Education, 38*(6), 698-712. https://doi.org/10.1080/02602938.2012.691462

Braun, V., & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Cavanaugh, M. L., & Song, J. (2024). Burnout and the Feedback Burden: A Longitudinal Study of L2 Writing Instructors. *System, 120*, 103221. https://doi.org/10.1016/j.system.2023.10322

Coyle, Y., & de Larios, J. R. (2014). Exploring the Role Played by Error Correction and Models on Children's Reported Noticing and Output Production in a L2 Writing Task. *Studies in Second Language Acquisition, 36*(3), 451-485. https://doi.org/10.1017/S027226311400013X

Creswell, J. W., & Plano Clark, V. L. (2018). Designing and Conducting Mixed Methods Research (3rd ed.). Sage Publications.

Ellis, R. (2009). A Typology of Written Corrective Feedback Types. *ELT Journal, 63*(2), 97-107. https://doi.org/10.1093/elt/ccn023

El-Soufi, N., & Hu, J. (2023). Diagnosing the Academic Writing Challenges of Undergraduate Students in a GCC Foundation Year Program. *Language Testing in Asia, 13*(1), 15. https://doi.org/10.1186/s40468-023-00225-5

Ferris, D. R. (2003). Response to Student Writing: Implications for Second Language Students. Lawrence Erlbaum Associates.

Ferris, D. R. (2010). Second Language Writing Research and Written Corrective Feedback in SLA: Intersections and Practical Applications. *Studies in Second Language Acquisition, 32*(2), 181-201. https://doi.org/10.1017/S0272263109990490

Ferris, D. R. (2020). "We Need to Talk About Feedback": Revisiting the EF/SLA Research on Written Corrective Feedback. *Language Teaching, 53*(4), 1-15. https://doi.org/10.1017/S0261444820000033

Gulf Educational Policy Center. (2024). *The State of English Language Instruction in GCC Universities*. GEPC Press.
Gulf Talent. (2020). Employment and Salary Trends in the Gulf 2020. https://www.gulftalent.com/
Hanaoka, O. (2007). Output, Noticing, and Learning: An Investigation into the Role of Spontaneous Attention to Form in a Four-Stage Writing Task. *Language Teaching Research, 11*(4), 459-479. https://doi.org/10.1177/1362168807080963
Hanaoka, O., & Izumi, S. (2012). Noticing and Uptake: Addressing Pre-Articulated Covert Problems in L2 Writing. *Journal of Second Language Writing, 21*(4), 332-347. https://doi.org/10.1016/j.jslw.2012.09.008
Hyland, K., & Hyland, F. (Eds.). (2006). Feedback in Second Language Writing: Contexts and Issues. Cambridge University Press.
Ivankova, N. V., & Creswell, J. W. (2009). Mixed Methods. In J. Heigham & R. A. Croker (Eds.), Qualitative Research in Applied Linguistics: A Practical Introduction (pp. 135-161). Palgrave Macmillan.
Johnson, M. D. (2021). Noticing and uptake: A meta-analysis of the effects of pre-task and post-task metalinguistic explanation. *Language Teaching Research, 25*(6), 891–916. https://doi.org/10.1177/1362168821994081
Kang, E. (2020). The Efficacy of Model Texts in Written Corrective Feedback: A Study of Chinese EFL Learners. *System, 95*, 102368. https://doi.org/10.1016/j.system.2020.102368
Kellogg, R. T. (1996). A Model of Working Memory in Writing. In C. M. Levy & S. Ransdell (Eds.), The Science of Writing: Theories, Methods, Individual Differences, and Applications (pp. 57-71). Erlbaum.
Khan, S. (2021). Massification of Higher Education in the Middle East: A Bahrain Case Study. *Journal of Higher Education Policy and Management, 43*(2), 210-225. https://doi.org/10.1080/1360080X.2020.1856332
Lantolf, J. P., & Thorne, S. L. (2006). Sociocultural Theory and the Genesis of Second Language Development. Oxford University Press.
Lee, I. (2021). Teachers' Frequently Asked Questions About Written Corrective Feedback. *TESOL Journal*, *12*(3), e00607. https://doi.org/10.1002/tesj.607
Li, Y., & Yuan, F. (2024). The Effects of Model-Based Comparison on Syntactic Complexity and Accuracy in L2 Writing: The Role of Proficiency. *The Modern Language Journal, 108*(1), 223-241. https://doi.org/10.1111/modl.12912
Lu, X., & Liu, Y. (2022). The Effects of Model Texts on L2 Writers' Complexity, Accuracy, and Fluency: A Meta-Analysis. *Studies in Second Language Acquisition, 44*(4), 1073-1100. https://doi.org/10.1017/S0272263121000860
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
Shintani, N., & Ellis, R. (2020). Does Language Analytical Ability Mediate the Effect of Written Feedback on Grammatical Accuracy in L2 Writing? *System, 91*, 102244. https://doi.org/10.1016/j.system.2020.102244
Smith, B. E., Toma, A., & Zhu, Y. (2024). Digital Annotation Tools for Comparative Reflection in L2 Writing. *ReCALL, 36*(1), 78-94. https://doi.org/10.1017/S0958344023000143

Swain, M. (1985). Communicative Competence: Some Roles of Comprehensible Input and Comprehensible Output in Its Development. In S. M. Gass & C. G. Madden (Eds.), Input in Second Language Acquisition (pp. 235-253). Newbury House.

Swain, M. (1995). Three Functions of Output in Second Language Learning. In G. Cook & B. Seidlhofer (Eds.), Principle and Practice in Applied Linguistics: Studies in Honour of H. G. Widdowson (pp. 125-144). Oxford University Press.

Teng, L. S., & Huang, J. (2021). The Effects of Data-Driven Learning Activities on Young EFL Learners' Writing Processes. *Language Teaching Research.* Advance online publication. https://doi.org/10.1177/13621688211024356

Teng, M. F. (2019). The Effects of Video Caption Types and Advance Organizers on Incidental EFL Vocabulary Learning. *Computer Assisted Language Learning, 32*(6), 621-651. https://doi.org/10.1080/09588221.2018.1532911

Tharp, R. G., & Gallimore, R. (1988). Rousing Minds to Life: Teaching, Learning, and Schooling in Social Context. Cambridge University Press.

Tieu, L., & Baker, A. (2022). The Impact of Model Texts on L2 Learners' Written Output. *Language Teaching Research, 26*(5), 1001-1021. https://doi.org/10.1177/1362168821994080

Vygotsky, L. S. (1978). Mind in Society: The Development of Higher Psychological Processes. Harvard University Press.

Winstone, N. E., & Boud, D. (2022). The Need to Disentangle Assessment and Feedback in Higher Education. *Studies in Higher Education, 47*(3), 656-667. https://doi.org/10.1080/03075079.2020.1779687

Zhang, Z., & Hyland, K. (2022). Fostering Student Engagement with Feedback: An Integrated Approach. *Assessment & Evaluation in Higher Education, 47*(2), 213-225. https://doi.org/10.1080/02602938.2021.1908529

Zheng, Y. (2023). The Differential Effects of Pre-Task and Post-Task Model Texts on Chinese EFL Learners' Argumentative Writing. *Reading and Writing, 36*(5), 1267-1290. https://doi.org/10.1007/s11145-022-10322-1

Zheng, Y., & Yu, S. (2022). What Has Been Assessed in Peer Assessment? A Taxonomy of PA Constructs. *Assessment & Evaluation in Higher Education, 47*(3), 372-385. https://doi.org/10.1080/02602938.2021.1924662

Appendices

Appendix A

Writing Task A (Pretest)

Duration: One Hour

"People should read for pleasure."

Write an essay (300-350 words) discussing the statement above. Use the following points and support them with specific examples from your own experience.

- *reading expands knowledge*
- *reading provides fun and improves imagination*
- *reading expands vocabulary and enhances fluency*

……
……

Appendix B
Writing Task B (Post-test)
Duration: One Hour

"Social media has destroyed communication among friends and family."

Write an essay (300-350 words) discussing the statement above. Use the following points and support them with specific examples from your own experience.

- *social media negatively affects communication*
- *social media distracts us from people who are with us*
- *social media can lead to misunderstandings and conflicts*

……
……
……
……
……

Appendix C
Control Group Questionnaire

Directions:
Please indicate the extent to which you agree or disagree with the following statements by selecting the appropriate number.
Scale: 1 = Strongly Disagree 2 = Disagree 3 = Neutral 4 = Agree 5 = Strongly Agree

Section	Question No.	Item	
A. Demographic Data		Full Name	
		ID	
		Age	
		Gener	
		Faculty	
		Mother Tongue Language	
B. Perceived Effectiveness of Model-Based WCF Provided Before Writing	1	The model text provided was helpful in organizing my ideas before writing. 1 2 3 4 5	
	2	The model text made it easier for me to start writing. 1 2 3 4 5	
	3	Using the model text before writing helped me understand correct language usage specific to persuasive writing. 1 2 3 4 5	
	4	I felt more confident writing after seeing the model text. 1 2 3 4 5	
	5	Studying the model text before writing improved my understanding of how to structure a persuasive essay.	
C. Noticing and Awareness	6	The model text helped me notice specific grammatical structures that needs to be used in my persuasive writing. 1 2 3 4 5	
	7	I became more aware of the organization and structure of persuasive writing after studying the model text.	

		1 2 3 4 5
	8	The model text helped me identify vocabulary errors which I could avoid in my writing. 1 2 3 4 5
	9	Reviewing the model text helped me identify areas where my writing could be more coherent. 1 2 3 4 5
	10	Using the model text increased my awareness of sentence variety and complexity. 1 2 3 4 5
	11	The model text helped me recognize correct language usage that I was not previously aware of. 1 2 3 4 5
D. **Engagement with Feedback**	12	While writing, I frequently applied the knowledge I gained from the model text. 1 2 3 4 5
	13	The model text motivated me to structure my essay more carefully. 1 2 3 4 5
	14	I found it easy to apply the knowledge gained from the model text into my own writing. 1 2 3 4 5
	15	I prefer using model texts over other types of feedback. 1 2 3 4 5
	16	I enjoy writing more after using a model text. 1 2 3 4 5
E. **Attitudes and Preferences**	17	Using the model text before the writing task made me feel more prepared. 1 2 3 4 5
	18	The model text helped reduce my anxiety about writing in a Second Language. 1 2 3 4 5
	19	I would like to continue using model texts in future writing tasks. 1 2 3 4 5
	20	Overall, I am satisfied with the feedback provided by the model text. 1 2 3 4 5

Link on Google Form https://forms.gle/qBbYubkF6Jj3jh5h9

Appendix D
Experimental Group Questionnaire

Directions:
Please indicate the extent to which you agree or disagree with the following statements by selecting the appropriate number.
Scale: 1 = Strongly Disagree 2 = Disagree 3 = Neutral 4 = Agree 5 = Strongly Agree

Section	Question No.	Item	
A. Demographic Data		Full Name	
		ID	
		Age	
		Gener	
		Faculty	
		Mother Tongue Language	
B. Perceived Effectiveness of Model-Based WCF Provided After Writing	1	After seeing the model text, I could easily compare it to my own writing. 1 2 3 4 5	
	2	The model text helped me understand how to improve my writing after the first draft. 1 2 3 4 5	
	3	The feedback from the model text was more useful than traditional teacher corrections. 1 2 3 4 5	
	4	Using a model text after writing improved my understanding of correct language usage specific to persuasive writing. 1 2 3 4 5	
	5	I feel that my writing has improved as a result of comparing it with a model text. 1 2 3 4 5	
C. Noticing and Awareness	6	The model text helped me notice specific grammatical errors in my own writing. 1 2 3 4 5	
	7	I became more aware of the organization and structure of a persuasive essay after comparing my own writing with the model text. 1 2 3 4 5	
	8	I was able to avoid making the same mistakes in my revised draft after using the model text. 1 2 3 4 5	
	9	Using the model text increased my awareness of sentence variety and complexity. 1 2 3 4 5	
	10	The model text highlighted correct language usage that I was not previously aware of. 1 2 3 4 5	
D. **Engagement with Feedback**	11	While revising my writing, I frequently applied the knowledge I gained from the model text. 1 2 3 4 5	
	12	The model text motivated me to make more careful revisions. 1 2 3 4 5	
	13	I found it easy to apply the feedback from the model text into my own writing. 1 2 3 4 5	
	14	I felt more confident revising my own writing when comparing it with the model text.	

		1 2 3 4 5
	15	I prefer using model texts over other types of feedback. 1 2 3 4 5
E. **Attitudes and Preferences**	16	I enjoy writing more after using a model text. 1 2 3 4 5
	17	Using a model text after the writing task made me feel more confident about my revisions. 1 2 3 4 5
	18	The model text helped reduce my anxiety about making errors in my writing. 1 2 3 4 5
	19	I would like to continue using model texts for feedback in future writing tasks. 1 2 3 4 5
	20	Overall, I am satisfied with the feedback provided by the model text. 1 2 3 4 5

Link on Google Form https://forms.gle/kFemhzi9ZVQVL7sB7

Arab World English Journal

INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp. 479-492
DOI: https://dx.doi.org/10.24093/awej/vol16no4.26

AI Tools in Academic Research: Perceptions, Adoption, and Challenges among Moroccan Student Researchers

Lamiae Mamoune
Department of English Studies
Sidi Mohamed Ben Abdellah University, Fez, Morocco
Corresponding Author: lamiae.mamoune@gmail.com

Safae Kharmach
Department of English Studies
Sidi Mohamed Ben Abdellah University, Fez, Morocco
Email: safaekharmach00@gmail.com

Abderrahim El Karfa
Department of English Studies
Sidi Mohamed Ben Abdellah University, Fez, Morocco
Email: elkarfa@yahoo.fr

Mohammed Bennis
Department of English Studies
Sidi Mohamed Ben Abdellah University, Fez, Morocco
Email: mohammed.bennis3@usmba.ac.ma

Received:08/14/2025 Accepted:11/25/2025 Published:12/10/2025

Abstract:
The evolving landscape of academic research worldwide has seen a new revolution with the development of artificial intelligence (AI). This article illuminates the perceptions, adoption and challenges of the use of AI tools among Moroccan student researchers. This study employed a cross-sectional quantitative approach for the collection and interpretation of data. A structured questionnaire was administered to a body of student researchers enrolled in master's and PhD programs across disciplines within the University of Sidi Mohamed Ben Abdellah of Fez. Built on the valid responses of 100 participants, the data obtained was analysed statistically with SPSS. The data analysis's results were divided into sections, revealing demographic information, then descriptive data, while also exploring correlations. The findings show that PhD students constituted the majority of respondents (n=69) and that overall participants have positive attitudes towards the integration of AI in their academic research. Participants frequently embed AI tools in their work with a remarkable preference for two AI tools: ChatGPT and Grammarly AI. This study also revealed that student researchers primarily adopt AI for information retrieval, idea generation, language refinement, and citation management, with limited use in advanced tasks such as data analysis and translation. Nevertheless, participants are aware of reliability and ethical concerns, and express an interest in further use of AI, yet not at the expense of diminished quality of writing and analysis. This study calls for guidelines and regulations for the responsible use, as well as addresses emerging issues such as AI citation in academic research.

Keywords: Academic research, artificial intelligence, adoption, challenges, Moroccan context, perceptions, research practices, student researchers
How to cite: Mamoune, L., Kharmach, S., El Karfa, S., & Bennis, M. (2025).AI Tools in Academic Research: Perceptions, Adoption, and Challenges among Moroccan Student Researchers. *Arab World English Journal, 16*(4): 479-492. https://dx.doi.org/10.24093/awej/vol16no4.26

Introduction

Among the growing technologies, artificial intelligence (AI) has gained popularity in academia within different fields. With the ability to perform human-like tasks, saving energy and time (Livberber & Ayvaz, 2023), AI is increasingly infiltrating research in its different stages. Considering that academic research involves handling a considerable amount of information, theories, and data, researchers often seek tools that will facilitate these operations (Khalifa & Albadawi, 2024). Various tools for academic research optimise the performance of specific tasks such as information retrieval, writing, literature reviewing, tracking citations, and analysing data. It is therefore apparent that artificial intelligence is revolutionising the field of academic research.

The scope of artificial intelligence is not limited to a particular field; it has a broader reach, as it can be used in the medical field, the agricultural field, finance, science, humanities, and more (Moukhliss et al., 2024). AI has thus triggered a radical shift in the way research is conducted. Researchers are increasingly leveraging artificial intelligence tools to advance their studies, leading to an upsurge in academic tools. We can cite many that are popular with researchers, for instance, ChatGPT, Grammarly AI, Semantic Scholar, Zotero, Mendeley, Jenny AI, and Quillbot, each facilitating the performance of a certain task.

However, amidst this evolution in research, other aspects of AI are to be considered. Therefore, this study seeks to explore the perceptions of researchers, and especially student researchers, toward the use of artificial intelligence in academic work, the perceived benefits and the challenges faced. Conducting the current study will serve to identify the extent to which artificial intelligence is implemented in student researchers' studies, as well as inform policymakers on how to best support the use of AI in academic research, especially in the Moroccan context.

Literature Review

The integration of artificial intelligence in academic research has instigated debate and controversies surrounding its validity and potential (Aich et al., 2025). Advocates for the use of artificial intelligence in academia defend AI technologies for the support they provide in science discovery, that is, literature management and synthesis, easy access to databases and the wealth of what has already been discovered and even data analysis and interpretation (Hastings, 2023). Indeed, many researchers have shown a newfound interest in the various AI tools that have reshaped the way research is conducted, which range from academic writing to actual research processes, including information retrieval, peer review and feedback on academic work (Aich et al., 2025). According to Madanchian and Taherdoost (2025): "the integration of AI into research processes offers substantial benefits, including significant reductions in time spent on tasks such as literature screening and data extraction" (p. 9).

As researchers encounter issues with academic writing, the use of chatbots and AI tools reduces the burden by serving as personal assistants, assisting with writing tasks, handling errors, and enhancing language (Kamal & Khaidzir, 2024). Thus, it is apparent that efficiency in research is highly required to augment production, foster creativity and facilitate processes (Madanchian & Taherdoost, 2025). Bouzar et al. (2024) view AI technologies like GPT-3 and ChatGPT as becoming easy to reach, scalable, and practical. ChatGPT is remarkable for its ability to generate human-like conversations and provide thorough responses to complex questions, as well as assist in promoting programs, customer service, and business. Nevertheless, they acknowledge that, at times, it lacks understanding of context, critical thinking, and ethical decision-making, and can inadvertently disseminate societal biases.

However, despite the upsurge of AI tools that perform various tasks, there are still concerns that arise and issues that researchers need to consider, with ethical concerns being a major one. Indeed, the use of artificial intelligence tools is bound to their trustworthiness, reliability, and conformity with ethical standards (Daoudi, 2025). These concerns are reasonable since the very essence of academic research is rooted in credibility, reliability, and transparency. Regardless of the agreed-upon merits of artificial intelligence, research has shown that academics fear diminished quality of writing as well as plagiarism, with many worrying it would re-conceptualise practices of creativity, research integrity and originality (Malik et al., 2024; Perkins & Roe, 2024). Therefore, it is of dire importance to make sure that using AI would not result in undermining the worth of academic work while still profiting from the numerous services it provides. Dahal (2014) posits that "Transparency in the use of GenAI and/or AI in research and the limitations of these tools are also critical ethical requirements" (p. 727).

The steady growth in the number of AI tools targeted at facilitating research has led to a proliferation of studies investigating its potential in academic writing, data analysis, and other research processes. For instance, Yan et al. (2024) conducted a user study involving ten qualitative researchers to explore the potential of ChatGPT as a collaborative tool in thematic analysis. Their findings demonstrated that ChatGPT had value in enhancing coding efficiency, facilitating initial data exploration, providing granular quantitative insights, and aiding comprehension for non-native speakers and non-experts. On the other hand, the researchers also identified persistent issues regarding ChatGPT's trustworthiness, accuracy, reliability, and contextual understanding, as well as concerns about its broader acceptance within the research community. Similarly, Gao and Wang (2023) conducted a study on quantifying the benefit of artificial intelligence for scientific research. They found that disciplines overall used more AI in their research over the past two decades. This increase occurred not only in computer science but also in a wide range of other disciplines, including physics, biology, and economics. Moreover, this increase had not been linear; there had been a notably sharp increase in the use of AI since 2015 across many disciplines.

Even though research on the use of AI tools in academic work is rapidly expanding, there is still limited empirical evidence on how student researchers in developing contexts, particularly in Morocco, perceive and adopt AI in their academic work. While various studies have highlighted potential benefits and ethical concerns, little is known about the extent of adoption, perceived usefulness, and challenges faced by the target population in Moroccan higher education. This gap restricts institutions from developing informed strategies and policies for responsible AI integration in research training. Therefore, three research objectives have been set to address this gap in the literature:

1. To investigate Moroccan student researchers' perceptions of AI tools in academic research.
2. To measure the extent and purposes of AI tools' adoption in academic research processes.
3. To identify the challenges and ethical concerns student researchers face when using AI tools.

Method

Research Design

The current study explores the perceptions, adoption and challenges of the integration of AI tools in academic research among Moroccan student researchers. Therefore, this study has adopted a quantitative approach to allow for systematic measurement of variables and gauging relationships amongst them, as well as conducting statistical analysis, and generalisation of

findings within the study population (Creswell & Creswell, 2017). This design was cross-sectional in nature as it sought to capture the researchers' perceptions at a single point in time.

Participants

This study was conducted during the spring semester of the academic year 2024/2025 at the University of Sidi Mohamed Ben Abdellah of Fez. The target population of this study were students of higher education, specifically master's and doctoral students across disciplines. The sample consisted of 100 Moroccan student researchers. The majority were PhD candidates (69%), with the remaining 31% at the Master's level. Gender distribution varied across age groups: the 20–25 group was predominantly female (79.3%), while the 31+ group was mostly male (84.1%) (see Table 1 and Figure 1).

Table 1. *Age group distribution (N = 100)*

	N	%
20-25	29	29.0%
26-30	27	27.0%
31 and above	44	44.0%

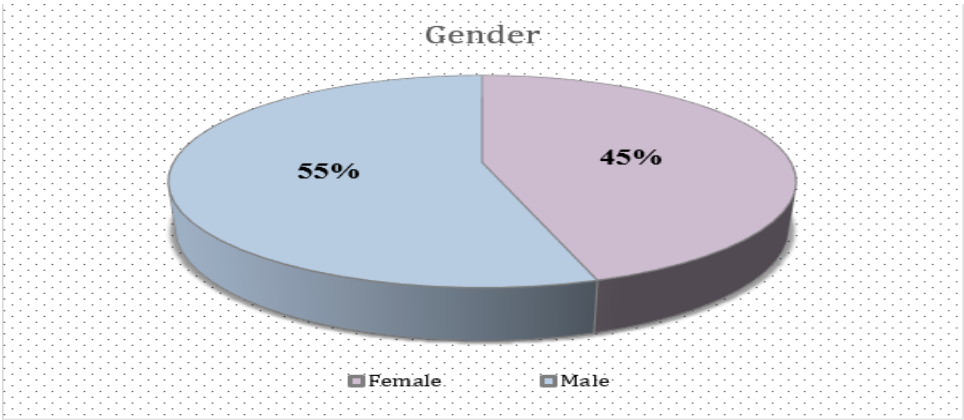


Figure 1. Gender distribution of participants (N = 100)

Instruments

This study employed a structured questionnaire as the primary data collection instrument, focusing on both descriptive and inferential statistics to identify patterns and relationships among participants' demographics, perceptions, and experiences with AI tools. The items explored in the questionnaire were built on the existing literature and adapted to the Moroccan context. The questionnaire was designed using Google Forms and was shared widely with graduate and postgraduate students through social media platforms. The questionnaire came back with a total of 100 valid responses that were coded and analysed using the Statistical Package for the Social Sciences (SPSS).

Results

Demographic Information

For the age group distribution, the age of participants was coded into three categorical groups. Most participants were aged 31 and above (44%), followed by 20–25 (29%) and 26–30 (27%). The average age group was between 26–30 and 31+ (M = 2.15, SD = 0.85), indicating a slightly older academic sample, with a median category of 2.00 (see Table 1).

Figure 2 shows the distribution of participants by area of study. The study targeted Moroccan student researchers enrolled in graduate and postgraduate programs across different academic disciplines. While the sample is predominantly composed of students from the Humanities, the inclusion of participants from a variety of academic backgrounds adds diversity to the dataset, allowing for a more nuanced understanding of how AI tools are perceived across different fields of study.

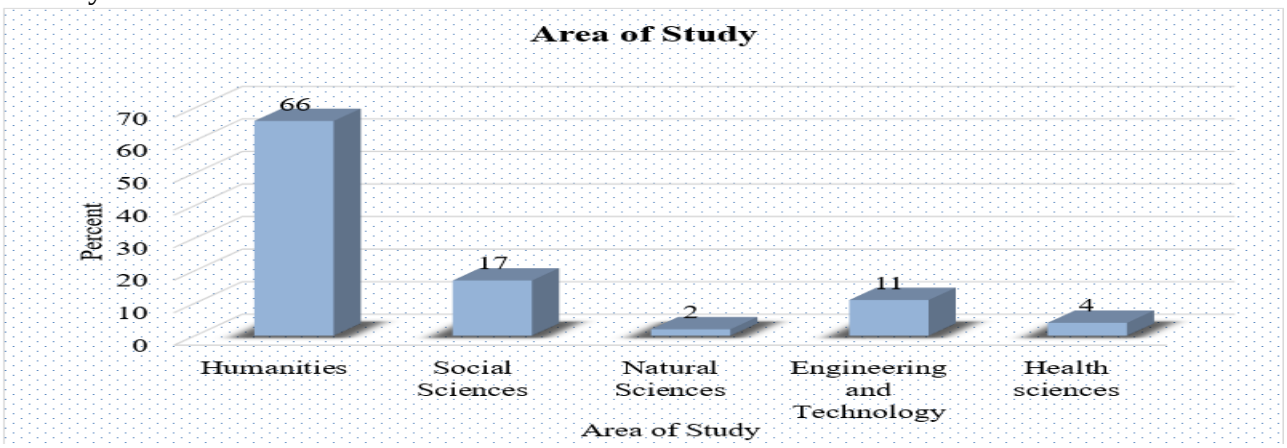


Figure 2. Distribution of participants by area of study (N = 100)

Frequency of AI tool usage

Figure 3, which shows the frequency of AI tool usage among student researchers, indicates a clear trend toward the regular and habitual integration of AI tools into academic research workflows. With over 60% of participants using AI tools either daily or several times a week, it is evident that these technologies are no longer peripheral – they are becoming central to students' research practices.

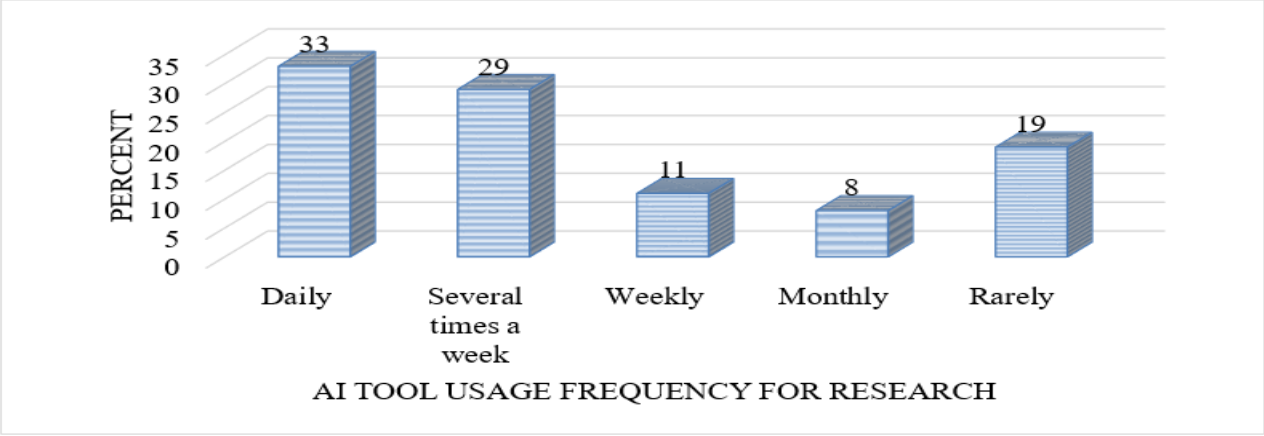


Figure 3. Frequency of AI tool usage for research among participants (N = 100)

ISSN: 2229-9327

Among the AI tools that were presented for participants to select from, the bar chart clearly highlights that ChatGPT dominates usage among student researchers (89%), followed by Grammarly AI (40%) and Quillbot (36%). The remaining tools were selected by fewer than 30% of participants, indicating limited adoption.

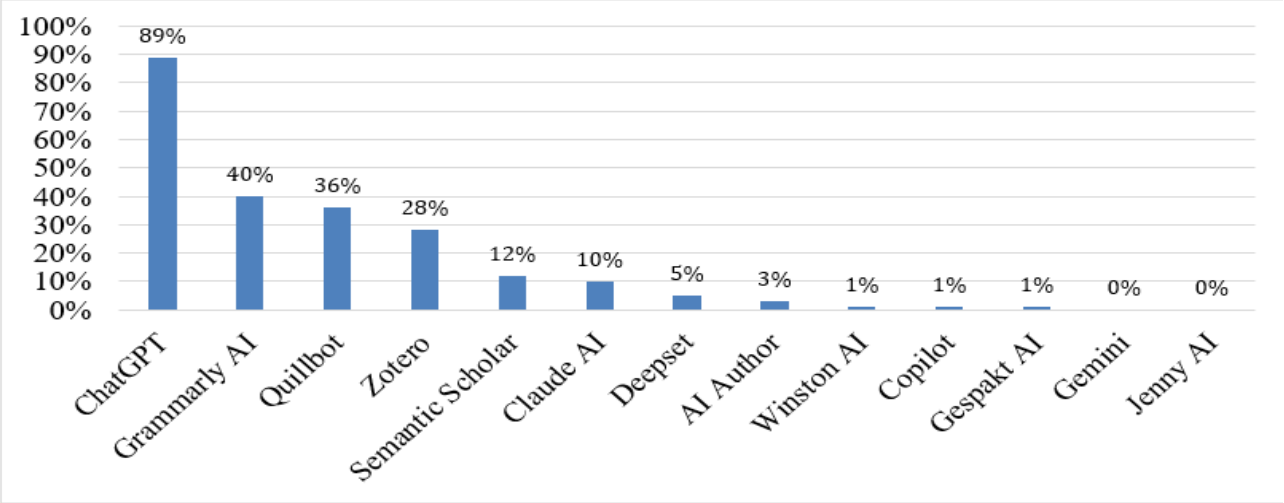


Figure 4. AI tools most frequently used by researchers

The data suggest that students rely on AI mainly for idea generation, language support, and research organisation, rather than for advanced analytical or translation tasks.

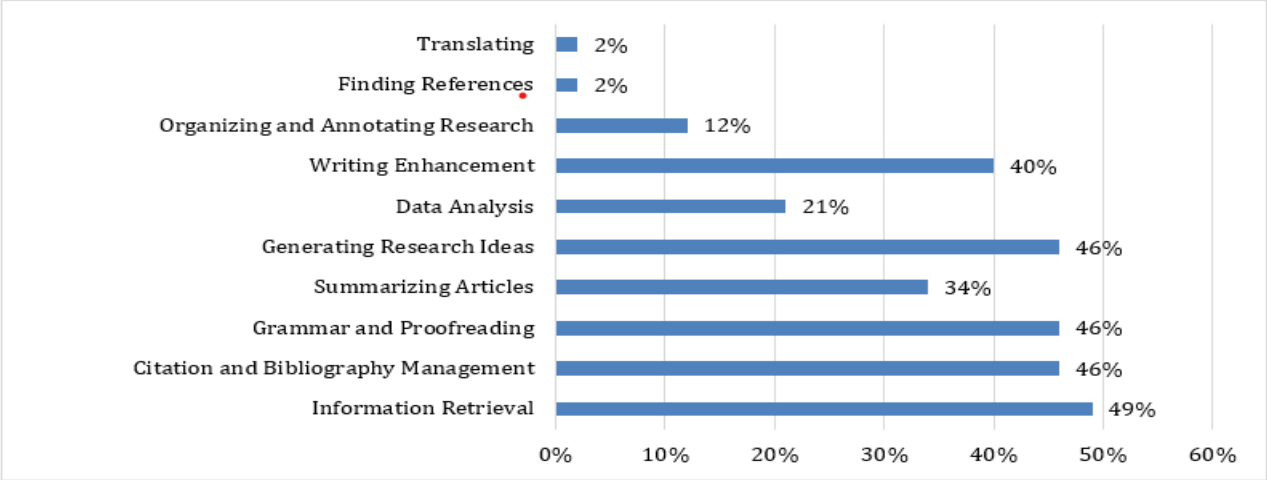


Figure 5. Participants' selected purposes for using AI tools in research

Perceived Benefits of AI Tools in Academic Research

The following figure presents the results of the Likert scale questions that sought to examine participants' perceived benefits of AI tools in Academic Research.

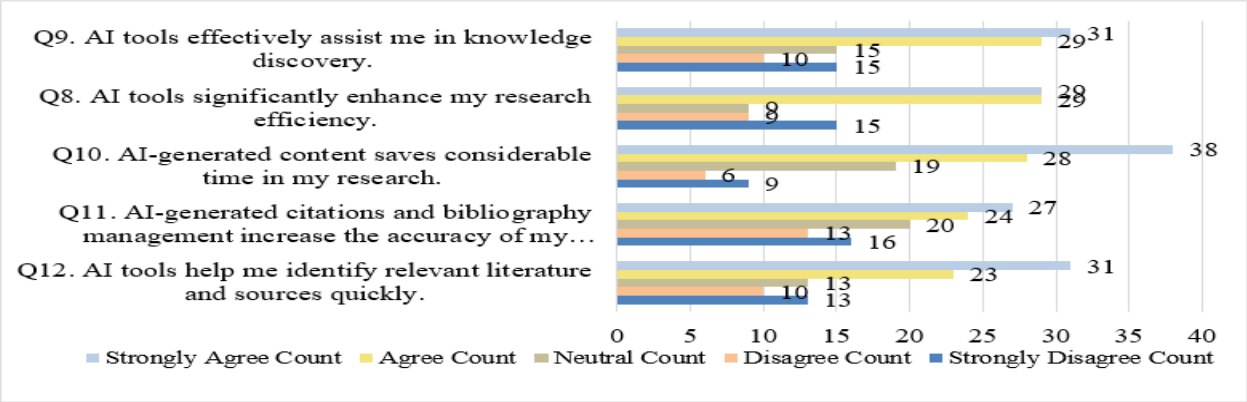


Figure 6. Participants' perceived benefits of AI tools in academic research

Participants' Satisfaction with AI Tools

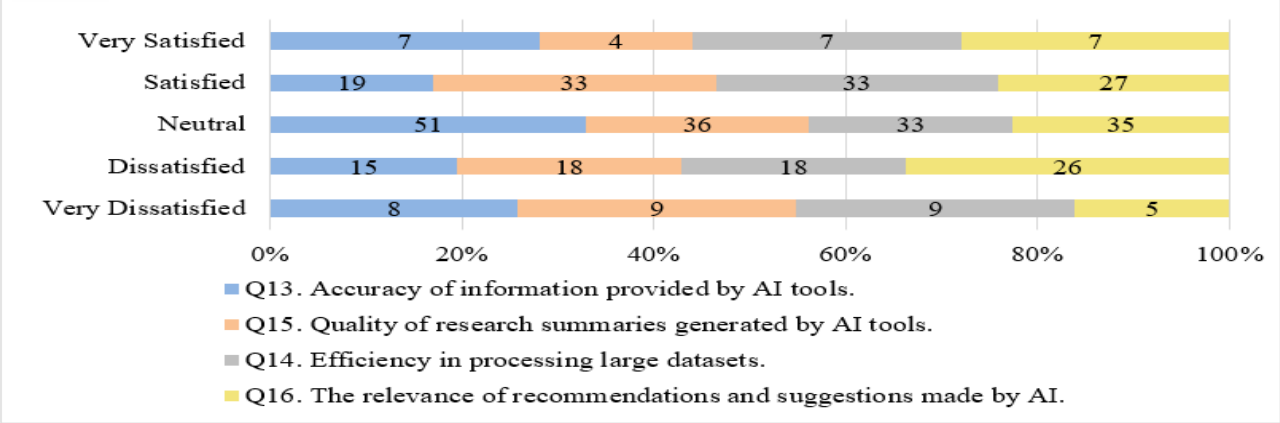


Figure 7. Participants' satisfaction with AI Tools

Challenges and Ethical Concerns

The responses of participants have reported an acute awareness of issues that arise with the continuous use of AI tools, as many have selected strongly agree with the statements presented in the Likert scale. However, statement Q18 has attracted a higher count of agreement, which shows that student researchers are mindful of the issues related to the reliability and accuracy of AI-generated insights.

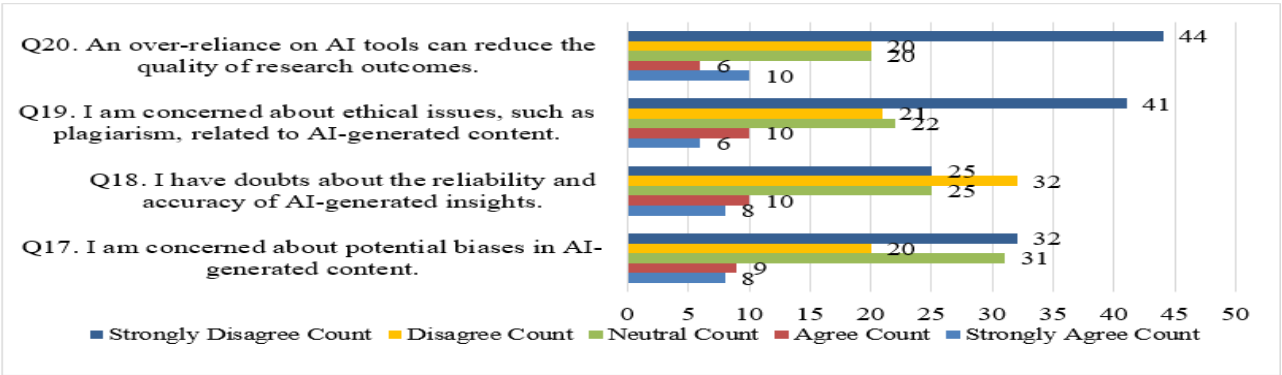


Figure 8. Students' concerns about the reliability and risks of AI tools in research

ISSN: 2229-9327

Integration and future use of AI Tools in Academic Research

Participants expressed overwhelmingly positive attitudes regarding the integration and future use of AI tools in academic research (see Figure 9). These findings align with the broader theme of AI receptiveness, while also reinforcing the need for supportive academic infrastructure to maximise AI's value.

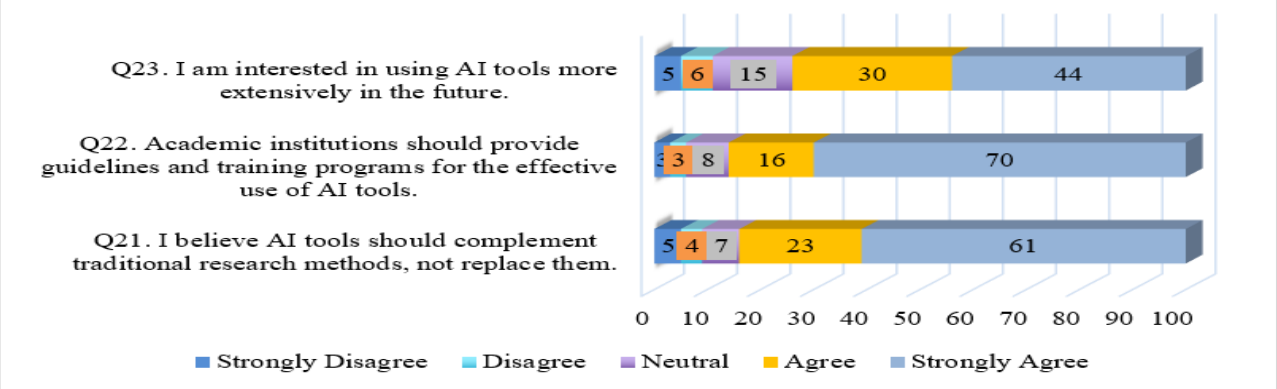


Figure 9. Student attitudes toward the integration and future use of AI tools in academic research

Correlations

Table 2. *Correlation between age group and AI tool usage frequency for research*

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	17.461[a]	8	.026
Likelihood Ratio	19.265	8	.014
Linear-by-Linear Association	11.899	1	.001
N of Valid Cases	100		

a. 6 cells (40.0%) have expected count less than 5. The minimum expected count is 2.16.

A chi-square test of independence was conducted to examine the relationship between age group and AI tool usage frequency. The association was statistically significant, $\chi^2(8, N = 100) = 17.46$, $p = .026$, indicating that age is related to how often students use AI tools for research. Younger participants (20–25) were more likely to use AI tools daily, while older participants (31 and above) tended to report less frequent usage. The Linear-by-Linear Association also shows a clear directional trend, suggesting that as age increases, the frequency of daily AI tool usage may decline.

An independent-sample t-test was conducted to examine whether perceived benefits of AI tools differed between participants who use them frequently (daily/several times a week) and those who use them infrequently (rarely/monthly). The results revealed statistically significant differences on all five benefit items, suggesting that the frequency of AI tool usage is positively associated with higher perceived benefits in academic research. Participants who use AI tools more frequently tend to rate them more favorably across all dimensions.

Discussion

The findings of the current study capture the perceptions of student researchers toward the adoption of AI tools in academic research. Although the sample size is not large, it still

encapsulates attitudes and perceptions of the target population; it was also relatively balanced in terms of gender, with females slightly outnumbering males.

The data set shows that the sample of student researchers across disciplines has positive attitudes toward the integration of AI in academia, as they employ it frequently in their own works of research. They are also aware of the various tools available and how beneficial they are to their own work. However, they still show a preference to ChatGPT first, due to its generative nature and its highly conversational, adaptable interface and Grammarly AI second, for its assistance in improving writing skills. These inclinations reveal that participants lean more towards human-like content generated by conversational and natural language prompts (as stated by Bouzar et al., 2024). It also reveals that critical attention is paid to accuracy in academic writing.

The high frequency of AI usage and high levels of satisfaction with their performance suggest that AI tools are valued not just for their convenience but also for their perceived effectiveness in enhancing productivity, simplifying complex tasks, polishing writing, proofreading, and accelerating knowledge access. These findings correspond to the existing body of knowledge as they confirm the arguments advanced by Kamal & Khaidzir (2024) and Madanchian & Taherdoost (2025). They also show that participants are active and critical evaluators of AI tools, given their selective adoption and awareness of challenges and ethical concerns that arise with continuous reliance on artificial intelligence. Nevertheless, the limited use of advanced applications of AI tools in tasks such as translation, data analysis and finding references suggests that student-researchers use AI tools for foundational rather than specialised and analytical purposes. This caution stems from the reliability and ethical concerns highlighted previously, as well as risks of fabricated content.

Moreover, the presence of less frequent users (weekly, monthly, or rarely) reveals that while adoption is widespread, it is not yet universal. This could reflect disciplinary differences, varying digital literacy levels, or differing institutional support for AI integration. Taken together, the results confirm one of the core premises of the study: that while AI tools are broadly accepted and frequently used by student researchers, there remains a need for structured guidance, training, and critical understanding to ensure their use is effective, ethical, and equitable across research communities.

Conclusion

The ongoing debate about the integration of artificial intelligence in academic research raises multiple questions. The findings of the current study present an overview of the perceptions, adoption and challenges in the use of AI tools in academic research among Moroccan student researchers. This study targeted a specific population of researchers, a group that represents the emerging generation of academics and scholars in Moroccan higher education. Therefore, their insights serve as valuable data for a critical study of the integration of AI in academic research. This study highlighted an existing and widespread use of AI tools among these students, with varying frequencies among age groups. The overall acceptance and recognition of the benefits of using artificial intelligence, coupled with the awareness of its risks and challenges, reveal that student researchers are enthusiastic adopters of artificial intelligence, yet do not fully consider it as a research partner but rather an assistant.

Participants' recommendations for future use and integration of AI in research emphasise that it is welcome, albeit contingent upon proper training, clear institutional policies, and responsible

practices. More importantly, the data underscore the participants' view of AI as complementary rather than a replacement for traditional research practices.

The perspective of student researchers regarding this issue is of great value for mapping the current situation of academic research in Moroccan higher education. This study revealed much about the adoption and patterns of AI use, but above all, it set the ground for future research on the regulations for AI use and anticipated how the next wave of researchers will integrate AI into academic research. The findings of this study also raise questions about proper AI citation and the recognition of its contribution while maintaining academic integrity.

Funding: This research is not funded.

Acknowledgement: The corresponding author Lamiae Mamoune wishes to acknowledge the CNRST (Centre National de la Recherche Scientifique et Technique) for the PhD Associate Scholarship (PASS), awarded in support of the author's academic training.

Conflicts of Interest: The authors declare no conflict of interest.

Authenticity: This manuscript is an original work.

Artificial Intelligence Statement: During the preparation of this work, the author(s) used AI and AI-assisted technologies, specifically ChatGPT- 4.0 in order to enhance the language and increase readability. The authors reviewed and edited the content as needed and they take full responsibility for the content of the publication.

About the authors:

Lamiae Mamoune, PhD candidate at the University of Sidi Mohammed ben Abdallah, Faculty of Letters and Humanities Dhar El Mehraz, Fez, Morocco. She is part of the PhD associate scholarship granted by CNRST. She is working on the teaching of literature in the digital age. Her research interests also include applied linguistics, media studies and artificial intelligence.
ORCID: https://orcid.org/0009-0007-9126-187X

Safae Kharmach, a PhD candidate at the University of Sidi Mohammed ben Abdallah, Faculty of Letters and Humanities Dhar El Mehraz, Fez, Morocco. Her research focuses on the Enhancement of Academic Orientation for Improved Student Preparedness and Success. She harbors a special interest in technology, AI, and its intersection with the educational field.
ORCID: https://orcid.org/0009-0007-0415-9966

Abderrahim El Karfa, Professor of Applied Linguistics at the Faculty of Letters and Human Sciences, Dhar El Mehraz, Sidi Mohammed Ben Abdellah, Fes, Morocco. Research Interests: Applied Linguistics, Leadership Education, Critical Thinking Education, Gender and Language, Gender and Education, English Language Education, AND lifelong Learning.
ORCID: https://orcid.org/0000-0002-5927-2575

Mohammed Bennis, Associate Professor at Sidi Mohammed Ben Abdellah University, Faculty of Letters & Human Sciences, Fes, Morocco, Department of English Studies. Research interests:

Media Studies, Digital Humanities, Cybercultures, Cultural Studies, Literary Studies, Cultural Anthropology. ORCID: https://orcid.org/0009-0009-6214-7757

References

Aich, F. Z., Boumahdy, L., Benfilali, I., & Sandy, K. (2025). AI Usage in Academic Writing Among Moroccan Doctoral Researchers at Sidi Mohamed Ben Abdellah University. *Arab World English Journal (AWEJ) Special Issue on CALL, 11*, 192–209. https://doi.org/10.24093/awej/call11.12

Bouzar, A., El Idrissi, K., & Ghourdou, T. (2024). ChatGPT and Academic Writing Self-Efficacy: Unveiling Correlations and Technological Dependency among Postgraduate Students. *Arab World English Journal (AWEJ) Special Issue on ChatGPT, (April 2024)*, 225–236. https://doi.org/10.24093/awej/ChatGPT.15

Creswell, J. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (5th ed.)*. Thousand Oaks, CA: SAGE Publications.

Dahal, N. (2024). How Can Generative AI (GenAI) Enhance or Hinder Qualitative Studies? A Critical Appraisal from South Asia, Nepal. *The Qualitative Report, 29(3)*, 722–733. https://doi.org/10.46743/2160-3715/2024.6637

Daoudi, M. (2025). Ethical limits and suggestions for improving the use of AI in scientific research, academic publishing, and the peer review process, based on deontological and consequentialist viewpoints. *Discover Education, 4*, 241. https://doi.org/10.1007/s44217-025-00696-z

Gao, J., & Wang, D. (2023). Quantifying the benefit of artificial intelligence for scientific research. *arXiv preprint, arXiv:2304.10578*. https://doi.org/10.48550/arXiv.2304.10578

Hastings, J. (2023). *AI for Scientific Discovery*. CRC Press. https://doi.org/10.1201/9781003226642

Kamal, M. A. A., & Khaidzir, M. F. S. (2024). Publish or Perish: The Impact of ChatGPT on Future Academics. *Arab World English Journal (AWEJ) Special Issue on ChatGPT, (April 2024)*, 319–330. https://doi.org/10.24093/awej/ChatGPT.22

Khalifa, M., & Albadawy, M. (2024). Using artificial intelligence in academic writing and research: An essential productivity tool. *Computer Methods and Programs in Biomedicine Update, 5*, 100145. https://doi.org/10.1016/j.cmpbup.2024.100145

Livberber, T., & Ayvaz, S. (2023). The impact of Artificial Intelligence in academia: Views of Turkish academics on ChatGPT. *Heliyon, 9(9)*, e19688. https://doi.org/10.1016/j.heliyon.2023.e19688

Madanchian, M., & Taherdoost, H. (2025). The impact of artificial intelligence on research efficiency. *Results in Engineering, 26*, 104743. https://doi.org/10.1016/j.rineng.2025.104743

Malik, M. A., Amjad, A. I., Aslam, S., & Fakhrou, A. (2024). Global insights: ChatGPT's influence on academic and research writing, creativity, and plagiarism policies. *Frontiers in Research Metrics and Analytics, 9*, 1486832. https://doi.org/10.3389/frma.2024.1486832

Moukhliss, G., Lahyani, K., & Diab, G. (2024). The Impact of Artificial Intelligence on Research and Higher Education in Morocco. *Journal of Education and Learning (EduLearn), 18(4)*, 1292–1300. https://doi.org/10.11591/edulearn.v18i4.21511

Perkins, M., & Roe, J. (2024). Generative AI tools in academic research: Applications and implications for qualitative and quantitative research methodologies. *arXiv preprint, arXiv:2408.06872*. https://doi.org/10.48550/arXiv.2408.06872

Yan, L., Echeverria, V., Fernandez-Nieto, G. M., Jin, Y., Swiecki, Z., Zhao, L., Gašević, D., & Martinez-Maldonado, R. (2024). Human–AI collaboration in thematic analysis using ChatGPT: A user study and design recommendations. *In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '24)* (pp. 1–7). ACM, Honolulu, HI. https://doi.org/10.1145/3613905.3650732

Appendix

Students' researchers' questionnaire:

AI Tools in Academic Research: Perceptions, Adoption and Challenges among Moroccan Student Researchers

The advent of digital technologies and especially the emergence of artificial intelligence (AI) has reshaped the field of academia, presenting tools that facilitate information retrieval, archiving and knowledge discovery. There is a multitude of AI tools available for academics that enhance the research experience, such as ChatGPT, Claude, Grammarly AI, Semantic Scholar, Zotero, Mendeley, Winston AI, Jenny AI, AI Author and Quillbot. These AI tools help in different stages of research, and each offers certain functionalities that benefit the researcher, ranging from conversation and bibliography generators to writing assistants. However, as scholars make use of artificial intelligence in academic research, the line between reaping its benefits and abusing it blurs, as issues of ethical considerations, bias and reliability arise. Therefore, the purpose of the current study is to investigate the different perceptions of student researchers regarding the use of AI-powered tools in research, with a focus on the adoption, perceived benefits, and challenges.

Your contribution to this study is highly appreciated. It is voluntary, and each response is anonymous. The Data collected is guaranteed to be kept private and will only be used for academic purposes. Results obtained from analysing this data may be published in articles or presentations. Thank you again for your time and contribution to this research. Your input is highly valuable in shaping the future of academic research.

Section 1: Demographic Information

1. Gender
   - o Male
   - o Female
2. Age group
   - o 20-25
   - o 26-30
   - o 31 and above
3. Level of Study
   - o Master's
   - o PhD
4. Area of Study
   - o Humanities
   - o Social Sciences

- Natural Sciences
- Engineering and Technology
- Other :

5. University/ Higher Education institution

_____

Section 2: AI Tool Usage Frequency and Purpose

6. How frequently do you use AI tools for research?

- Daily
- Several times a week
- Weekly
- Monthly
- Rarely
- Never

7. Select the AI tools you regularly use for your research (Select all applicable)
You can select more than one answer.

- ChatGPT
- Claude AI
- Grammarly AI
- Semantic Scholar
- Zotero
- Mendeley
- Winston AI
- Jenny AI
- AI Author
- Quillbot

8. What are your primary purposes for using AI tools in research? (Select all applicable)
You can select more than one answer.

- Information retrieval
- Citation and bibliography management
- Grammar and proofreading
- Summarizing articles
- Organising and annotating research
- Generating research ideas
- Data analysis
- Writing enhancement

Other: _____

Section 3: Perceived Benefits of AI Tools

(1 = Strongly Disagree, 5 = Strongly Agree)

9. AI tools significantly enhance my research efficiency: [1] [2] [3] [4] [5]
10. AI tools effectively assist me in knowledge discovery. [1] [2] [3] [4] [5]
11. AI-generated content saves considerable time in my research. [1] [2] [3] [4] [5]
12. AI-generated citations and bibliography management increase the accuracy of my references. [1] [2] [3] [4] [5]
13. AI tools help me identify relevant literature and sources quickly. [1] [2] [3] [4] [5]

14. Accuracy of information provided by AI tools. [1] [2] [3] [4] [5]
15. Efficiency in processing large datasets. [1] [2] [3] [4] [5]
16. Quality of research summaries generated by AI tools. [1] [2] [3] [4] [5]
17. The relevance of recommendations and suggestions made by AI. [1] [2] [3] [4] [5]

Section 5: Challenges and Ethical Concerns

(1 = Strongly Disagree, 5 = Strongly Agree)

18. I am concerned about potential biases in AI-generated content. [1] [2] [3] [4] [5]
19. I have doubts about the reliability and accuracy of AI-generated insights. [1] [2] [3] [4] [5]
20. I am concerned about ethical issues, such as plagiarism, related to AI-generated content. [1] [2] [3] [4] [5]
21. Over-reliance on AI tools can reduce the quality of research outcomes. [1] [2] [3] [4] [5]

Section 6: Recommendations and Future Use

(1 = Strongly Disagree, 5 = Strongly Agree)

22. I believe AI tools should complement traditional research methods, not replace them. [1] [2] [3] [4] [5]
23. Academic institutions should provide guidelines and training programs for the effective use of AI tools. [1] [2] [3] [4] [5]
24. I am interested in using AI tools more extensively in the future. [1] [2] [3] [4] [5]
25. Please share any further suggestions and recommendations for improving the responsible use of AI tools in research.

Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 4, December 2025 Pp. 493 -514
DOI: https://dx.doi.org/10.24093/awej/vol16no4.27

Autonomous Digital Language Learning Strategies of EFL Students Beyond the Classroom

Hanadi A. Khadawardi
King Abdulaziz University, English Language Institute
Jeddah, Saudi Arabia
Email: hakhadawardi@kau.edu.sa

Received: 08/22/2025 Accepted:12/07/2025 Published:12/10/2025

Abstract

This study investigates the autonomous digital language learning practices of English as a Foreign Language (EFL) learners outside traditional classroom settings, with a focus on the roles of motivation, academic tracks, and English proficiency levels. Using a mixed-methods approach, the research examines how learners incorporate digital tools into their language learning, the types of activities they engage in, and the factors shaping their autonomy. The findings show that informal digital activities, such as listening to music, watching TV, and browsing the internet, are the most common, largely driven by intrinsic motivation. In contrast, more structured activities, such as attending self-study centers or writing diaries, are less frequently pursued. The study underscores motivation as a key factor in fostering learner autonomy and recommends educational strategies that promote both informal and structured digital learning. These insights offer valuable implications for enhancing learner autonomy and improving language learning outcomes in an increasingly digitalized educational environment.

Keywords: Autonomous language learning, learning beyond classroom, digital learning strategies, English as a foreign language, motivation, technology integration, autonomous digital learning strategies

How to cite: Khadawardi, H. (2025). Autonomous Digital Language Learning Strategies of EFL Students Beyond the Classroom. *Arab World English Journal*, *16*(4): 493 -514. https://dx.doi.org/10.24093/awej/vol16no4.27

Introduction

Autonomous learning is a multifaceted concept in language education that encompasses learner independence, self-directed learning, and autonomous engagement in the learning process. Rooted in constructivist theory, which emphasizes active learner participation, this pedagogical approach also draws on humanistic and cognitive psychology, focusing on self-actualization through educational experiences. Recent shifts towards learner-centered approaches have underscored the critical importance of autonomy in language education (Benson & Reinders, 2023; Ushioda, 2022). The advent of digital technologies has profoundly transformed language learning, providing extensive opportunities for fostering autonomy. However, the integration of autonomous learning practices varies significantly across different contexts particularly in regions such as Saudi Arabia and other Asian English-as-a-Foreign-Language (EFL) settings, where traditional teacher-centered methods often prevail (Wang et al., 2024). Promising developments have been observed in promoting learner autonomy through targeted training and structured support. For instance, learners in EFL contexts have shown substantial improvements in autonomy when provided with appropriate autonomy-supportive resources and guidance (Rahmasari, 2025; Noori, 2025). Advancing learner autonomy requires sustained efforts to empower students, integrate technology effectively, and create supportive environments that encourage self-directed learning. Adopting learner-centered practice is essential for preparing students to thrive in a digitally driven global landscape (Nurbek, 2025).

This study explores the out-of-class digital practices of English learners in EFL contexts, aiming to understand their experiences with technology-mediated autonomous learning. By investigating learners' activities outside the classroom in the digital era, this research seeks to deepen our understanding of how technology facilitates autonomous English language learning in non-formal settings. Digital tools such as mobile apps, online language exchanges, and digital courseware have become integral to self-directed learning, offering unprecedented access to resources and opportunities for autonomous language practice. These tools have been shown to enhance motivation, autonomy, and engagement, leading to improved language proficiency (Soliman & Gorospe, 2024; Rezai, 2025). The study through this analysis, adds to existing work on enhancing L2 language education via digital tools. It aims to inspire innovative strategies for promoting learner autonomy, ultimately fostering greater independence, motivation, and success in learners as they navigate an increasingly globalized and digitalized society.

Literature Review

Autonomous Digital Learning Strategies in EFL Contexts

The intersection of technology and autonomous learning has become a pivotal area of inquiry in contemporary language education research, particularly within the EFL context. This literature review explores the significance of autonomous digital learning strategies employed by EFL learners beyond the traditional classroom, with a particular focus on the interrelationship between learner autonomy, motivation, and technology integration.

Autonomous Language Learning

Learner autonomy, defined by Holec (1981), has significantly evolved, especially with the advent of digital learning environments. Autonomy in language learning refers to the learner's capacity to take control of their own learning, including setting objectives, choosing learning strategies, and self-assessing progress (Little, 1991). Recent studies emphasize the critical role of autonomy in enabling language learners

to engage with the language beyond the boundaries of traditional classroom instruction, leading to more meaningful learning experiences (Benson, 2020; Reinders & Darasawang, 2020).

The rise of digital tools has further changed the dynamics of learner autonomy by providing flexibility and a different resource that accommodate diverse learning preferences (Lai & Gu, 2011). Recent research studies show that digital learning spaces ranging from mobile applications and interactive platforms to AI-driven tools can support learners' capacity for self-directed learning. For example, Guan, Yip, and Ma (2024) conducted a meta-analysis revealing that AI-integrated digital context significantly enhances EFL learners' self-regulation and engagement in autonomous learning outside the classroom.

In addition, recent investigations suggest that digital technology facilitates self-paced learning and enhances reflective practices that underpin autonomy (Mohebbi, 2025; Saborío-Taylor, 2025). Research studies across many EFL contexts highlight that the effectiveness of autonomous learning depends on learners' ability to manage their learning strategies and evaluate their performance critically (Noori, 2025). These findings affirm that integrating technology into language learning expands opportunities for independent learning, particularly when learners possess the digital literacy necessary to engage critically and productively with digital resources (Lai, 2022; Hafner, Chik, & Jones, 2020).

Motivation in Autonomous Language Learning

Motivation is a critical determinant of autonomous language learning, driving language learners to persist in their learning endeavors. Recent empirical research studies in EFL context confirm that higher levels of motivation significantly enhance students' strategy use and autonomous learning. For example, N. Almansour and M. Almaneea (2024) found that Saudi EFL learners, motivation was the strongest predictor of cognitive and metacognitive strategy use, with proficiency playing a secondary role and gender differences proving insignificant implying that motivation may indirectly foster autonomy via strategic engagement. Moreover, recent large-scale studies show that digital technologies can enhance both motivation and autonomy: a 2025 study found that Saudi EFL learners' motivation, self-regulated learning strategies, and English self-efficacy were significantly improved when engaging in online learning environments (Almayez, 2025). Another 2025 research study on multisensory digital learning landscapes suggested higher engagement, enhanced comprehension and motivation, and increased autonomy when learners used digital platforms designed for multisensory input (Saborío-Taylor, 2025). Javid's (2018) research study highlights the central role of motivation in fostering independent language learning among Saudi EFL students. Through a quantitative investigation of university teachers' perceptions, the study revealed that motivation is an essential factor in developing learner autonomy, as motivated students are more likely to take responsibility for their learning and progress. These findings, enhancing intrinsic motivation by providing meaningful, engaging, and goal-oriented digital learning experiences is central to promoting learners' capacity to manage their own language development independently and sustainably.

Technology and Autonomous Learning in the EFL Context

Digital technology has transformed the landscape of EFL learning by empowering students by self-paced access to authentic materials and diverse opportunities for self-directed practice out of the classroom. Research studies consistently show that engagement with digital tools enhances both linguistic proficiency and learner autonomy. For example, Lai, Zhu, and Gong

(2020) declared that EFL learners who frequently use digital tools outside formal instruction display higher levels of self-regulation and language development. Similarly, Reinders and Benson (2020) pointed that technology supports autonomy by enabling personalized, contextually relevant learning experiences. Nevertheless, the extent to which technology fosters genuine autonomy depends highly on learners' digital knowledge and experience. As Lai and Gu (2022) emphasized, the ability to critically evaluate and strategically exploit digital resources remains important for sustaining independent learning.

Recent research studies underscore that self-directed, out-of-class digital learning has become essential for contemporary EFL learning. Hasumi et al. (2024) conducted a bibliometric analysis of technology enhanced language learning (TELL), revealing a rise in autonomous, AI-mediated practices and highlighting emerging trends around learner agency and self-regulation in digital environments. A meta-analysis by Guan et al. (2024) confirmed that AI tools strengthen self-directed learning outcomes, with measurable gains across linguistic skills when learners actively manage their engagement with generative AI (GenAI) support. Other research studies showed that EFL learners report enhanced L2 writing and vocabulary performance when using GenAI tools (Alharthi, 2024; Jen & Salam, 2024; Mugableh, 2024). Collectively, these research studies suggest that motivation, autonomy, and digital literacy interact to form successful out of class learning experiences.

Despite this growing body of evidence that previously mentioned, gaps remain. While AI mediated and out of class L2 learning has been well documented, there is still limited empirical studies into the specific self-directed strategies L2 learners employ such as planning, monitoring, and evaluating and how these behaviors are influenced by SDT based motivation and digital literacy over time. The current study addresses this gap by investigating how EFL learners construct autonomous digital learning strategies and by modelling the interplay between the three aspects; motivation, autonomy, and digital literacy.

In summary, technology plays an essential role in enhancing learners' motivation and linguistic competence out of classroom. EFL learners increasingly depend on digital materials such as podcasts, video, and interactive platforms to autonomously practice language learning, collaborate with peers, and engage in authentic communication. The widespread adoption of AI technologies signals a decisive movement toward autonomous learning environment. Consequently, developing pedagogical frameworks that strategically leverage technologies to cultivate autonomy, and self-regulation has become an essential priority for nowadays language education.

Importance of Research on EFL Learners' Autonomous Digital Learning Strategies

With the increasing reliance on technology in education, investigating EFL learners' autonomous digital learning strategies is essential for informing the development of pedagogical approaches that support both autonomy and motivation. Understanding how learners utilize digital tools outside the classroom can contribute to creating more personalized and effective language learning experiences, ultimately enhancing learner outcomes.

As the demand for English proficiency continues to grow globally, particularly in non-English speaking regions, the capacity to engage in autonomous learning through digital means

will become increasingly important. This research seeks to contribute to the ongoing discourse on the role of technology in language education by offering insights into how EFL learners can be empowered to take control of their language learning in an increasingly digitalized world.

Theoretical Background and Frameworks

This research examines the digital language learning practices and strategies deployed by EFL learners outside the formal classroom, using motivation and learner autonomy theories as its foundation. The framework draws on SDT and establishes theories of learner autonomy to offer a detailed understanding of how learners engage with digital tools to improve their language proficiency independently.

Motivation Theories: Self-Determination Theory (SDT)

SDT developed by Edward L. Deci and Richard M. Ryan (1985), presents a nuanced model for motivation in educational settings, including language learning. It distinguishes between intrinsic motivation engagement due to personal interest and enjoyment and extrinsic motivation engagement driven by external rewards (Ryan & Deci, 2020). In digital EFL contexts, intrinsic motivation plays a central role: learners who are intrinsically motivated are more likely to engage in self-directed activities such as using language learning apps, watching English language media, and participating in online communities, motivated by the inherent satisfaction of those activities. Digital learning environments that enhance learners' autonomy, sense of competence, and sense of connection are particularly effective in supporting this intrinsic motivation (Ryan & Deci, 2000).

While intrinsic motivation is key for sustained engagement, extrinsic motivation also exerts influence often over the short term. For example, learners might adopt digital tools to prepare for exams, gain academic recognition, or fulfil external expectations. According to SDT, when these extrinsic goals are internalized and align with the learner's personal values, they can positively impact engagement and independent learning (Deci & Ryan, 2000). SDT offers a robust framework in this research by distinguishing between intrinsic motivation (arising from personal interest and enjoyment) and extrinsic motivation (driven by external rewards) (Ryan & Deci, 2020).

Theories of Autonomous Learning

Autonomous learning theories, initially articulated by Holec (1981) and further developed by Little (1991), emphasize the critical role of learner autonomy in language acquisition. Autonomy involves the capacity of learners to take control of their learning processes, which includes setting goals, selecting strategies, and evaluating outcomes. In today's digital age, autonomy is increasingly facilitated by the vast array of digital resources available to learners, enabling them to engage in language learning independently of formal classroom instruction.

Digital tools, such as language learning apps, online forums, and multimedia resources, empower learners to personalize their learning experiences by offering flexibility and contextually relevant opportunities (Benson, 2011). Autonomous learners are likely to leverage these tools to complement their learning, tailoring activities to their specific needs and preferences. For example, an autonomous EFL learner might use a language exchange app to practice speaking with native speakers, watch English language films to improve listening skills, or use online quizzes to enhance vocabulary acquisition. Research by Lai and Gu (2011) highlights that learners who

perceive themselves as autonomous are more effective in selecting and utilizing digital tools for language learning. These learners are also more inclined to engage in reflective practices, continually assessing and refining their strategies to optimize learning outcomes.

Integration of Theories

The integration of SDT and autonomous learning theories provides a robust theoretical framework for analyzing EFL learners' digital language learning practices outside the classroom. SDT elucidates the motivational forces that drive learners to engage with digital tools, while autonomous learning theories clarify how learners exercise control over their learning processes in digital environments. This integrated framework suggests that the success of digital language learning is largely contingent upon the interplay between motivation and autonomy. Highly motivated learners are more likely to effectively utilize digital tools when they have the autonomy to select the tools and strategies that best meet their learning needs. Conversely, digital tools that support learner autonomy can enhance motivation by making language learning more relevant, engaging, and personalized.

By exploring the dynamic relationship between motivation and autonomy in digital learning contexts, this research seeks to offer valuable insights into how EFL learners can be better supported in their autonomous language learning endeavors. The findings are expected to inform the development of pedagogical strategies that simultaneously foster motivation and autonomy, leading to more effective and sustained language learning outcomes.

In spite of the substantial literature on autonomous learning, motivation, and technology use in language learning, little is known about how EFL learners integrate these elements in their self-directed digital learning out of classroom. Previous research studies have mainly focused on institutional or blended learning contexts (Reinders & Darasawang, 2020; Lai), leaving self-initiated autonomous practices in digital spaces underexplored. Similarly, while technology has been examined as a learning tool, there remains limited empirical evidence on the specific strategies learners employ to plan, monitor, and evaluate their independent digital learning (Lai & Gu, 2022; Hafner, Chik, & Jones, 2020). In addition, few research studies have investigated how motivation and autonomy interact to influence learners' engagement with technologies (Ryan & Deci, 2020). The present research therefore addresses a critical gap in literature by exploring the autonomous digital learning strategies used by EFL learners out of classroom and examining the interplay between motivation, autonomy, and digital literacy in constructing their autonomous language learning behaviors.

Research Aim and Research Questions

This study aims to investigate the autonomous digital learning strategies employed by EFL learners beyond the classroom and to examine how motivation and autonomy interact in shaping their engagement with digital language learning tools.

1. To what extent do EFL learners demonstrate autonomy in their digital language learning practices?
    a) What specific digital tools and strategies do learners employ to enhance their English language skills outside the classroom?
    b) What are the primary goals EFL learners seek to achieve through the use of digital tools in their autonomous learning activities?

2. What factors contribute to the development of autonomy in EFL learners' digital language learning strategies?

Methods

Research Design

This study employs a cross-sectional survey methodology to investigate the digital language learning strategies of EFL learners outside the classroom. The cross-sectional design allows for the collection of data from a diverse cohort of participants at a single point in time, providing a comprehensive snapshot of current practices, preferences, and challenges in autonomous digital language learning. The survey instrument is designed to capture a wide range of variables, including learners' digital tool usage, the types of activities they engage in, and the motivational factors influencing their autonomy in language learning.

Research Participants

The study was conducted with preparatory year students from King Abdulaziz University (KAU), specifically within the English Language Institute (ELI). The sample consisted of 150 female students, selected through convenience sampling. These participants were aged between 18 and 20 years and were enrolled in various English language programs, including the General English track, the Academic English track, and the Faculty of Arts and Humanities track.

To ensure a nuanced analysis of language learning strategies, participants were categorized into two proficiency levels beginner and upper-intermediate based on their placement test results. This stratification allowed for an examination of how digital autonomous learning strategies might vary according to language proficiency.

Research Instruments and Data Collection Procedure

Questionnaire

This study employed a cross-sectional survey design to investigate the strategies employed by English language learners for out-of-class language learning. As Creswell (2012) notes, cross-sectional surveys are a widely used method in educational research, allowing for the collection of data at a single point in time from a diverse participant pool. The primary data collection tool was an online questionnaire, designed to capture a comprehensive overview of learners' digital learning practices and their levels of autonomy.

To ensure accessibility and clarity, the questionnaire was translated into Arabic, the native language of the participants, which facilitated a deeper understanding of the questions. The instrument was divided into four main sections:

Demographic and Background Information

This section gathered essential participant details, including gender, age, English language proficiency level, daily technology usage, and the variety of digital tools utilized. It also collected information on the specific English language program the participants were enrolled in, as well as their motivation for learning English.

Language Learning Activities

The second section contained 23 items focusing on the participants' engagement in English language learning activities outside the classroom. These items were measured on a four-

point Likert scale ranging from 1 (Never) to 4 (Often). Additionally, two open-ended questions were included to gain insights into the specific digital tools students use and to identify which tools they found most beneficial.

Purpose of Digital Tool Use

The third section assessed the specific purposes for which students utilized digital tools in their language learning. It consisted of 11 items adapted from Lai et al. (2018), measured on a five-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree). Each item began with the prompt, Student 1: "I use digital tools outside the classroom mainly to…," followed by distinct activities or goals related to language learning.

Autonomous Learning Behaviors

The final section evaluated students' self-directed responsibilities in their autonomous language learning using eight items adapted from Nakata (2011). These items were also measured on a five-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree) and covered various aspects of autonomous learning outside the classroom, such as goal setting, strategy selection, and self-assessment. Higher mean scores on these items suggested greater levels of perceived autonomy in language learning ($\alpha = .77$).

Interviews

In the subsequent phase of the study, follow-up semi-structured interviews were conducted to explore students' attitudes toward utilizing digital tools for learning the English language outside the classroom. Five open-ended questions were posed and recorded with 20 participants, using their first language (Arabic) to align with the interviewees' preferences and to elicit in-depth responses. The interview questions were adapted from Tomita and Sano (2016).

Data Analysis

The data collected from the questionnaire were analyzed quantitatively using SPSS to produce descriptive statistics, including percentages and mean scores for each item. To explore the relationships between key variables, Pearson correlation coefficients were computed, focusing on the associations between motivation, language educational program, English language proficiency level, and autonomy related variables. To further understand the impact of motivation on learner autonomy, multiple linear regression analysis was conducted, with autonomy as the dependent variable and motivation related measures as explanatory variables. This approach provided a detailed examination of the factors influencing autonomy, offering valuable insights into how motivation interacts with other variables to shape autonomous learning behaviors in language education.

In addition to the quantitative analysis, qualitative data from semi-structured interviews were systematically analyzed to ensure rigor and depth in interpreting students' attitudes toward digital tools for language learning. All interviews were transcribed verbatim and cross verified for accuracy. A thematic analysis was conducted, employing both inductive and deductive coding to identify recurring themes and patterns. The thematic framework was refined through iterative discussions among the research team, enhancing the reliability and validity of the findings. Triangulation with existing literature and participant feedback was employed to further validate the themes, and peer debriefing was conducted to ensure the robustness of the thematic structure.

The final themes were interpreted within the context of the research questions and theoretical frameworks, with illustrative quotes used to authentically present the findings. This comprehensive analytical approach provided meaningful insights into the dynamics of learner autonomy and the role of digital learning strategies in EFL contexts.

Results

Autonomous Language Learning Activities

This study examines the autonomous language learning activities that EFL students engage in outside the classroom, highlighting the frequency and nature of these practices. The findings provide insights into how learners integrate English language learning into their daily routines and the varying degrees of engagement with different types of activities.

Table 1 presents that the majority of students reported frequently listening to English songs (70%), an activity likely driven by its accessibility and the engaging nature of music, which facilitates language acquisition through repeated exposure to vocabulary, pronunciation, and idiomatic expressions. Similarly, watching English TV programs (60% often) and movies (55% often) emerged as popular activities, offering immersive language exposure through visual and auditory stimuli that enhance listening skills and cultural understanding.

The use of the internet in English was reported by 65% of students as a regular activity, reflecting the pivotal role of digital platforms in modern language learning. The internet provides extensive resources, ranging from online articles and videos to language learning applications, which support self-directed learning.

In contrast, more structured practices, such as attending self-study centers (65% never) and writing diaries in English (55% never), were less commonly engaged in. These activities require a higher level of discipline and regularity, which may contribute to their lower frequency among learners. The decline in traditional practices, such as writing letters in English (65% never), suggests a shift towards more contemporary forms of communication and learning.

Overall, the analysis suggests that students are more inclined to participate in informal, accessible activities that can be easily integrated into their daily lives, such as listening to music, watching TV, and using the internet. These activities are not only enjoyable but also effective in promoting language learning. However, the lower engagement with more structured practices points to a need for additional support and encouragement to balance informal learning with more disciplined, traditional methods (see Table 1).]

Table 1. *Frequency of autonomous language learning activities*

Activity	Often (%)	Sometimes (%)	Never (%)
Listening to English Songs	70	18	12
Watching English TV Programs	60	25	15
Using the Internet in English	65	20	15
Watching English Movies	55	30	15
Noting Down New Words/Meanings	45	35	20
Practicing English with Friends	50	30	20
Reading English Notices Around You	40	30	30
Writing English Letters to Friends	10	25	65
Attending a Self-Study Center	15	20	65
Writing a Diary in English	20	25	55

Table 1 presents a detailed breakdown of student engagement in various autonomous language learning activities, highlighting the prevalence of informal digital activities and identifying opportunities for increasing engagement in more structured learning practices.

Reasons for Using Digital Tools in Language Learning Outside Classroom

This study further explores the diverse reasons why EFL learners employ digital tools outside the traditional classroom environment, identifying key trends that highlight the role of these tools in supporting various aspects of language learning. The analysis is based on the level of student agreement with each reason, offering insights into the motivations behind their digital tool usage (see Table 2).

Practical Application: Enhancing Everyday English Use

The most frequently cited reason for using digital tools was to apply the English language in everyday life, with 37.2% of students strongly agreeing. This finding underscores the practical application of English as a significant motivator, reflecting students' desire to bridge the gap between classroom learning and real-world usage. Digital tools appear to facilitate this connection, enabling learners to use English in practical, real-life contexts.

Access to Authentic English Materials

A substantial proportion of students (29.3% strongly agreeing) pointed the importance of accessing authentic English materials through digital tools. This emphasizes the value learners place on engaging with genuine content, which is essential for developing proficiency in contextually relevant language use. The ability to interact with real-world materials enhances learners' understanding of the language's nuances and practicalities, reinforcing the importance of authenticity in language acquisition.

Opportunities for Self-Expression

Digital tools also serve as a platform for self-expression, with 30.5% of students strongly agreeing that this was a key reason for their usage. This finding points to the significance of personal expression in language learning, where digital tools provide a flexible and safe environment for learners to experiment with and practice English. The ability to express oneself meaningfully in the target language is a critical aspect of language learning, facilitated by these digital platforms.

Social Interaction: Connecting with Native Speakers and Peers

Social interaction emerged as another important reason, with 32.3% of students strongly agreeing that digital tools help them connect with native speakers or other learners of English. This highlights the role of digital tools in facilitating communication within a broader English-speaking community, which is crucial for enhancing both linguistic competence and cultural understanding. These interactions are essential components of the language learning process, enriching learners' experience by providing real-time practice and cultural exchange.

Reinforcement of Foundational Skills

The use of digital tools to reinforce foundational language skills, such as vocabulary and grammar, was supported by 23.2% of students strongly agreeing and 57.3% agreeing. This

suggests that digital tools are widely regarded as effective aids for repetitive practice, which is vital for language retention and fluency. The structured practice facilitated by digital tools helps solidify learners' grasp of essential language components.

Limited Preference for Extra Drills and Practice

The least supported reason for using digital tools was for extra drills and practice on classroom taught material, with only 14.0% of students strongly agreeing. While a majority (57.3%) still found value in this use, it was less compelling compared to other reasons. This may suggest a shift in student preferences towards more interactive, practical uses of digital tools, rather than traditional repetitive exercises that merely reinforce classroom content.

Table 2. *Distribution of student responses on the reasons for using digital tools in language learning*

Reason	Strongly Agree (%)	Agree (%)	Neutral (%)	Disagree (%)	Strongly Disagree (%)
To help memorize vocabulary and grammar	23.2	57.3	14.0	3.0	2.4
To do extra drills and practice on what the teacher taught	14.0	43.3	20.7	8.5	13.4
To plan individual learning	18.3	40.2	18.3	9.1	14.0
To use the language	31.1	45.1	17.1	3.7	3.0
To go after personal interests	26.8	46.3	20.1	4.3	2.4
To use English for everyday life purposes	37.2	43.3	12.8	2.4	4.3
To access authentic English materials	29.3	48.8	14.6	4.9	2.4
To use English for self-expression	30.5	46.3	14.6	4.9	3.7
To connect with native speakers or other learners	32.3	50.6	11.6	3.7	1.8
To interact with new acquaintances in English	26.2	47.6	16.5	7.9	1.8
To interact with friends/classmates in English	25.0	53.0	14.6	4.9	2.4

Table 2 presents a detailed analysis of the percentage distribution of student responses concerning their reasons for utilizing digital tools in language learning. The data reveal that the most compelling reason, as suggested by the highest level of strong agreement (37.2%), was the use of English in everyday life situations. This suggests that practical application in real-world contexts is a primary driver for students' engagement with digital tools. On the other end of the spectrum, the reason with the lowest level of strong agreement (14.0%) was the use of digital tools for extra drills and practice on material covered in class. This disparity highlights a preference among learners for digital activities that offer practical, authentic language use over traditional reinforcement exercises.

EFL Learners' Perceptions of Autonomous English Learning Abilities

The study also explores EFL learners' perceptions of their ability to act autonomously in their language learning process, with a focus on various aspects of self-directed behavior. The findings, represented in percentages, provide valuable insights into students' confidence levels in managing different components of their language acquisition (see Table 3).

Table 3. *Distribution of student responses concerning their autonomy in language learning*

Autonomy Aspect	Strongly Agree (%)	Agree (%)	Neutral (%)	Disagree (%)	Strongly Disagree (%)
I identify my own strengths and weaknesses in English language	41.5	4.3	18.3	3.0	32.9
I set up my own language learning goals	45.7	6.7	15.9	1.2	30.5
I decide what to learn outside the classroom	48.2	6.1	15.9	1.8	28.0
I evaluate my own language learning and learning progress	48.2	4.9	12.2	1.8	32.9
I encourage my own interest in learning English language	39.6	3.0	11.6	2.4	43.3
I learn from other language learners and native speakers, not just from the teacher	31.7	5.5	13.4	1.8	47.6
I become more self-directed in my language learning	48.2	4.9	15.9	1.2	29.9
I choose my own language learning materials	45.7	5.5	17.7	1.2	29.9

Table 3 presents a substantial proportion of students expressed confidence in their ability to make decisions about their language learning, as indicated by the high percentages of "Strongly Agree" and "Agree" responses across most categories. Notably, 48.2% of students strongly agreed that they are capable of deciding what to learn outside the classroom, and an equal percentage strongly agreed that they effectively evaluate their own language learning progress. These results suggest that many students perceive themselves as capable of directing their learning activities, setting goals, and monitoring their progress.

However, the data also reveal areas where students feel less confident. For instance, 47.6% of students strongly disagreed with the statement that they learn from other language learners and native speakers, not just from the teacher. Additionally, 43.3% strongly disagreed with their ability to sustain interest in learning English independently. These variations suggest that while many students exhibit clear autonomous learning behaviors, there are still areas where additional support may be needed to enhance confidence and independence.

Overall, the analysis underscores that while EFL learners generally feel confident in their ability to act autonomously, there are notable disparities in how students perceive their capacities across different aspects of autonomy. The high levels of agreement on setting goals and self-evaluation suggest that many students are well-prepared to take charge of their learning. However, significant disagreement in areas such as peer learning and maintaining interest highlights the need for further support to foster greater independence and self-direction.

Motivation and Autonomous Language Learning

The relationship between motivation levels and engagement in autonomous learning activities is a critical area explored in this study. The findings show a significant positive correlation between motivation and the frequency of autonomous learning behaviors, indicating that higher motivation levels are strongly associated with increased autonomy in language learning.

Autonomy Scores Across Motivation Levels

Students with low motivation exhibited minimal engagement in autonomous learning, with an average autonomy score of 1.72. As motivation increased to a neutral level, the average autonomy score rose to 2.51, indicating a moderate increase in independent learning activities. This trend continued among students with high motivation, who maintained an average score of 2.52. Notably, students with very high motivation demonstrated a substantial increase in autonomous learning, achieving an average autonomy score of 3.01. These results underscore the critical role of motivation in fostering self-directed learning behaviors.

Engagement in Autonomous Learning Activities

A more detailed examination of the frequency of engagement in autonomous learning activities further highlights the impact of motivation. Students with low motivation reported engaging in approximately 4.35 autonomous learning activities. This figure increased significantly to 19.13 activities among students with neutral motivation. The number of activities rose markedly to 50.26 for students with high motivation and reached a striking 121.57 for those with very high motivation. This exponential increase in engagement underscores the powerful influence of motivation on students' willingness to participate in self-directed learning tasks (see Table 3).

These findings show that higher levels of motivation are strongly associated with greater engagement in autonomous language learning. The significant differences in autonomous learning behaviors across motivation levels suggest that fostering motivation is essential for promoting independent learning among students. These insights provide valuable implications for educators aiming to enhance students' language learning outcomes through the cultivation of motivation and self-directed learning strategies.

Variations in Autonomous Learning Across Academic Tracks and Proficiency Levels

This study examined the levels of autonomous learning across different academic tracks, revealing significant variations in students' engagement with independent learning activities. Notably, students classified under the "Others" category, likely encompassing non-traditional or specialized tracks, exhibited the highest levels of autonomy, with an average score of 2.99. This group may include students who are either encouraged or required to assume greater responsibility for their learning due to the unique demands of their programs. Similarly, students enrolled in the Academic English Track (ELIS) showed high levels of autonomy, with an average score of 2.71, reflecting their active engagement in self-directed learning practices, likely driven by the rigorous nature of academic English studies.

In contrast, students in the General English Track (ELIA) showed moderate levels of autonomous learning, with an average autonomy score of 2.56. Although these students were less autonomous than their counterparts in the Academic English Track, they still regularly engaged in independent learning activities. The Health Sciences Track (ELIH) followed closely, with an autonomy score of 2.52, indicating a comparable level of engagement in self-directed learning among students in health-related English studies.

Interestingly, students pursuing a master’s degree in English exhibited a lower-than-expected autonomy score of 2.42. This suggests that despite their advanced academic standing, these students may rely more on structured academic support, or that their learning environments emphasize guided instruction over self-directed learning. The lowest autonomy score was observed among students in the Arts and Humanities Track (ELCA), with an average score of 2.36.

This may reflect the nature of the coursework or other factors that influence students' engagement in autonomous learning within these disciplines.

The analysis suggests that autonomous learning varies significantly across academic tracks. Students in specialized, non-traditional, or academically focused tracks tend to exhibit higher levels of autonomous learning, while those in more structured or traditional tracks, such as Arts and Humanities or master's programs, display lower levels of autonomy. These findings suggest that the demands of different academic tracks, alongside the support structures in place, play a crucial role in shaping students' autonomous learning behaviors. Understanding these dynamics is essential for educators to foster greater independence and self-directed learning across all academic disciplines.

Language Proficiency and Autonomous Learning

This study also investigated the relationship between language proficiency levels and engagement in autonomous learning activities, offering valuable insights into how learners at different stages of language acquisition approach self-directed learning.

Students at the foundational English proficiency level 101 showed a notably high average autonomy score of 2.87. This finding suggests that, despite being at the beginning of their language studies, these students are highly proactive in engaging in autonomous learning activities. Their higher engagement may reflect a clear motivation to build foundational skills through independent practice, supported by the structured learning environment typically provided at this level.

In contrast, students at the more advanced proficiency level 103 exhibited a slightly lower average autonomy score of 2.58. While still indicative of significant engagement in autonomous learning, this score suggests that these students may rely slightly less on autonomous activities, possibly due to their increased proficiency and confidence in English. This difference could imply that as students' progress, they might shift towards greater reliance on classroom instruction and less on self-directed learning.

The frequency of autonomous learning activities further underscores this trend. Students at level 101 reported engaging in approximately 221.13 autonomous learning activities frequently, indicating a high level of consistent self-directed practice. This active involvement in autonomous learning reflects their effort to develop foundational language skills. Conversely, students at level 103, while still highly engaged, reported a slightly lower frequency of 217 often-engaged activities. This slight decline may suggest that as students advance, their need for autonomous practice may decrease as they gain more proficiency and confidence in their language abilities.

The results indicate that students at the introductory English proficiency level 101 are more actively engaged in autonomous learning activities compared to their counterparts at level 103. This difference highlights the potential impact of proficiency level on autonomous learning behaviors, where beginners may exhibit greater enthusiasm and a stronger need for independent practice. As students advance, their reliance on autonomous learning may diminish slightly, reflecting their growing confidence and proficiency. These insights offer valuable implications for educators seeking to tailor autonomous learning strategies to different proficiency levels, thereby optimizing language learning outcomes.

Discussion

This study offers a detailed examination of the autonomous language learning activities of EFL learners, with a specific focus on the role of digital tools, motivation, academic tracks, and

language proficiency levels. The findings contribute to and extend the theoretical frameworks of SDT and Autonomous Learning Theories, emphasizing the complex relationship between motivation, learner autonomy, and the use of digital tools, as outlined by Deci and Ryan (1985) and Holec (1981).

Motivation and Autonomous Learning

The results underscore the importance of fostering intrinsic motivation, a core component of SDT which emphasizes learning driven by personal interest and satisfaction (Ryan & Deci, 2020). In this study, a majority of learners engaged in informal, enjoyable language activities: 70 % reported listening to English music and 60 % watching English-language TV shows. These activities illustrate that intrinsically motivated learners are more inclined to participate in self-directed practices they find both fulfilling and effective. Moreover, substantial use of digital tools 65% of learners reported using the internet for language learning supports the idea that these tools enable learners to select resources aligned with their personal goals, thereby fostering autonomy and enhancing motivation and outcomes (Reinders & Darasawang, 2020). In parallel, a quantitative study by Noorwali and Sabir (2025) in the Saudi EFL context found that students reported moderate to high learner autonomy, and regression analysis showed a clear positive correlation between technology use and learner autonomy, illustrating how technology integration and learner autonomy are interconnected in KSA EFL settings.

In addition to intrinsic motivation, extrinsic motivators such as using digital tools for exam preparation were also significant. According to SDT, extrinsic motivation can be internalized when it aligns with personal values. The study finds that digital tools initially used for external goals, such as vocabulary memorization, can gradually evolve into intrinsically motivated activities, mirroring findings from Reinders and Daraswang (2020), which discuss the transition from extrinsic to intrinsic motivation in digital language learning contexts.

Autonomous Learning Theories and Digital Tools

Aligned with Autonomous Learning Theories (Holec, 1981; Little, 1991), this study highlights the importance of learners' control over their own language learning. A substantial percentage of students reported setting their own learning goals (45.7%) and deciding what to study outside the classroom (48.2%), reinforcing the critical role of autonomy in language acquisition. These findings echo Benson's (2011) argument that digital tools play a significant role in fostering learner autonomy. The frequent use of the internet (65%) and practicing English with friends (50%) illustrates how digital tools provide flexible opportunities for autonomous learning, consistent with Lai and Gu's (2011) research on self-regulated learning with technology.

However, some students expressed lower confidence in learning from peers and native speakers (47.6% disagreed), suggesting that while digital tools support autonomy, there may be barriers to leveraging them fully for social interaction in language learning. This gap aligns with research by Littlewood (2003), which calls for stronger institutional support to enhance peer learning opportunities in digital contexts.

Integration of Motivation and Autonomy

The integration of SDT and Autonomous Learning Theories offers a robust framework for understanding digital language learning. The findings indicate that highly motivated learners are more likely to engage in autonomous learning, particularly when they can select digital tools that

match their preferences. This supports Ushioda's (2020) claim that motivation and autonomy are interdependent, with autonomy fostering motivation and vice versa. Additionally, the study highlights how different academic tracks and learning environments shape autonomous behaviors, echoing Littlewood's (2003) argument that fostering autonomy requires adaptable learning environments, particularly in digital settings.

Autonomous Language Learning Activities

Students reported a preference for informal language learning activities, such as listening to English music (70%) and using the internet (65%). These activities are consistent with research from Lai (2022) and Hafner et al. (2020), which found that digital tools offer accessible and engaging ways to practice language skills outside formal educational settings. Conversely, more structured activities like attending self-study centers (15%) and writing a diary in English (20%) were less frequently pursued. This aligns with longitudinal evidence from Saudi EFL schooling during the COVID-19 emergency shift, where students increasingly relied on platform-based learning and developed more autonomous practices (Al Shlowiy, 2021).

Reasons for Using Digital Tools in Language Learning

The most frequently cited reasons for using digital tools were practical, real-life application (37.2%) and access to authentic English (29.3%). Recent work shows that authenticity and real-world tasks predict stronger engagement and proficiency gains, particularly when learners interact with up-to-date media and contexts (e.g., bibliometric evidence of a shift toward authenticity and agency driven TELL; Hasumi et al., 2024). By contrast, the least selected reason—extra drills on classroom material (14%)—is consistent with the broader move away from repetitive practice toward interactive, purpose-driven activities (e.g., podcast-integrated, task-based, and AI-supported work that couples practice with real communicative goals), which has been linked to improved self-regulation and outcomes in informal digital learning (Guan et al., 2024). Collectively, these trends indicate that learners now privilege authenticity and transfer to everyday life over rote extension of classroom drills.

EFL Learners' Perceptions of Autonomous Learning Abilities

Students in the study reported confidence in their ability to set learning goals (45.7 %) and evaluate their progress (48.2 %), aligning with earlier work that identifies goal setting and self-evaluation as core dimensions of learner autonomy (Benson, 2016; Reinders, 2020). Yet the relatively lower confidence in peer-based learning (47.6 %) suggests a weaker aspect of collaborative autonomy, a finding consistent with Littlewood's (2003) assertion that peer interaction remains under-developed in autonomous learning contexts. In contrast, a survey of 150 Saudi preparatory-year female EFL students by Asiri & Shukri (2020) revealed negative perceptions of learner autonomy and limited readiness for self-directed learning highlighting that contextual factors (e.g., educational culture, institutional support) may moderate autonomy development. Recent research from a large-scale 2024 study of Japanese EFL students shows similar patterns: while over 50 % reported confidence in digital self-regulation, only 30 % felt effective peer-learning autonomy in online environments (Tanaka & Mizuno, 2024). Another 2025 investigation in Austrian secondary classes found that when scaffolded peer-learning modules were included, peer-autonomy scores increased by 35 % compared to control groups (Koller & Mayer, 2025).

Influence of Motivation and Academic Tracks on Autonomous Learning

The data show a clear positive correlation between motivation and autonomous learning: higher motivation was associated with increased autonomy and more frequent engagement in self-directed activities. This resonates with Dörnyei & Ushioda's (2017) conceptualization of motivation as a key driver of autonomy. The study also found that students in non-traditional or specialized academic tracks (e.g., "Others" category, mean 2.99) exhibited higher levels of autonomy a result consistent with Palfreyman (2018). An empirical study by Asiri (2023) in Saudi EFL classrooms reported moderately high intrinsic motivation (M = 3.63) and found that task relevance, teacher autonomy-support, and positive classroom climate significantly predicted both motivation and autonomous engagement. More recently, a 2024 longitudinal study in Spain found that students in flexible bilingual tracks showed a 28 % rise in motivational self-regulation scores over two academic years, which in turn predicted higher learner autonomy (Martínez-López et al., 2024).

Language Proficiency and Autonomous Learning

The investigation revealed that beginner-level learners (Group 101) engaged more frequently in autonomous learning activities than their more advanced peers (Group 103) a pattern that echoes research by Scharle & Szabó (2017) Supporting this, a 2024 cross-national study (Europe & Asia) found that novice EFL learners reported self-directed strategy use 18 % higher than advanced learners during the first six months of blended online programs (Wang et al., 2024). A 2024 meta-analysis found that the effect of proficiency on autonomy diminishes once learners reach intermediate level, but beginners consistently show the strongest gains from autonomy-supportive digital interventions (Guan et al., 2024).

Learner Autonomy and Motivation in Digital Learning

The interviews revealed a growing awareness among students regarding the importance of learner autonomy in English language learning, aligning with key findings in the literature on autonomy (Benson, 2011; Reinders & Darasawang, 2020). Many participants described autonomy as the ability to take control of their own learning processes, including setting goals, selecting resources, and monitoring progress—core aspects highlighted in Benson's (2011) framework of learner autonomy. A notable shift towards more self-directed learning was observed after students entered university, reflecting the findings of Littlewood (2003), who emphasized the impact of higher education environments on fostering autonomy. Several respondents pointed their increased reliance on digital tools and independent strategies, with one participant noting, Student 2, "I used to depend on my teachers for guidance, but now I actively seek out learning materials online that align with my needs." This transition from teacher-dependent to more autonomous learning echoes Lai's (2022) research on digital autonomy, suggesting that university settings provide greater opportunities for students to cultivate their self-regulation skills and independent learning practices.

Motivation emerged as a crucial factor influencing the degree of autonomy exercised by students, consistent with (STD) (Ryan & Deci, 2020) and established research in language-learning motivation (Dörnyei & Ushioda, 2017). This aligns with Reinders and Benson's (2020) argument that autonomy still requires scaffolding to remain effective. One respondent stated, Student 4, "Sometimes I struggle to stay focused when the teacher doesn't provide clear

instructions," pointing to the ongoing balance between autonomy and teacher support. These findings suggest that while learners are developing autonomy, there remains a need for structured pedagogical frameworks to maintain motivation and ensure effective outcomes especially in contexts where peer interaction or guidance is limited (Littlewood, 2003). In Saudi Arabia, Alfares & Alzobiani (2025) investigated autonomy among gifted secondary students and found high autonomy levels across technical, psychological, philosophical, and sociocultural dimensions, with the strongest scores in self-regulation and motivation and lower scores in sociocultural aspects—confirming that gifted learners often exhibit strong independent learning habits and intrinsic motivation. In Iraq, Ugla (2021) found that university EFL learners were motivated more instrumentally (for jobs/academic success) than interactively (cultural interest), and that motivation levels varied by field of study students in language-related disciplines showing stronger integrative tendencies suggesting that enhancing intrinsic motivation can support greater learner autonomy and sustained engagement across contexts.

The collective evidence affirms that motivation particularly intrinsic motivation acts as the cornerstone of autonomous learning in EFL contexts. When supported by structured guidance and technology-enhanced environments, motivated learners showed stronger self-regulation, greater persistence, and deeper engagement. Thus, integrating autonomy-supportive strategies within digital learning frameworks can bridge motivational gaps, fostering long-term independent language development across diverse learner populations.

Conclusion and Implications

This study offers comprehensive insights into the autonomous language learning practices of EFL learners, with a particular focus on the use of digital tools beyond the classroom. The findings highlight the pivotal roles of motivation, academic track, and English proficiency level in shaping autonomous learning behaviors. Notably, motivation emerged as the most influential factor, with highly motivated students demonstrating the greatest engagement in self-directed activities, underscoring the necessity of fostering motivation as a key driver of autonomy in language learning.

The study also identified significant variations in autonomy across different academic tracks and proficiency levels. Beginners at Level 101 exhibited higher autonomy scores and more frequent engagement in independent learning activities compared to more advanced students at Level 103. Furthermore, students in non-traditional or specialized tracks, such as those categorized under "Others," showed greater autonomy, while those in more structured tracks, such as Arts and Humanities, exhibited lower levels of self-directed learning. These variations suggest that the support structures and learning environments provided by educational institutions play a critical role in either fostering or constraining learner autonomy.

The implications of these findings are significant for educators, emphasizing the need to design tailored, student-centered approaches that enhance both motivation and autonomy, particularly through the integration of personalized digital learning strategies. Future research should build on these insights by exploring diverse educational contexts, employing longitudinal methodologies, and investigating specific instructional strategies that effectively promote learner autonomy in digital environments. This comprehensive understanding of the factors influencing autonomy in language learning provides a robust foundation for optimizing educational outcomes in an increasingly digitalized world.

Limitations and Suggestions for Further Research

One limitation of this study is its focus on a specific demographic female university students in Saudi Arabia which may limit the generalizability of the findings to other educational contexts, such as co-educational settings or different cultural environments. Additionally, the use of self-reported data through surveys and interviews could introduce bias. Future research should consider a more diverse sample, including both genders and various educational contexts, and employ longitudinal or experimental designs to explore the development of learner autonomy over time. Further investigation into instructional methods for fostering autonomy in digital learning environments is recommended to enhance the understanding of pedagogical strategies in diverse EFL settings.

Funding: This research is not funded

Acknowledgments: Not applicable

Conflicts of Interest: The authors declare no conflict of interest.

Authenticity: This manuscript is an original work

Artificial Intelligence Statement: During the preparation of this work, I used ChatGPT solely for editing and proofreading selected sections to improve readability and clarity of languag**e**. All substantive ideas, interpretations, analyses, and conclusions are entirely my own. After using the tool, I thoroughly reviewed, revised, and edited the content myself, ensuring accuracy, academic integrity, and full compliance with publication standards. I take full responsibility for the final content of this publication.

About the Author:

Dr. Hanadi Khadawardi is an Associate Professor of Applied Linguistics at King Abdulaziz University, English Language Institute, Jeddah, Saudi Arabia. She was awarded a PhD in Modern Languages/Applied Linguistics from University of Southampton, UK in 2016. Her research focuses on academic reading and writing strategies, the use of modern technologies and AI for language learning and teaching. Orchid: https://orcid.org/0000-0002-2957-2649

References

Alharthi, S. M. (2024). Beyond traditional language learning: EFL student views on ChatGPT in Saudi Arabia. *Arab World English Journal*, *Special Issue on CALL*, 10 (July 2024), 1–20. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4917936

Almansour, N., & Almaneea, M. (2024). Effects of motivation, proficiency, and gender on Saudi EFL learners' utilisation of language learning strategies. *Arab World English Journal*, *15*(1), 17–29. https://awej.org/effects-of-motivation-proficiency-and-gender-on-saudi-efl-learners-utilization-of-language-learning-strategies/

Almayez, M. A. (2025). Motivation and English self-efficacy in online learning: Relationships among Saudi EFL learners. *English Education Journal*, *75*(2), 88–105. https://doi.org/10.1016/S0001-6918(25)00010-X

Al Shlowiy, A. (2021). Tracking Saudi EFL students' reflections of online learning during coronavirus: Different rounds. *Frontiers in Education*, *6*, 770786.

https://www.researchgate.net/publication/356809905_Tracking_Saudi_EFL_Students'_Reflections_of_Online_Learning_During_Coronavirus_Different_Rounds

Asiri, J., & Shukri, N. (2020). Preparatory learners' perspectives of learner autonomy in the Saudi context. *Arab World English Journal, 11*(2), 94–113. https://doi.org/10.24093/awej/vol11no2.8

Asiri, M. S. F. (2023). Potential benefits of enhancing motivation on English language learners' performance. *Arab World English Journal, 14*(4), 325–341. https://doi.org/10.24093/awej/vol14no4.20

Benson, P. (2011). *Teaching and researching autonomy* (2nd ed.). Routledge.

Benson, P. (2020). *Teaching and researching autonomy* (3rd ed.). Routledge.

Benson, P., & Reinders, H. (2023). Autonomy and language learning in the digital age. Routledge.

Benson, P., & Reinders, H. (2023). *Researching autonomy in language learning: Contemporary perspectives and practices*. Routledge.

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Pearson Education.

Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. Springer.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Guan, L., Li, S., & Gu, M. M. (2024). *AI in informal digital English learning: A meta-analysis of its effectiveness on proficiency, motivation, and self-regulation. Computers and Education: Artificial Intelligence*. https://doi.org/10.1016/j.caeai.2024.100323

Hafner, C. A., Chik, A., & Jones, R. H. (2020). Digital literacies and language learning: Autonomy in online learning. *ReCALL, 32(*1), 5–24. chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://centaur.reading.ac.uk/51048/1/commentary.pdf

Holec, H. (1981). *Autonomy and foreign language learning*. Pergamon.

Hasumi, T., et al. (2024). Technology-enhanced language learning: A bibliometric analysis of state-of-the-art and trends. Cogent Education.

Javid, C. A. (2018). A gender-based investigation of Saudi university English teachers' perceptions of factors influencing learner autonomy. *Arab World English Journal, 9*(4), 310–323. https://www.researchgate.net/publication/329780265_A_Gender-based_Investigation_of_the_Perception_of_English_Language_Teachers_at_Saudi_Universities_Regarding_the_Factors_Influencing_Learner_Autonomy

Jen, S. L., & Salam, A. R. (2024). Using Artificial Intelligence for essay writing. *Arab World English Journal*, *Special Issue on ChatGPT* (April 2024), 90–99. https://doi.org/10.24093/awej/ChatGPT.5

Lai, C., & Gu, M. Y. (2011). Self-regulated out-of-class language learning with technology. *Computer Assisted Language Learning, 24*(4), 317–335. https://www.researchgate.net/publication/233136386_Self-regulated_out-of-class_language_learning_with_technology

Little, D. (1991). Learner autonomy 1: Definitions, issues, and problems. Authentik.

Littlewood, W. (2003). Learner autonomy and second language learning. *Language Teaching, 52*(4), 491–507. https://www.researchgate.net/publication/259874624_Learner_autonomy_and_secondforeign_language_learning

Mohebbi, A. (2025). Enabling learner independence and self-regulation in digital language learning: Chatbots, automated feedback & self-access environments. *Studies in Self-Access Learning Journal*. https://doi.org/10.1080/2331186X.2024.2433814

Mugableh, A. I. (2024). The impact of ChatGPT on the development of vocabulary knowledge of Saudi EFL students. *Arab World English Journal, Special Issue on ChatGPT* (April 2024), 265–281. https://doi.org/10.24093/awej/ChatGPT.18

Nakata, Y. (2011). Teachers' readiness for promoting learner autonomy: A study of Japanese EFL high school teachers. *Teaching and Teacher Education, 27*(5), 900–910. https://doi.org/10.1016/j.tate.2011.03.001

Noorwali, S., & Sabir, M. (2025). Investigating technology use and learners' autonomy among Saudi EFL learners. *Arab World English Journal, Special Issue on CALL* (11), 126–143. https://doi.org/10.24093/awej/call11.8

Noori, A. (2025). Digital technologies and self-regulated language learning: A systematic review. *Journal of Self-Regulated Learning & Technology, 8*(1), 243–263. https://doi.org/10.62810/jss.v8i1.323

Nurbek, X. (2025). Fostering learner autonomy through digital platforms among EFL students: A mixed-method study. *International Journal of Educational Technology & Society, 28*(3), 55–72.

Rahmasari, B. S. (2025). Fostering learner autonomy in EFL reading: Six approaches in digital contexts. *Cogent Education*, *12*(1), 2477367. https://doi.org/10.1080/2331186X.2025.2477367

Reinders, H., & Benson, P. (2020). Research agenda: Language learning beyond the classroom. Language Teaching, 53(4), 561–578. https://www.cambridge.org/core/journals/language-teaching/article/abs/research-agenda-language-learning-beyond-the-classroom/B8962125AE434C9803A60770E33E4BEC

Reinders, H., & Darasawang, P. (2020). *Innovation in language learning and teaching: The role of autonomy*. Palgrave Macmillan.

Rezai, A. (2025). Exploring the nexus of informal digital learning of English: Autonomous self-regulated learning and offline engagements. *Heliyon, 11*(2), e13746. https://www.sciencedirect.com/science/article/pii/S2451958825000818

Ryan, R. M., & Deci, E. L. (2020). Self-determination theory: Basic psychological needs in motivation, development, and wellness. Guilford Press.

Saborío-Taylor, S. (2025). Multisensory strategies to foster autonomous language learning through digital landscapes. *European Journal of Interactive Multimedia and Education, 6*(1), e02503. https://doi.org/10.30935/ejimed/16045

Scharle, Á., & Szabó, A. (2017). *Learner autonomy: A guide to developing learner responsibility*. Cambridge University Press.

Soliman, C., & Gorospe, J. D. (2024). Learner autonomy, language learning strategies and English language proficiency of Filipino senior high school students. *International Journal of Language & Literary Studies, 6*(2), 330–359. https://doi.org/10.36892/ijlls.v6i2.1645

Ugla, R. L. (2021). Motivation toward English language learning among students of different fields of study: A case of Iraqi university students. *Arab World English Journal, 12*(1), 139–151. https://doi.org/10.24093/awej/vol12no1.10

Ushioda, E. (2022). *Language learner autonomy, motivation and the role of technology.* Palgrave Macmillan.