Arab World English Journal

INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

AWEJ Volume.16 Number.16 September 2025

www.awej.org

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025, Pp.-1-2

Contents

Article Titles & Authors	Pages
Contents	1-2
Digital Tools and Writing Education: A Thematic Analysis of Technology's Role in Writing Skills Development Nagamurali Eragamreddy & Rachel Joseph	3-25
Self-Regulated Learning Strategies as Predictors of PERMA Well-being in MOOC-Based English Language Learning Danyang Yue & Mohd Rashid Bin Mohd Saad	26-44
Students and Instructors' Views on the Use of Portfolios in Teaching Writing at the Tertiary Level in Saudi Arabia Hussein Taha Assaggaf	45-59
Research Trends of Digital Game-based Learning for Technical Vocabulary Acquisition: A Systematic Decennial Literature Review (2014-2024) Feng Tian & Irfan Naufal Umar	60-88
Innovative Linguistic Strategies of EFL University Students' Performance in Grammar Teaching Raghad Khalaf Ulaywi Sahaw AlDulaimi	89-112
First Language Transfer Among Chinese High School Students in the English Writing Process YANQIN MA & MAHANI STAPA	113-128
Effects of English Songs on EFL Learners' Speaking Performance: Lexical Complexity, Accuracy, and Fluency Tantip Kitjaroonchai & Krittat Sukman	129-143
Harnessing AI in CALL: Theories, Pedagogies, and Applications for Adult Language Learning Lakshmi Kakumanu	144-158
Which Engages Students More in English Writing: Teacher+Peer Feedback or Teacher-Only Feedback? -A Study in Chinese Tertiary Education Min Deng, Nur Ehsan Mohd Said, Nur Ainil Sulaiman & Yunyun Zhao	159-174
A Comparative Study of Human and AI-Assisted Assessment using ChatGPT: The Case of Moroccan EFL Learners Hajar Ennadir & Salma Habiballah	175- 187
Global Trends in Language for Specific Purposes Curriculum Development: A Systematic Review Mior Syazril Mohamed Sapawi & Nik Mohd Rahimi Nik Yusoff	188-210
From Fragmentation to Coherence: A Knowledge Graph-Driven Integration Framework for Blended Learning Lingling Lou	211-226
Exploring Students' Perception Towards the Use of Grammarly to Improve Grammatical Accuracy in Essay Writing Nur Hallyda Mohd Sham, Nooreen Noordin & Norhakimah Khaiessa Ahmad	227-243
Preserving Authorial Voice in Academic Texts in the Age of Generative AI: A Thematic Literature ReviewJokha Al Hosni	244-258
The Trends in the Integration of AI Technologies in English Language Teaching: Systematic Literature Review Wani Nurfahani Mohd Sapuan, Nur Ainil Sulaiman & Maslawati Mohamad	259-269

Contents

Realizing ESL-oriented Education in the Singaporean Setting: A Scoping Review Phan Ngoc Tuong Vy, Trinh Quoc Lap, Le Cong Tuan, Dang Vu Hoai Nhan & Le Xuan Mai	270-284
Preference Structure in Refusals by Chinese Female International Students He Zhai, Ridwan Wahid & Woo Wai Sheng	285-300
Unpacking Learner Autonomy: Instructor Insights, Sustainable Benefits, and Challenges in Tertiary EFL Context Lara Arafat & Zeinab Al Khshali	301-314
Needs Analysis of English-Language Use of Tourism Personnel in Xi'an, China Boyu Pang, Faizah Binti Abd Majid & Geethanjali A/P Narayanan	315-328
Investigating the Difficulties Encountered by Saudi English Majors in Transcribing Schwa Sound \ə\ Ghaida Khalid Albahlal & Ansa Hameed	329-344
Investigating the Relationship Between Self-Regulated Writing Strategies and Academic Writing Performance Xiaolan Wang, Supyan Hussin & Azlina Abdul Aziz	345-358
Enhancing L2 Writing Pedagogy Through AntWordProfiler and Flesch Reading Score Analyses of Vocabulary and Readability Princess Gerbie C. Durante	359 -372
'We Divided the Questions Evenly among Ourselves to Answer them Properly': Collaborative Learning in EFL Classroom Status Quo Nada Bin Dahmash	373-387
Past Tense Forms in CEFR-Aligned English Textbooks Used at Malaysian Primary Schools and Their Directive Functions Patrick Duffy Bayuong, Harwati Hashim & Melor Md. Yunus	388-406
The Impact of Watching English Shows on Enhancing English Language Acquisition Among Undergraduate Students Daianne S. Gloria & Elenita P. Galvez	407-425
Investigating the Role of AI Tool Interactions in Enhancing English Language Acquisition Among Saudi College Students Amnah Abdulrahman Alhamam	426-440
Literary Vocabulary of Charlotte Bingham's Novels: Stylistic and Functional Aspects Yuliia Haidenko, Oksana Serheieva , Olesia Sribna & Olha Hrabar	441- 460
Learning Self-Efficacy, Regulation, and Resilience among English-Majored Students in the Vietnamese Context Do Minh Hung	461-471
Mentoring Teachers: Supporting Learning, Wellbeing and Retention Angi Malderez & Reviewer: Hamdan F. Alzahrani	472-475

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp. 3-25
DOI: https://dx.doi.org/10.24093/awej/vol16no3.1

Digital Tools and Writing Education: A Thematic Analysis of Technology's Role in Writing Skills Development

Nagamurali Eragamreddy
General Requirements Unit, Preparatory Studies Center, University of Technology and Applied Sciences, Salalah, Sultanate of Oman
Corresponding Author: muralieragamreddi@gmail.com

Rachel Joseph
General Requirements Unit, Preparatory Studies Center, University of Technology and Applied Sciences, Salalah, Sultanate of Oman
Email: Rachel.Joseph@utas.edu.om

Received: 06/25/2025 Accepted:08/26/2025 Published:09/20/2025

Abstract
This study investigates how technology facilitates the development of higher-order writing capacities, specifically critical thinking and creativity, to benefit diverse groups of learners, including English as a Foreign Language (EFL) students and those with dyslexia. Limited research on the effect of technology on writing skills, creativity, and critical thinking, which fails to focus on the different needs of learners and fails to acknowledge the individual needs of different learner groups. This study, which identifies a shortage of literature that predominantly addresses lower-level writing mechanics, poses the primary research question: to what extent do digital tools support higher-order writing skills? By utilizing a systematic review study design and thematic analysis approach, the study selected and analyzed 40 peer-reviewed articles. The researchers gathered the data by systematically reviewing various academic databases. This helped the authors identify the latest developments in digital writing tools, collaborative writing environments, and AI-based feedback systems. The five grand themes surfaced after an iterative coding scheme: computer-based writing tools, interactive writing environments, collaborative writing environments, automatic feedback systems, and virtual writing communities. The study's results show that computational tools not only significantly augment elementary mechanics but also lead to a substantial rise in creativity and critical thinking when applied in interactive and collaborative environments. However, the effectiveness of these tools depends on digital literacy and ethical considerations related to AI applications. This study finds that technology has immense possibilities for re-inventing writing pedagogy if combined with conventional pedagogical approaches to improve fundamental and advanced writing skills.

Keywords: AI writing tools, critical thinking, learner diversity, thematic analysis, writing education

How to cite: Eragamreddy, N., & Joseph, R. (2025). Digital Tools and Writing Education: A Thematic Analysis of Technology's Role in Writing Skills Development. *Arab World English Journal, 16*(3):3-25. https://dx.doi.org/10.24093/awej/vol16no3.1

Introduction

Across various educational settings, the growing popularity of modern technology in language instruction has become the norm rather than the exception (Aljuaid, 2024). Examples include multimedia learning, online learning, web-based learning, ICT, computer-mediated communication (CMC), and, more recently, technology-enhanced language learning (TELL). Instructions from governments worldwide have prompted tertiary institutions, in particular, to integrate information technology (IT) into language instruction. It is now more critical than ever for instructors to assess the learning opportunities provided by technology, according to applied linguists such as Chapelle (2003), who believe that advances in information and communication technologies have substantially enlarged the alternatives for English language acquisition and instruction. As Tran and Nguyen (2021) highlighted, technology developments have significantly impacted written communication. Scholars have observed that technology may serve as a valuable writing aid, given the transition from paper and pen to an LCD screen. According to Pennington (1996), computer technology serves as both a collaborator in the development and retrieval of information and a tool. Therefore, there is an expanding demand to investigate the various elements of technology adoption due to the ever-growing usage of technology in language instruction settings. Of these, the influence of technology on traditional instructional methods in language education, particularly in writing, appears to be significant.

Despite the wealth of technological writing tools, their subtle effects on the complete range of writing competencies—particularly higher-order competencies such as creativity and critical thinking—have been understudied. Most research focuses on advancements in mechanics and grammar, without considering how technology affects the cognitive and creative aspects of writing. Additionally, most studies generalize writing development and do not consider the varied needs of students across settings and populations. This need is even more urgent in the face of swift technological evolution, including the emergence of AI-driven writing support tools, whose instructional efficacy has not yet been extensively investigated. Filling these gaps is crucial for the development of comprehensive and equitable writing instruction. By examining how emerging technologies impact higher-level writing ability and serve diverse learner profiles, researchers can supply practical information for instructors and developers. This knowledge is fundamental for the development of focused, evidence-based intervention that maximizes writing growth in varied learning settings (Bhusal & Shrestha, 2024; Nazari et al., 2021; Al Balushi, 2024). Therefore, the following research objectives and questions are formulated to answer the research gaps in the current study.

1. To explore how technology affects advanced writing abilities like creativity and critical thinking.
2. To investigate technology-based writing training and the unique requirements of various learner populations, such as EFL students and dyslexic writers.
3. To assess how well developing writing abilities can be aided by growing technology, such as AI-powered writing tools.
4. To inform the creation of focused, empirically supported interventions by synthesizing evidence on how technology influences writing skill development across diverse learner demographics and educational contexts.

1. What effect does technology have on advanced writing skills like creativity and critical thinking?

2. How can technology be used to support different learners, such as dyslexic writers and EFL students, in writing instruction?
3. How effective are new technologies, like AI writing tools, in helping learners improve their writing?
4. How can current evidence on the impact of technology on writing skills inform the design of effective interventions for diverse learners?

Literature Review

Technology Tools and Writing Mechanics

The number of research articles on the use of technology to enhance English language proficiency, particularly writing performance, has increased over the past few years (Benlaghrissi & Ouahidi, 2024). Because writing requires such complexity, it is the most complex skill to master. This is because writing abilities necessitate the ability to generate and arrange ideas, as well as the mastery of numerous writing-related skills, including syntax, vocabulary, word selection, and punctuation (Dendup & Onthanee, 2020; Valizadeh, 2022). A considerable body of research recognizes the value of digital technologies in enhancing lower-order writing capacities, including grammar, spelling, and organization (Navarette & Elan, 2025; Tran, 2025). Automated Writing Evaluation Systems (AWESs), including PEG Writing, IntelliMetric, and e-Rater, are commonly incorporated into classrooms to enable scalable, immediate feedback (Miranty et al., 2025; Huawei & Aryadoust, 2023; Koltovskaia, 2023; Geçkin et al., 2023; Sari & Han, 2022; Link et al., 2022; Hazelton et al., 2021). These programs improve spelling, punctuation, and syntax correctness, providing iterative writing experiences. However, while helpful for mechanics, AWES's role in supporting creativity or critical thinking is less researched. Cloud-based systems and Online Writing Checkers (OWCs), such as AccurIT, also demonstrate potential for assisting in technical writing dimensions. Grami (2020) found that OWCs helped ESL students improve collocation and mechanics, although student attitudes varied according to their prior experience with computer-based tools. However, the emphasis on surface error correction highlights a drawback: the lack of attention to metacognitive or ideational writing processes. These tools lack the sophistication to inform higher-order thinking honestly. This deficiency is also recognized by McCarthy et al. (2022), who advocate for a “strategy-based feedback that transcends grammar correction.”

Recently, attention has shifted to investigating the advantages of combining different technology tools to improve students' writing abilities (Li, 2023). The significance of e-learning in fostering academic writing skills is emphasized by Hoang and Hoang (2024). According to Ellison and Drew (2020), computer games can help authors develop their critical thinking and organizing abilities. Numerous writers emphasize the importance of technology in establishing the best possible learning environment, encouraging students to write, and enhancing the caliber and volume of student outputs. Furthermore, it is recognized that technology may reduce errors through enhanced editing capabilities and store time for both teachers and students (Goldman et al., 2024; Marzuki et al., 2023). This growing corpus of research highlights the numerous benefits of various technological instruments in forming a more successful and efficient approach to writing instruction.

Web-Based and Collaborative Writing Environments

Computer-mediated collaborative writing spaces are highly commended for facilitating substantive language use, peer interaction, and negotiation of meaning. Chukwuere (2024) notes

that there is increasing interest in ChatGPT and other large language models (LLMs) and how they can affect academic writing. Due to the novelty of ChatGPT, the study likely doesn't examine a large corpus of earlier research; however, it does contribute to the ongoing discussion about how technology can enhance writing. The advantages of ChatGPT, including its ability to conduct literature reviews, facilitate cooperation, and support data analysis, are highlighted. The report, however, acknowledges several limitations, including concerns about plagiarism. This emphasizes the need for further study on the moral and practical integration of LLMs into academic writing processes, aiming to supplement human judgment and critical thinking abilities rather than replace them. Studies indicate that tools like Google Docs, forums, and blogs promote both peer feedback and synchronous co-construction of texts and therefore increase students' motivation and language awareness (Kasih et al.,2021; Warschauer, 2020; Purwanto et al., 2023; Al-Shaboul et al., 2023; Zhang & Zou, 2022; Zioga & Bikos, 2020; Li, 2023). The benefits of a web-based writing environment, as enumerated by Dousti and Amirian (2023), include: allowing students to examine and learn from one another, providing an opportunity for students to give and receive feedback, enabling pupils to publish their work, offering an effective editing environment for pupils, and providing an educational setting.

The virtual forum served as a venue for students to communicate, reflect on, and debate pedagogical concerns raised by their reading and use of various web-based resources in the classroom. According to Hallam (2022), the primary purpose of a virtual forum is to facilitate genuine conversation and encourage participation. In the meantime, reciprocal engagement among pupils in communal pursuits fosters peer review and real-world learning environments (Zewail-Foote & Gonzalez, 2023). Computer-mediated collaborative writing (CMCW) has been particularly effective in fostering interaction and learner autonomy. As Kim and Kim (2020), Liu et al. (2022), and Han (2023) argue, CMCW environments enable learners to cultivate audience awareness, coherence, and argumentation competencies—ingredients of higher-order thinking. Zhai (2021), Zou et al. (2023), and Zhang (2023) established that students' affective engagement in collaboration dramatically enhanced their motivation and reflective awareness, which underpin the development of creativity. However, the effectiveness of collaborative tools is usually predicated on learners' ability to plan and their digital literacy. Zhang and Zou (2022) warn that "in the absence of adequate goal-setting or collaboration training, students are unlikely to reap the tools' full benefits." This implies the requirement for pedagogical support in the introduction of collaborative writing technologies.

Feedback Systems and Writing Development

Feedback—especially formative and process-based—plays a central role in writing instruction. While teachers remain primary providers of feedback, AWEs are increasingly seen as complementary tools. Woodworth and Barkaoui (2020) and McCarthy et al. (2022) note that AWEs improve autonomy, increase writing practice opportunities, and support strategic revision. However, concerns persist about over-reliance, lack of customization, and potential erosion of teacher responsibility. Online tools, such as Padlet, have also been adopted for writing testing, with the potential to deliver instant and accessible feedback (Jong & Tan, 2021). However, issues of internet connectivity and students' limited ICT competency hinder their widespread usage, especially in under-resourced settings.

Artificial Intelligence (AI) and Higher-Order Writing Skills

According to Prayudi et al. (2021), a crucial element in accelerating language acquisition is the method teachers employ in the classroom. Current research highlights the emergence of AI writing tools—such as Sudowrite and Wordtune—as new aids for writing development. These AI writing tools provide sentence rephrasing, style modification, and content creation, raising questions about their cognitive effect. Fang et al. (2024) and Zhao et al. (2024) argue that AI tools have the potential to foster critical thinking if students are taught to evaluate and select AI suggestions with discernment. This positions technology less as a correcting aid and more as a scaffold for greater metacognitive involvement. However, these advantages are not inherent. Zhao et al. (2024) emphasize that students need to develop evaluative and interpretive approaches to effectively utilize AI, underscoring the necessity for incorporating digital literacy instruction into writing courses. Chairunisa et al. (2024) also draw attention to the relatively unexplored nexus between digital literacy and writing ability, identifying a research gap regarding the mediating role of digital competence on writing performance. Despite these promising findings, research on AI tools tends to be exploratory or small-scale, with limited sample sizes and/or context-specific concentration (e.g., Chinese or Indonesian students), which constrains generalizability. Large-scale, longitudinal research is necessary to confirm these findings in more extensive contexts.

Blended and Genre-Based Approaches

Some researchers advocate for blended learning models that combine traditional teaching with technological support. Mahmood et al. (2024) found that ESL learners in a blended setting demonstrated improved paragraph writing, vocabulary, and reflective thinking. Likewise, Wicaksono et al. (2022) state that combining genre-based teaching with online forums enhanced learners' awareness of the audience, creativity, and involvement in genuine discourse. These investigations demonstrate that technology works best when it augments, rather than substitutes for, good pedagogical practice. Embedded within structured frameworks such as genre-based instruction or scaffolded peer review, technology can facilitate the development of higher-order capabilities in powerful ways.

Student-Centered Engagement and Emotional Factors

The effects of technology on affective engagement and learner attitudes constitute another emerging research strand. Nazari et al. (2021) investigate the effects of AI tools on the writing behavior and motivation of postgraduate students, highlighting the scarcity of research linking affective aspects to technology use in writing. Rahimi and Fathi (2021) and Li and Zhang (2023) illustrate how collaborative digital writing promotes learner agency and emotional investment. In addition, blogs and forums encourage introspection, self-expression, and a sense of community—all of which are essential for long-term creative engagement in writing (Mardiana et al., 2022; Caprara and Caprara, 2022). More empirical research is necessary, however, to establish how these affective and social advantages map onto quantifiable improvements in creativity or critical analysis.

Although the literature supports technology's facilitative role in enhancing the mechanical aspects of writing, evidence for its impact on higher-order skills remains fragmented and primarily indirect. Numerous tools are praised for enhancing engagement or efficiency, yet few studies directly examine their impact on creativity, idea generation, or evaluative reasoning. In addition, the field is without critical comparisons among various technologies or control-group designs that

would establish the relative effectiveness of such tools. Moreover, a significant gap exists in research that covers various learner needs—most studies focus on tertiary-level EFL learners from Asia, and cultural and contextual generalizability is a concern. Lastly, although several studies indicate the potential of AI tools and collaborative software, few provide strong longitudinal evidence or theoretical underpinning in writing pedagogy.

Methodology

Procedure

This review employed a systematic approach, synthesizing findings through thematic analysis of recent peer-reviewed articles on the application of digital technologies in writing teaching. An extensive and comprehensive literature search was conducted on various academic databases, including ERIC, Scopus, Web of Science, and PsycINFO. The literature search was restricted to articles published between 2020 and 2024 to ensure coverage of the latest advancements in writing pedagogy and educational technology. The search strategy employed Boolean operators and various synonymous terms to enhance the inclusiveness of the results. Keywords used were combinations of: "technology" AND "writing education", "AI" OR "artificial intelligence" OR "large language models" AND "writing skills", "digital tools" AND "composition instruction", "EFL" OR "ELL" AND "writing development", "assistive technology" AND "dysgraphia"

Inclusion criteria were specified to ensure the relevance and quality of sources: Peer-reviewed empirical studies (qualitative, quantitative, or mixed methods), published from 2020 to 2024, focused on writing skill development involving technology, including human learners at any level of education, studies were excluded if they were not peer-reviewed (e.g., editorials, opinion pieces, book chapters), did not report empirical results (i.e., purely conceptual or theoretical articles), focused solely on reading, speaking, or general language skills without a substantial focus on writing and lacked methodological transparency (e.g., absent information about participants, data collection, or analysis) Thematic analysis was conducted according to Braun and Clarke's (2006) guidelines, including an iterative coding, categorizing, and theme-refining process. In this way, the research was able to uncover patterns and trends in how different types of technology—and especially AI-supported tools and learning platforms—impact writing teaching, including grammar, structure, creativity, and critical thinking. Thematic analysis's transparency and flexibility allowed for several theoretical approaches, providing reliability through intercoder checks and consistency reviews (Morgan et al., 2023; Ntsobi et al., 2024). This methodology provided a rich, contextualized synthesis of contemporary research with the potential to inform both pedagogical practice and technological design in writing education.

Data Collection

During the first stage of data gathering, a systematic literature search was conducted in academic databases, including ERIC, Scopus, Web of Science, and PsycINFO, as well as authoritative online libraries and educational websites. The search was conducted using Boolean operators and synonymous keywords to retrieve a comprehensive body of related literature. Some of the keywords used in various combinations were: "technology in writing education," "digital tools for writing skills," "technology-enhanced writing," "AI-assisted writing," "writing development technology," "EFL/ELL writing," and "assistive writing technologies." This resulted in an initial corpus of 372 publications. There was an initial screening stage based on titles and abstracts, during which 92 studies were removed due to a lack of relevance to writing-specific outcomes, and 15 articles were excluded because of inaccessibility or paywall limitations, resulting in 265 full-text articles for closer analysis.

The researchers then imposed explicit inclusion and exclusion criteria on this narrowed dataset. The researchers only included peer-reviewed empirical research—qualitative, quantitative, or mixed methods—published between 2020 and 2024 that explicitly targeted writing skill development using technology. The study was excluded if it did not involve empirical data (e.g., conceptual models, theoretical

essays), did not transparently report methodological procedures or results, or targeted general language skills without a significant writing element. This secondary screening resulted in the elimination of 190 studies that lacked empirical data on writing outcomes, 14 for methodological transparency (e.g., lack of participant data, indeterminate analytic procedures), and 21 for marginal relevance to the research emphasis, leaving a final sample of 40 studies for detailed analysis.

Coding and Thematic Development

Thematic analysis adhered to Braun and Clarke's (2006) six-phase guideline, pursuing a hybrid inductive–deductive coding strategy. Two researchers independently performed open coding in the first round, developing preliminary codes line-by-line from the findings and discussion sections of the 40 chosen articles. This entailed identifying and naming significant concepts, including instructional strategies, learner outcomes, technologies employed, and the integration of pedagogy. In phase two, a constant comparative approach was employed to group comparable codes and merge redundant labels. Via axial coding, these provisional codes were sorted into upper-level categories according to conceptual connections (i.e., codes concerning peer-feedback software and collaborative revision were subsumed under a general category of computer-based collaboration). Discrepancies in coding were addressed via discussion and inter-coder reliability checks. The move from categories to higher-order themes entailed iterative refinement through team meetings. Researchers explored the co-occurrence of codes in several studies, re-read the original papers to confirm emerging trends, and employed theoretical triangulation by contrasting the emergent themes with existing frameworks in writing pedagogy and educational technology. Preliminary themes were repeatedly checked against the raw data to ensure coherence, internal homogeneity, and external heterogeneity, ensuring that each theme represented a distinct conceptual pattern. This multi-layered, systematic process permitted the creation of clearly defined and analytically strong themes without theoretically constraining them in advance, thus allowing for a grounded synthesis of the ways digital technologies impact writing development.

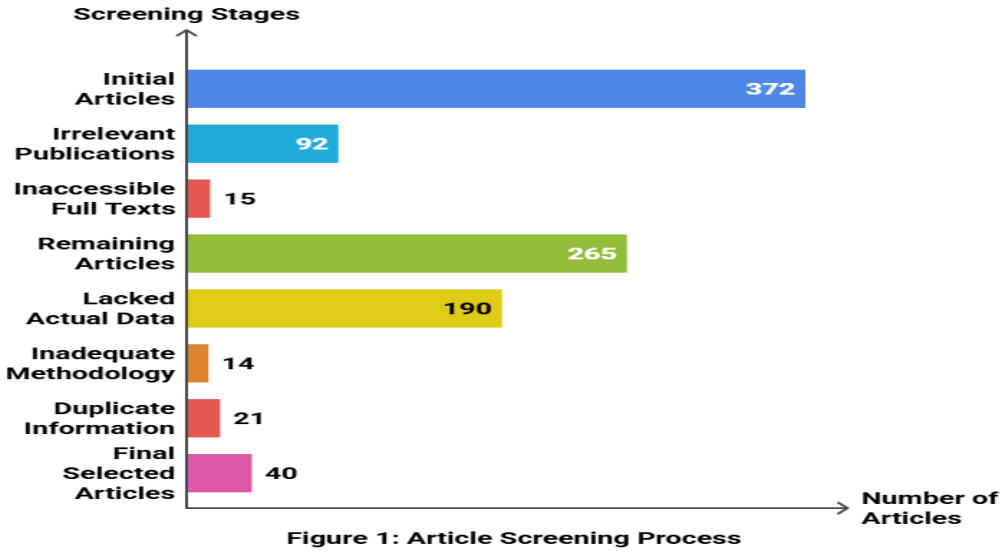


Figure 1. Article screening process

Analytical Framework

The researchers reviewed the materials, focusing on the sections relevant to the study's topic, to create the initial codes (Schmieder & Rogers, 2024). In analyzing the data, the researchers employed a thematic analysis methodology, drawing on the principles outlined by Braun and

Clarke (2006). The dataset comprised peer-reviewed journal articles, along with purposively selected grey literature—namely, institutional reports, conference proceedings, and practitioner-based publications—to enable triangulation of insights. Although the initial sweep of the search primarily targeted peer-reviewed sources, grey literature was added in a follow-up sweep to obtain innovative practices and practical interventions that may not have been captured in academic journals. These additional sources were identified based on relevance, recency, and credibility. Coding commenced with open coding of a pilot sample of 30 sources. The two researchers separately read the same 10 articles to develop initial codes. This exercise was designed to extract key concepts, including the utilization of digital tools, feedback modes, and learner engagement. Initial codes, such as "technology tools," "feedback mechanisms," and "writing skill improvement," were created. Through a constant comparison method, differences between independent coding by the authors were ironed out through negotiated agreement. Inter-coder reliability was measured using percentage agreement, which was achieved at 87% during pilot coding. NVivo 14 software was used to manage the coding exercise and facilitate the development of themes.

After open coding, researchers split the remaining texts for axial coding and then collapsed related codes into larger categories. These categories were examined iteratively, with weekly meetings between the researchers to compare notes, refine code definitions, and clarify emerging themes. Through this refining process, themes began to emerge around recurring patterns in the data. To this point, both explicit (readily apparent) and implicit (conceptually inferred) themes were determined through close interpretive reading. In refining themes, the authors undertook an iterative process of theme development. The authors revisited each candidate theme based on the initial coded data to check internal coherence and external distinctiveness. Themes were either combined, divided, or redefined based on their representational strength and theoretical clarity. For instance, themes related to writing tools and feedback mechanisms were delineated into separate categories to highlight their specific pedagogical functions.

To ensure thematic consistency, a clear definition of and consistent labeling for each theme were established. The finalized themes are as follows (though the aim of this section is methodological and their content is not discussed here): (1) Digital Writing Tools, (2) Collaborative Writing Platforms, (3) Automated Feedback Systems, (4) Interactive Writing Environments, and (5) Virtual Writing Communities. These themes served as the organizing framework for making sense of the findings and investigating various ways that technology assists in writing development. The study integrated ongoing consultation with field practitioners, especially in the formulation and refinement of the coding scheme. This was accomplished through two formal feedback sessions, during which veteran writing instructors and educational technologists critiqued the initial coding schema and recommended improvements. Their feedback ensured the relevance and clarity of the codes and themes, enhancing the credibility of the study. The authors also noted the limitations of the coding process, including subjective thematic boundary interpretation and the exclusion of studies outside the targeted scope. Nonetheless, the systematic analytic process, diverse literature inclusion, and expert validation all combined to make the thematic framework rigorous and representative.

Validity and Reliability

The thematic analysis presented in this review is made more valid and robust through several methodological strategies. Triangulation incorporates data from various sources, including peer-reviewed

literature, conference papers, and grey literature. This approach confirms findings across settings, minimizing the possibility of source-specific bias and improving overall validity (Ntsobi et al., 2024). Even though the study employs a dual-author approach, researchers achieve inter-coder reliability checks through ongoing collaborative monitoring and cross-checking with external specialists on a subsample of the coded data. These processes ensure that the coding process remains consistent and that the themes identified are representative of the underlying data (Morgan et al., 2023). Still, specific limitations and potential biases do exist. Thematic analysis's subjectivity may lead to confirmation bias, where the researchers' presumptions influence the identification of themes. While triangulation and inter-coder reliability checks help alleviate this concern, dependence on literature released after 2020 may limit the historical coverage, risking the omission of previous formative studies. Furthermore, the exclusion of the work of several researchers may limit the range of interpretations. To remedy these weaknesses, the research approach entails extensive documentation of coding choices and regular consultations with experts, thereby enhancing the transparency and reproducibility of the research process. The combined approaches strengthen the study's validity and provide a balanced synthesis of current findings.

Results

Technology has emerged as a potent ally in the ever-evolving field of writing, providing numerous resources and tools to enrich the writing experience.

Digital Writing Tools

Digital Writing Tools (DWTs), such as online writing checkers (OWCs) and grammar websites like Grammarly, rank among the most commonly implemented technologies for writing instruction. These tools tend to be effective in enhancing surface-level mechanics, such as spelling and grammar. For example, McCarthy et al. (2022) found that OWCs led to minor improvements in grammatical accuracy but had a minimal impact on broader aspects of writing, such as argumentation and creativity. Grami (2020) notes that the success of using these tools often depends on students' prior exposure to and familiarity with digital writing assistance. Figure 2 shows usage rates and satisfaction levels across student populations by proficiency level. However, DWTs are not limitless. Their algorithmic design tends to overlook tone subtleties, audience awareness, or rhetorical sophistication. As a result, they can hamper instead of helping to foster critical thinking when used in sole reliance. For this reason, DWTs need to be placed as supplemental supports that help with revision but do not replace instruction in writing principles or higher-order thinking.

Figure 2: Integrating Technology in Education

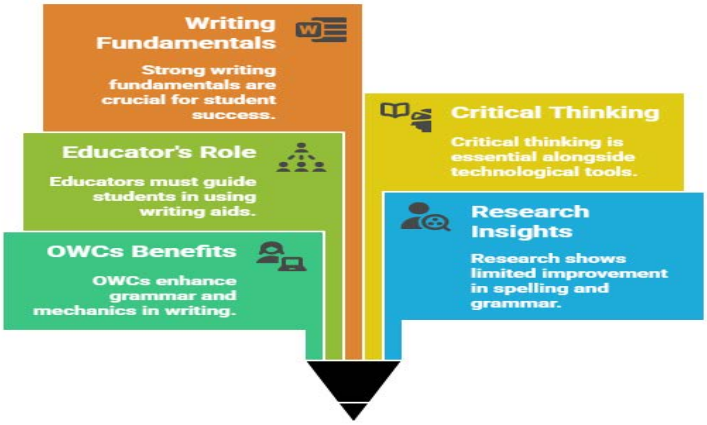


Figure 2. Integrating technology in education

ISSN: 2229-9327

Collaborative Writing Platforms

Collaborative Writing Platforms (CWPs) are characterized by their ability to facilitate real-time, synchronous co-authoring of texts by two or more users. Google Docs, wikis, and online forums are among the tools that enable students to construct collective documents, view changes, and negotiate meaning collaboratively. Zhang and Zou (2022) highlight how CWPs enhance learner motivation and engagement through a sense of "collective accountability." CWPs also foster peer feedback and discussion, which encourage learners to explain and defend their ideas, behaviors strongly associated with critical thinking. As Figure 3 illustrates, these collaborative editing environments significantly improve peer interaction frequency and depth. Zioga and Bikos (2020) found that EFL learners with CWPs demonstrated a better argumentation framework and evidence, possibly due to their exposure to various perspectives during peer collaboration. In the same vein, Ahmad (2020) discusses improvements in coherence and cohesion in essays produced through CWPs. Still, there are challenges. Unequal involvement, digital access, and ICT competency levels can constrain effectiveness (Jong & Tan, 2021). Nevertheless, CWPs provide an effective area for enabling students to challenge views, scrutinize claims, and integrate varied perspectives, all of which promote essential literacy.

Figure 3: Impact of Collaborative Writing Platforms

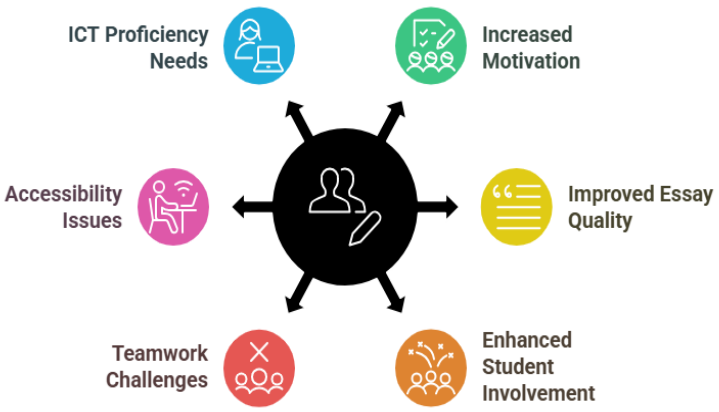


Figure 3. Impact of collaborative writing platforms

Automated Feedback Systems

Writing tools have transformed the way writing is taught and how individuals write (Graham, 2022). Automated Feedback Systems **(**AFSs) include software from basic grammar checkers to more sophisticated AI-powered tools such as Grammarly, Wordtune, and ChatGPT. These tools offer feedback on lexical range, structure, tone, and cohesion. McCarthy et al. (2022) illustrate that AFSs enable learners to self-correct and consider style choices. Notably, these tools allow the possibility of self-directed learning, developing metacognition and, in specific instances, more sensitive revisions. Yet, AI tool integration into feedback mechanisms provides both promise and peril. Chukwuere (2024) and Zhao et al. (2024) caution against overdependence, potential plagiarism, and ethical considerations, especially with generative AI such as ChatGPT. Figure 4 illustrates students' self-reported critical thinking levels when editing text with and without AI support. The use of AI, while potentially helpful, requires critical literacy competencies to review

suggestions instead of blindly accepting them. Woodworth and Barkaoui (2020) advocate for a hybrid feedback approach, where human judgment remains paramount. Therefore, AFSs, and especially those based on AI, can support higher-order revision competencies—if situated within a pedagogical model that infuses ethics, reflection, and teacher oversight.

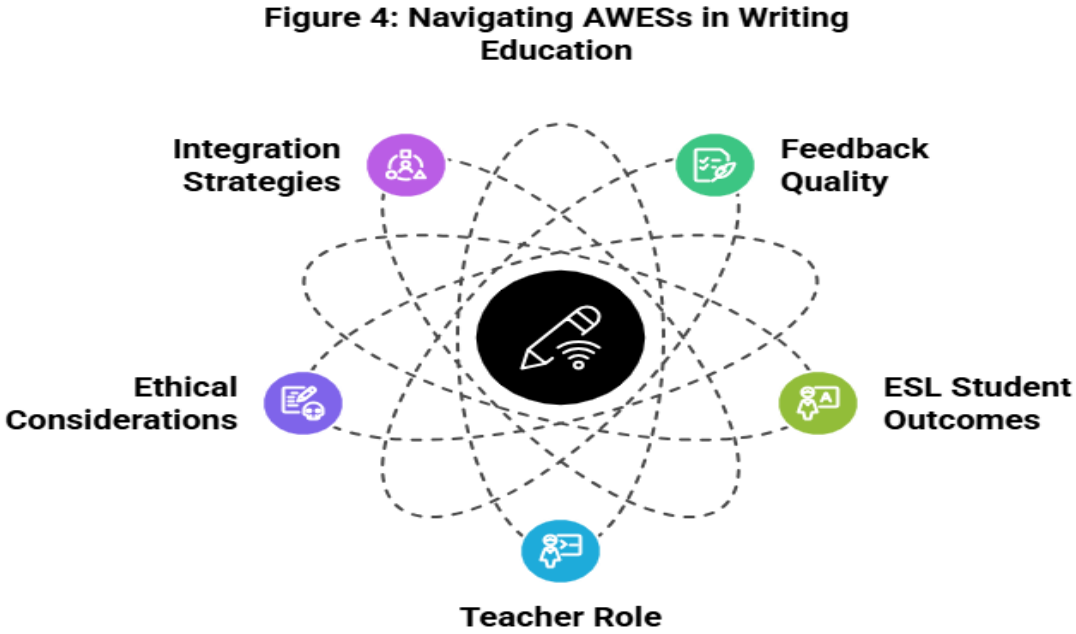


Figure 4. Navigating AFSs in writing education

Interactive Writing Environments

Interactive Writing Environments (IWEs) are defined by multimodal and reactive feedback systems, including gamified activities, real-time visual cues, and adaptive scaffolding tools that walk students through the writing process. In contrast to CWPs, which emphasize peer interaction, IWEs evoke writer-tool interaction. For instance, students can get immediate visual feedback on sentence variety, tone shifts, or organizational patterns. This encourages creative experimentation, particularly when students are invited to revise based on exploratory prompts instead of error correction exclusively. Ahmad (2020) found that IWEs enhanced the creative production of EFL students by exposing them to varied writing models and interactive exercises. Zhang and Zou (2022) also noted that IWEs enhanced student agency and motivation through interactive pathways within challenging writing tasks. Figure 5 illustrates a comparison of static digital tools and interactive environments in their influence on student idea generation. However, drawbacks involve some platforms' steep learning curve and variation in students' comfort with self-paced online settings. However, when carefully incorporated, IWEs facilitate creativity, metacognitive awareness, and more intense involvement with the writing process.

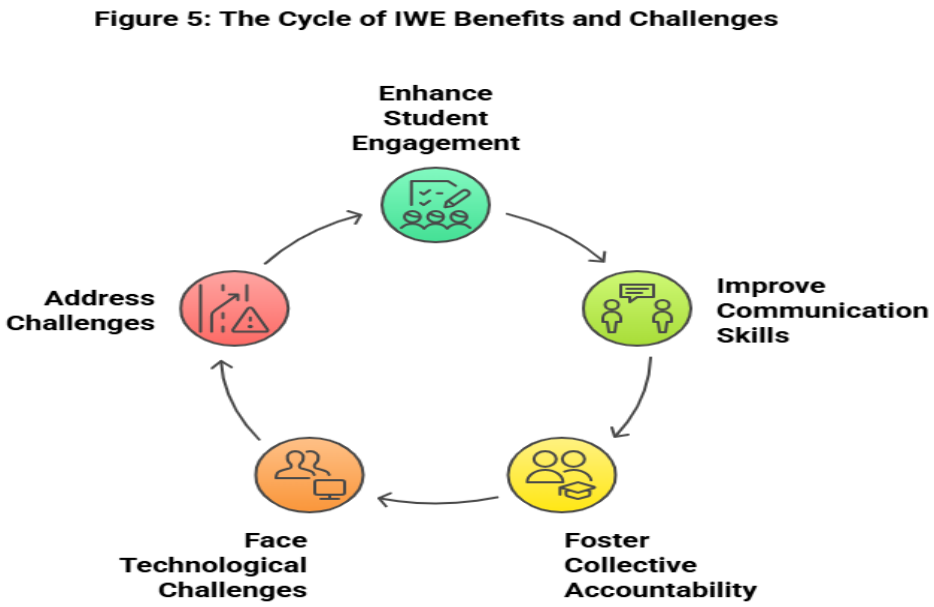


Figure 5. The cycle of IWE benefits and challenges

Virtual Writing Communities

Virtual Writing Communities (VWCs) are enduring, casual online environments where writers post, comment, and reflect on each other's writing, often extending beyond the classroom walls. VWCs can be found on websites such as Reddit, Wattpad, writing Discord servers, or institutional forums closed to the public. Unlike CWPs or IWEs, VWCs are more defined by sustained peer interaction, identity, and voluntary participation. Chairunisa et al. (2024) emphasize the role of digital literacy in enabling literature students to operate successfully within these communities. Nazari et al. (2021) examine students' affective engagement and self-concept in VWCs, noting increased confidence and control over writing. VWCs also often foster creativity, with users experimenting with voice, genre, and content in ways not always feasible in classroom writing. Trends in genre diversity in several VWCs are shown in Figure 6. Nevertheless, similar to AFSs, VWCs with AI writing partners also present ethical concerns. Chukwuere (2024) calls for integration models that support critical exploration, rather than passive reception. The role of community norms—especially regarding plagiarism, originality, and constructive feedback—is central to determining whether VWCs enhance or hinder critical thinking and creative agency.

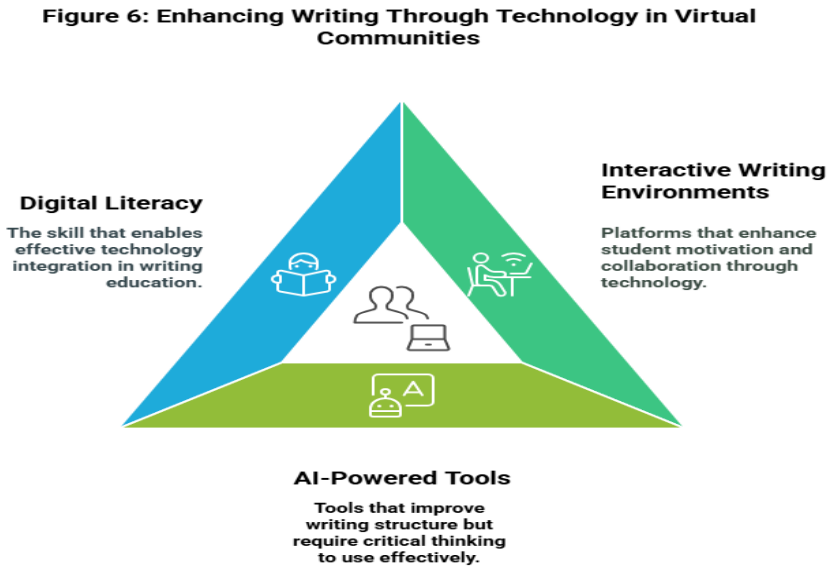


Figure 6. Enhancing writing through technology in virtual communities

To sum up, DWTs are also supportive of surface-level development but require teacher guidance to avoid stifling more profound competencies. CWPs are suitable for the collaborative development of critical thinking, especially in argumentative and analytical tasks. IWEs support creative thinking through interactive discovery and emergent feedback. AFSs are responsive in providing individualized feedback, with AI-supported systems showing potential for metacognitive growth but requiring ethical regulation. VWCs are offering informal, student-directed spaces that allow for creative risk-taking and community-based learning, mainly when supported by digital literacy.

Discussion

This research evaluated the impact of technology on the development of writing skills (Al Mahmud & Afzal Ur Rahmanu, 2023), specifically focusing on its influence on higher-order writing capacities, such as creativity and critical thinking. Although research has investigated the effect of computer-based tools on developing the lower-order mechanics of writing, such as grammar and punctuation (Grami, 2020; McCarthy et al., 2022), there is a lack of clear evidence on the impact of technology in facilitating the learning of higher-order cognitive processes. In addition, whereas past research has focused on general improvement in writing ability (Bhusal & Shrestha, 2024), less attention has been given to the variable needs of diverse learner groups, such as EFL learners and dyslexic writers. The rapidly evolving field of AI-based writing support software means its features under development are inadequately researched (Nazari et al., 2021; Al Balushi, 2024). Such gaps necessitate a thematic synthesis that, beyond accumulating evidence, provides a synthetic interpretation of the multifaceted way technology influences the creative and critical aspects of writing.

Thematic analysis of 40 studies produced five broad themes: Digital Writing Tools (DWTs), Collaborative Writing Platforms (CWPs), Automated Feedback Systems (AFSs), Interactive Writing Environments (IWEs), and Virtual Writing Communities (VWCs). The results show that DWTs, such as online writing checkers and grammar apps, improve mechanics

fundamentals and reduce errors. The research also showed that the tools alone cannot necessarily turn students into creative thinkers or enable them to develop critical thinking capacity (Grami, 2020; McCarthy et al., 2022). CWPs and IWEs showed a positive correlation with increased collaborative learning and motivation among students. In particular, the findings indicated that students who used CWPs, such as Google Docs, had a greater tendency to participate in reflective conversations and, consequently, improve their critical thinking activities (Zhang & Zou, 2022; Zioga & Bikos, 2020). Another significant finding is the greater use of AI writing tools. While these programs offer advanced feedback, including error correction and tips on style, tone, and content consistency, they also pose problems of dependency and potential plagiarism (Chukwuere, 2024; Zhao et al., 2024). The research results also suggest that teaching writing with technology has a much more preferred impact on digitally literate students, i.e., digital literacy is a key moderating variable in realizing the full potential of technological utilization.

RQ1: Effect of Technology on Advanced Writing Skills

This study's conclusions suggest that technology's influence on achieving higher-level writing skills is threefold. Computer programs enhance the mechanical aspects of writing and create environments that foster creative products and critical thinking. For instance, multimedia-enriched Interactive Writing Environments (IWEs) that offer immediate feedback facilitate higher-order creativity by testing mixed modes of narrative form and perspectives (Ahmad, 2020; Zhang & Zou, 2022). Theoretical models corroborate this, proposing that interactive digital environments can facilitate complex cognition by providing instant, context-based feedback (Chapelle, 2003). Nonetheless, the study also notes that while technology-based interventions facilitate creativity, they do not automatically result in greater critical thinking. Instead, acquiring essential skills appears to require an integration strategy where technology supports, but never replaces, conventional instructional methods (Woodworth & Barkaoui, 2020). This approach offers valuable novelty by encouraging educators to design technology-integrated curricula that promote higher-order thinking, rather than merely correcting errors.

RQ2: Supporting EFL Students and Dyslexic Writers Through Technology

One of the main contributions of this study is addressing the specific requirements of students, such as dyslexic authors and EFL learners. The discussed literature consistently emphasizes the promise of technology in providing individualized learning experiences. For instance, research has shown that computer writing aids and automated feedback systems (AFSs) provide customized support that may help dyslexic writers resolve precise issues with spelling and grammar (Marzuki et al., 2023). Research has shown that collaborative writing platforms (CWPs) benefit EFL learners by enabling simultaneous peer debate and criticism, thereby enhancing their linguistic proficiency and self-confidence (Ahmad, 2020; Zioga & Bikos, 2020). Understanding these diverse needs becomes increasingly crucial, as previous research has often overlooked the differentiated learning requirements of such groups (Bhusal & Shrestha, 2024). Through data triangulation across different technological tools, this paper asserts that adaptive learning environments, when complemented by AI writing support and other similar tools, can tailor learning to individual learners across various groups. However, the research cautioned that providing practical assistance to these students requires more than just offering e-tools; it also necessitates adequate training for both teachers and students to benefit from the technologies fully.

RQ3: Effectiveness of New AI Writing Tools

The efficacy of emergent technologies, particularly AI composition software, has become the central area of concern. Findings show that although AI instruments like ChatGPT and Wordtune yield significant gains in terms of providing stylistic recommendations and supporting the coherence of students' texts, their effectiveness is considerably dependent on learners' ability to critically review the feedback obtained (Chukwuere, 2024; Zhao et al., 2024). The research reaffirmed that the utilization of AI works best in the context of an overall pedagogical framework that emphasizes more human involvement. Existing research supports this, maintaining that AI should serve as an additive to, rather than a replacement for, standard teacher commentary (Woodworth & Barkaoui, 2020). Additionally, the potential concerns of AI abuse—such as causing addiction or facilitating immoral uses like plagiarism—underscore the necessity of well-defined ethical guidelines and pedagogy for mitigating these concerns. The evidence, therefore, necessitates a paradigm in which AI writing assistance catalyzes learning only when a rigorously designed pedagogy of critical thinking and appropriate use is in place to address its integration.

RQ4: Evidence-Based Treatments for Writing Skill Improvement

Finally, the study discusses what evidence-based interventions teachers and designers can provide to improve writing abilities. The discussion emphasizes the significance of various technological interventions in fostering an integrated learning environment. For example, integrating computer-mediated writing with CWPs and computer-mediated feedback systems can enhance a synergistic effect that enhances both technical correctness and creative expression. Evidence demonstrating blended learning environments yield higher writing performance outcomes informs this merging (Mahmood et al., 2024). Second, the study recommends that teachers implement a differentiated instruction model that utilizes technology to address the diverse learning needs of students, ensuring equal access to effective writing instruction (Al Balushi, 2024). Research from the included studies demonstrates that interventions, when combined with best practices in digital literacy and student-centered instruction, improve not only the mechanical processes of writing but also significantly enhance its strategic and reflective aspects.

Comparative Analysis with Prior Studies and Theoretical Perspectives

The current research largely aligns with prior studies in recognizing technology's contribution to improved writing performance (Alhujaylan, 2019; Tran & Nguyen, 2021; Chapelle, 2003). However, it contributes to prior research insofar as it focuses on higher-order writing skills and varying learning needs. Past studies have posited a bias towards the function of technology in reducing errors and the mechanics of writing. Still, this study integrates the knowledge that computer materials—when supplemented by collaborative and adaptive learning environments—can improve higher-order thinking and creative and critical thinking. For instance, while Grami (2020) achieved greater mechanical precision with online writing checkers, the current study concludes that the instruments' most significant value lies in their use as one component of a comprehensive instructional plan to foster higher-order abilities.

The study also corroborates sociocultural learning theories in illustrating how social contexts, such as CWPs and VWCs, afford knowledge construction through social negotiation (Li & Zhang, 2023; Rahimi & Fathi, 2021). Empirical research evidence corroborates the theory that peer review and group discussions can shape students' understanding of audience and genre,

thereby contributing to improved writing (Liu et al., 2022; Kim & Kim, 2020). With present argumentation, however, comes the caveat: the power of technology-facilitated writing instruction resides in part in the digital literacy of teachers and students. Without proper training and support, users will not optimize the leverage of these tools, a concern that Jong and Tan (2021) have concerning ICT access and skills. Furthermore, the conclusions of this study contribute to the ongoing discussion regarding the pedagogical and ethical concerns associated with such technologies. To the degree that AI technologies promote themselves as capable of delivering complex feedback and assisting with creative writing, their ability to spoil good sense if users do not critically examine their use poses an urgent threat (Chukwuere, 2024; Zhao et al., 2024). The situation demands a well-balanced, evidence-based answer that realizes the full potential of technological innovation without substituting for human judgment as the basic driving force of the writing process.

The present study is an in-depth study of how technology and electronic materials affect teaching writing. Discussing issues previously given short shrift, such as teaching writing skills acquisition and concerns regarding multicultural student groups, the research provides an equal assessment of the capabilities and weaknesses of technology in teaching writing. The study concludes that although computer writing assistant software and AI software can transform writing mechanics, their real potential lies in facilitating creative writing and critical thinking when utilized as part of an inclusive, student-centered pedagogical approach. Triangulation of issues validates that productive technology integration necessitates an in-depth examination of digital literacies, ethics, and the need to incorporate the human touch into machine-generated solutions. Developers and educators alike are interested in providing evidence-based, technology-aided interventions that increase writing capacity and stimulate critical thinking and creativity throughout a person's lifespan.

Implications

The results of this research have significant theoretical, practical, policy, and social implications, and they suggest several directions for future research. Theoretically, the results extend current theories of technology-enhanced writing pedagogy by highlighting the interconnections between digital tools and higher-order cognitive activities, such as creativity and critical thinking. This study challenges conventional models that place exclusive emphasis on the mechanical writing feature and validates emerging theories that recognize technology as a trigger for innovative skill acquisition (Chapelle, 2003; Woodworth & Barkaoui, 2020). In practical terms, educators can utilize this finding to develop integrated curricula that incorporate digital writing mechanisms, collaborative workspaces, and AI response mechanisms. In addition to boosting the mechanics of ordinary writing, such a curriculum also sparks imagination and critical thinking. For instance, incorporating interactive writing spaces can enhance learner involvement and provide targeted feedback tailored to individual learner needs (Ahmad, 2020; Nazari et al., 2021). In addition, designers of educational technology are motivated to develop usable and ethically sound AI writing support tools that enhance traditional instruction without replacing it (Chukwuere, 2024).

Policy implications entail state and institutional assistance in enhancing equitable access to technology in learning institutions. Decision-makers need to fund digital literacy packages and training initiatives for instructors to ensure they correctly use the potential of technology-enhanced learning. Adopting these types of policies allows closing the divide in technology-based access to

meet diverse populations with innovative features (Bhusal & Shrestha, 2024). Societally, integrating technology into writing lessons can level the playing field in education by enabling all students to become better writers. It fosters inclusive learning and equips learners for a digital world. Lastly, future studies should investigate the long-term implications of AI writing aids, the cross-cultural adoption of technology, and methods for striking a balance between automated and human feedback to continue exploring the development of critical thinking (Zhang & Zou, 2022; Zhao et al., 2024).

Limitations and Recommendations

Although this review provides insightful information regarding technology's potential to facilitate writing development, some limitations must be acknowledged, as they can influence interpretation and the scope of the findings. First, by overconcentrating on post-2020 research, the study may underrepresent pioneering research that continues to shape modern practices and pedagogies even now (Bhusal & Shrestha, 2024). This time-based limitation limits the contextual framework through which recent technological trends are considered. Second, although the thematic analysis was guided by a formal dual-author protocol and supported by professional consultation to ensure inter-coder reliability, the process remains interpretive in any event. Subjective bias in code selection and theme creation cannot be eliminated (Woodworth & Barkaoui, 2020). This can impact the consistency and impartiality of thematic outcomes.

Third, the exclusion of studies with unclear effects of technology on writing proficiency might have overlooked mixed-methods or qualitative studies that offer detailed, context-specific data (Nazari et al., 2021). In this way, some nuanced perspectives—particularly from varied learning environments—may be underrepresented. Fourth, even though systematic theme generation was possible under the analytical framework, it remains susceptible to partiality due to confirmation bias. Theoretical orientations of researchers may have influenced data interpretation in unapparent ways, even with iterative coding and open categorization methods (Chukwuere, 2024). Finally, the focus of the review on particular digital tools—e.g., AI writing assistants, collaborative spaces, and interactive environments—may limit the generalizability of the findings to broader or emerging technologies that are not highly controversial. These constraints suggest perceiving technology as a complement, rather than an alternative, to conventional pedagogy. Therefore, integration of digital tools should be followed by the establishment of critical thinking and digital literacy to enable effective and inclusive writing development.

Based on these limitations, future work would need to take a larger time frame. Covering historical and present-day work. Implementing a multi-author or multiple independent coder approach will strengthen inter-coder reliability and minimize subjective bias in coding. Broadening the criteria for data collection to include studies based on mixed-methodology and qualitative outcomes will also create richer and more representative analyses. Follow-up research could also investigate a more diverse range of computer-based instruments and platforms, such as contemporary ones, to ascertain their likely impact on lower-level and higher-order writing abilities. Ultimately, longitudinal study designs can reveal the long-term effects of technology integration on writing development. This approach will yield more substantial evidence for future practice and policy-making in the next century (Zhang & Zou, 2022; Zhao et al., 2024).

Conclusion

This research aimed to explore the impact of technology on advanced writing abilities- i.e., imagination and critical thinking—and computer-based software in assisting heterogeneous learner populations. The study stated that while computer programs improve rudimentary mechanics, their full potential emerges when combined with collaborative and interactive learning environments to stimulate higher-order thinking. Significant findings indicated that computer-based systems, collaborative platforms, and digital writing platforms have a positive impact on writing capacity. Nevertheless, they demonstrate varying degrees of effectiveness in enhancing higher-order skills such as critical thinking and creativity. Significantly, interactive writing environments and automated response systems enhance the rich potential for real-time learning and interaction but need critical human intervention to prevent over-reliance on technology (Nazari et al., 2021). Besides this, the research also introduced the need to accommodate the use of technology in a bid to address the needs of multi-modal learners like dyslexic and EFL learners, and also presented the spirit of adaptive, inclusive pedagogy (Zhang & Zou, 2022). The implications of these findings are tangible for theoretical, practical curriculum planning, and policy-making activities aimed at bridging digital divides. The disadvantage of being limited to post-2020 journals and a multi-author review design is that it minimizes generalizability. Future research should incorporate a wider temporal frame and utilize multiple independent coders to enhance reliability and investigate the long-term effects of emerging technologies on writing development (Woodworth & Barkaoui, 2020). In conclusion, technology holds great promise for transforming writing instruction; its practical use is contingent upon effective integration with conventional pedagogical practices to ensure that technology significantly develops both fundamental and advanced writing skills.

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: Artificial intelligence tools were utilized in the generation of the figures presented in this work.

About the Authors

Dr. Nagamurali Eragamreddy is a seasoned educator with 27 years of teaching experience, currently a senior lecturer at the University of Technology and Applied Sciences, Salalah, Sultanate of Oman. He holds a PhD in Linguistics and has extensive expertise in English Language Teaching (ELT), Linguistics. His research focuses on pragmatic markers, language skills development through technology, and the impact of semantics on communication. With over two decades of teaching experience, Dr. Eragamreddy has authored numerous publications in reputed journals, covering topics like Computer-Assisted Language Learning (CALL) and digital integration in language education. He actively contributes to international journals and continues influencing academic discourse through his scholarly work.
ORCID: https://orcid.org/0000-0003-1026-183X

Dr. Rachel Joseph holds a PhD in English Literature and has 25 years of teaching experience. Dr. Rachel is currently teaching at the University of Technology and Applied Sciences, Salalah, Sultanate of Oman. Email: rchl.jsph@gmail.com ORCID: https://orcid.org/0009-0006-5309-4321

References

Ahmad, S. Z. (2020). Cloud-based collaborative writing to develop EFL students' writing quantity and quality. *International Education Studies*, *13*(3), 51-64. https://doi.org/10.5539/ies.v13n3p51

Al Balushi, M. (2024). Effectiveness of Quizlet on the Writing Level of FFP Students in Oman. *Journal Of Teaching English for Specific and Academic Purposes*, *12(*1), 151–164. https://doi.org/10.22190/JTESAP230609013A

Alhujaylan, H. (2019). An assessment of the effectiveness of CALL in teaching English language writing skills in Saudi Arabia. *Arab World English Journal (AWEJ), Special Issue on CALL* (5). 18–27. https://dx.doi.org/10.24093/awej/call5.2

Aljuaid, H. (2024). The impact of Artificial Intelligence tools on academic writing instruction in higher education: A systematic review. *Arab World English Journal (AWEJ), Special Issue on ChatGPT*, April 2024:26-55. https://dx.doi.org/10.24093/awej/ChatGPT.2

Al Mahmud, F., & Afzal Ur Rahmanu, M. (2023). Academic writing of Saudi graduate students: Issues and improvements. *Arab World English Journal*, *14*(1), 409-427. https://dx.doi.org/10.24093/awej/vol14no1.26

Al-Shaboul, I. A., Kalsoom, T., Alshraah, S. M., & Khasawneh, M. A. S. (2023). Students' Perceptions on the effects of blogging sites in enhancing their motivation for foreign languages writing and reading proficiency. *Migration Letters*, *20*(S12), 157-168. https://doi.org/10.59670/ml.v20iS12.5869

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

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

Benlaghrissi, H., & Ouahidi, L. M. (2024). Effect of WhatsApp-based tasks on developing EFL Students' writing skills. *International Journal of Instruction*, *17*(2), 707-720. https://doi.org/10.29333/iji.2024.17239a

Bhusal, D., & Shrestha, K. (2024). Genre awareness in teaching writing: A case study of public High school teachers in Nepal. *Journal of Translation and Language Studies*, *5*(1), 10-22. https://doi.org/10.48185/jtls.v5i1.879

Caprara, L., & Caprara, C. (2022). Effects of virtual learning environments: A scoping review of Literature. *Education and Information Technologies*, *27*(3), 3683-3722. https://doi.org/10.1007/s10639-021-10768-w

Chairunisa, H., Harahap, M., & Anggreini, H. (2024). Utilisation of digital literacy to improve Writing skills of Indonesian literature students. In *Talenta Conference Series: Local Wisdom, Social, and Arts (LWSA)*, *7*(2), 154-160. https://doi.org/10.32734/lwsa.v7i2.2067

Chapelle, C. (2003). *English language learning and technology: Lectures on applied linguistics inthe age of information and communication technology*. Amsterdam: John Benjamin.

Chukwuere, J. E. (2024). Today's academic research: The role of ChatGPT writing. *Journal of*

Information Systems and Informatics, *6*(1), 30–46. https://doi.org/10.51519/journalisi.v6i1.639

Dousti, M., & Amirian, Z. (2023). The effect of web-mediated, blended, and purely online Learning on EFL learners' writing achievement in the Iranian context: A comparative study. *Education and Information Technologies*, *28*(2), 1675-1696. https://doi.org/10.1007/s10639-022-11215-0

Dendup, T., & Onthanee, A. (2020). Effectiveness of cooperative learning on English Communicative ability of 4th-grade students in Bhutan. *International Journal of Instruction*, *13*(1), 255–266. https://doi.org/10.29333/IJI.2020.13117A

Ellison, M. & Drew, C. (2020). Using digital Sandbox gaming to improve creativity within boys' Writing. *Journal of Research in Childhood Education*, *34*(2), 277-287. https://doi.org/10.1080/02568543.2019.1675823

Fang, X., Guo, K., & Ng, D. T. K. (2024). Sudowrite: Co-writing creative stories with Artificial Intelligence. *RELC Journal*, 1-5. https://doi.org/10.1177/00336882241250109

Geçkin, V., Kızıltaş, E., & Çınar, Ç. (2023). Assessing second-language academic writing: AI vs. Human raters. *Journal of Educational Technology and Online Learning*, *6*(4), 1096-1108. https://doi.org/10.31681/jetol.1336599

Goldman, S. R., Carreon, A., & Smith, S. J. (2024). Aligning the editing and revising of writing A strategy for technology support for students with learning disabilities. *Journal of Special Education Technology*, *39*(4), 567-575. https://doi.org/10.1177/01626434231211294

Graham, S. (2022). A walk through the landscape of writing: Insights from a program of writing research. *Educational Psychologist*, *57*(2), 55-72. https://doi.org/10.1080/00461520.2021.1951734

Grami, G. (2020). An evaluation of online and automated English writing assistants: Collocationsand idioms checkers. *International Journal of Emerging Technologies in Learning (IJET)*, *15*(4), 218-226. http://dx.doi.org/10.3991/ijet.v15i04.11782

Hallam, K. F. (2022). Moving on from trials and errors: A discussion on the use of a forum as an online focus group in qualitative research. *International Journal of Social Research Methodology*, *25*(4), 429-439. https://doi.org/10.1080/13645579.2021.1888402

Han, S. (2023). The contribution of blog-based writing instruction to enhancing writing performance and writing motivation of Chinese EFL learners. *Frontiers in Psychology*, *13*, 1069585. https://doi.org/10.3389/fpsyg.2022.1069585

Hazelton, L., Nastal, J., Elliot, N., Burstein, J., & McCaffrey, D. F. (2021). Formative automate Writing evaluation: A standpoint theory of action. *Journal of Response to Writing*, *7*(1), 37-91. https://scholarsarchive.byu.edu/journalrw/vol7/iss1/3

Hoang, D. T. N., & Hoang, T. (2024). Enhancing EFL students' academic writing skills in online Learning via Google Docs-based collaboration: A mixed-methods study. *Computer Assisted Language Learning*, *37*(7), 1504-1526. https://doi.org/10.1080/09588221.2022.2083176

Jong, B., & Tan, K. H. (2021). Using Padlet as a technological tool for the assessment of students' Writing skills in online classroom settings. *International Journal of Education and Practice*, *9(*2), 411-423. https://doi.org/10.18488/journal.61.2021.92.411.423

Kasih, E. N. E. W., Amelia, D., & Fitratullah, M. (2021, October). Unraveling students' attitudes

Using a blog for online learning in a literature class. In *2021, Universitas Riau International Conference on Education Technology (URICET)* (pp. 419-424). IEEE. https://doi.org/10.1109/URICET53378.2021.9865946

Kim, M. K., & Kim, S. M. (2020). Dynamic learner engagement in a wiki-enhanced writing Course. *Journal of Computing in Higher Education, 32(*3), 582–606. https://doi.org/10.1007/s12528-019-09248-5

Koltovskaia, S. (2023). Postsecondary L2 writing teachers' use and perceptions of Grammarly as a complement to their feedback. *ReCALL*, *35*(3), 290-304. https://doi.org/10.1017/S0958344022000179

Li, Y. (2023). The effect of online collaborative writing instruction on enhancing writing performance, writing motivation, and writing self-efficacy of Chinese EFL learners. *Frontiers in Psychology*, *14*, 1165221. https://doi.org/10.3389/fpsyg.2023.1165221

Li, M., & Zhang, M. (2023). Collaborative writing in L2 classrooms: A research agenda. *Language Teaching*, *56*(1), 94–112. https://doi.org/10.1017/S0261444821000318

Link, S., Mehrzad, M., & Rahimi, M. (2022). Impact of automated writing evaluation on the teacher Feedback, student revision, and writing improvement. *Computer Assisted Language Learning*, *35*(4), 605-634. https://doi.org/10.1080/09588221.2020.1743323

Liu, P. L., Ginting, A. M. G., Chen, C. J., & Yeh, H. C. (2022). Students' performance and Perceptions of wiki-based collaborative writing for learners of English as a foreign language. *Sage Open*, *12(*4), 1–12. https://doi.org/10.1177/21582440221144953

Mahmood, S., Sharif, H., & Aleem, F. (2024). Learning paragraph writing electronically: An Insight into the impact of blended learning strategies on ESL learners in Pakistan. *Forum for Linguistic Studies, 6* (1), 2057. https://doi.org/0.59400/fls.v6i1.2057

Mardiana, P. D., Sabiruddin, S., Gustia, A. Y., Yenis, M., & Langputeh, S. (2022). Forms of using The Internet supports student activities in online classroom learning in higher education. *International Journal of Islamic Studies Higher Education*, *1*(1), 14-29. https://doi.org/10.24036/insight.v1i1.90

Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools On the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*, *10*(2), 2236469. https://doi.org/10.1080/2331186X.2023.2236469

McCarthy, K. S., Roscoe, R. D., Allen, L. K., Likens, A. D., & McNamara, D. S. (2022). Automated writing evaluation: Does spelling and grammar feedback support high-quality writing and revision? *Assessing Writing*, *52*, 100608. https://doi.org/10.1016/j.asw.2022.100608

Miranty, D., Widiati, U., Cahyono, B. Y., & Sharif, T. I. S. T. (2025). The effectiveness of Grammarly application and teacher feedback for undergraduate EFL students' writing skills. *Innoeduca: International Journal of Technology and Educational Innovation*, *11*(1), 108-133. https://doi.org/10.24310/ijtei.111.2025.20521

Morgan, T. R., Gabler, C. B., & Manhart, P. S. (2023). Supply chain transparency: Theoretical Perspectives for future research. *The International Journal of Logistics Management*, *34*(5), 1422-1445. https://doi.org/10.1108/IJLM-02-2021-0115

Navarette, N., & Elan, H. J. (2025). The wite way forward: Learners' experiences on the utilization Of Artificial Intelligence tools in research writing. *International Journal of Language and Literary Studies*, *7*(3), 42-60. https://doi.org/10.36892/ijlls.v7i3.2111

Nazari, N., Shabbir, M. S., & Setiawan, R. (2021). Application of Artificial Intelligence

Powered digital writing assistant in higher education: Randomized controlled trial. *Heliyon*, *7*(5), e07014. https://doi.org/10.1016/j.heliyon.2021.e07014

Ntsobi, M. P., Costa, A. P., Kasperiuniene, J., Brandão, C., & Ribeiro, J. (2024, January). Digital Tools and techniques in qualitative research: Digital skills and research optimisation. In J. Ribeiro, C. Brandão, M. Ntsobi, J. Kasperiuniene, A. P. Costa (Eds.), Computer-supported qualitative research. In *World Conference on Qualitative Research* (pp. 1-25). Cham: Springer. https://doi.org/10.1007/978-3-031-65735-1_1

Pennington, M. C. (1996). *The computer and the non-native writer: A natural partnership*. Cresskill, NJ: Hampton Press.

Prayudi, R. A., Hakiki, A. K., Putra, N. R. D., Anzka, T. O., & Ihsan, M. T. (2021). The use of Technology in English teaching and learning process. *Jurnal Riset Dan Inovasi Pembelajaran*, *1*(2), 102–111. https://doi.org/10.51574/jrip.v1i2.38

Purwanto, A., Fahmi, K., & Cahyono, Y. (2023). The benefits of using social media in the The learning process of students in the digital literacy era and the education 4.0 era. *Journal of Information Systems and Management (JISMA)*, *2*(2), 1-7. https://doi.org/10.4444/jisma.v2i2.296

Rahimi, M., & Fathi, J. (2021). Exploring the impact of wiki-mediated collaborative writing on EFL students' writing performance, writing self-regulation, and writing self-efficacy: A mixed methods study. *Computer Assisted Language Learning*, *35*(9), 2627–2674. https://doi.org/10.1080/09588221.2021.1888753

Sari, E., & Han, T. (2022). Using generalizability theory to investigate the variability and Reliability of EFL composition scores by human raters and e-rater. *Porta Linguarum*: *International Journal of Foreign Language Teaching*, *38*, 27-45. https://doi.org/10.30827/portalin.vi38.18056

Schmieder, C., & Rogers, A. (2024). Using MAXQDA for a large qualitative dataset: Identifying Relevant data and analyzing it with thematic analysis. In S. Rädiker, & C. Gizzi (Eds.), *The practice of qualitative data analysis: Research examples using MAXQDA, Volume 2*, (pp. 127-149). Berlin: MAXQDA Press.

Tran, T. M. L., & Nguyen, T. T. H. (2021). The impacts of technology-based communication on EFL students' writing. *AsiaCALL Online Journal*, *12*(5), 54-76. https://www.asiacall.info/acoj/index.php/journal/article/view/80

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

Valizadeh, M. (2022). The effect of comprehensive written corrective feedback on EFL learners' written syntactic complexity. *Journal of Language and Education*. *8*(1), 196–208. https://doi.org/10.17323/jle.2022.12052

Warschauer, M. (2020). *Electronic literacies: Language, culture, and power in online education*. London: Routledge.

Wicaksono, F. A., Sulistyaningsih, S., & Syakur, A. (2022). A genre-based approach to improve The students' writing skills. *Journal of Basic Education*, *6*(3), 3391-3397. https://doi.org/10.31004/basicedu.v6i3.2622

Woodworth, J., & Barkaoui, K. (2020). Perspectives on using automated writing evaluation Systems to provide written corrective feedback in the ESL classroom. *TESL Canada Journal*, *37*(2), 234-247. https://doi.org/10.18806/tesl.v37i2.1340

Zewail-Foote, M., & Gonzalez, M. (2023). Crisscrossing learning experiences in an undergraduate Research-based laboratory course to promote reciprocal peer learning. *Journal of Chemical Education, 100*(3), 1092-1099. https://doi.org/10.1021/acs.jchemed.2c00341

Zhai, M. (2021). Collaborative writing in a Chinese as a foreign language classroom: Learners' perceptions and motivations. *Journal of Second Language Writing, 53*. 100836. https://doi.org/10.1016/j.jslw.2021.100836

Zhang, Z. (2023). Promoting student engagement with feedback: Insights from collaborative pedagogy and teacher feedback. *Assessment & Evaluation in Higher Education, 47*(4), 540-555. https://doi.org/10.1080/02602938.2021.1933900

Zhang, R., & Zou, D. (2022). Types, features, and effectiveness of technologies in collaborative Writing for second language learning. *Computer Assisted Language Learning, 35*(9), 2391-2422. https://doi.org/10.1080/09588221.2021.1880441

Zhao, X., Xu, J., & Cox, A. (2024). Incorporating Artificial Intelligence into student academic Writing in higher education: The use of Wordtune by Chinese international students. In *Proceedings of the 57th Hawaii International Conference on System Sciences* (pp. 2726-2735). https://hdl.handle.net/10125/106712

Zıoga, C., & Bıkos, K. (2020). Collaborative writing using Google Docs in primary education: Development of argumentative discourse. *Turkish Online Journal of Distance Education, 21*(1), 133-142. https://doi.org/10.17718/tojde.690372

Zou, D., Xie, H., & Wang, F. L. (2023). Effects of technology-enhanced peer, teacher, and self Feedback on students' collaborative writing, critical thinking tendency, and engagement in learning. *Journal of Computing in Higher Education, 35*(1), 166-185. https://doi.org/10.1007/s12528-022-09337-y

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.26-44
DOI: https://dx.doi.org/10.24093/awej/vol16no3.2

Self-Regulated Learning Strategies as Predictors of PERMA Well-being in MOOC-Based English Language Learning

Danyang Yue
Department of Language and Literacy, Faculty of Education
University of Malaya, Kuala Lumpur, Malaysia
Email: yuedanyang1996@gmail.com

Mohd Rashid Bin Mohd Saad
Department of Language and Literacy, Faculty of Education
University of Malaya, Kuala Lumpur, Malaysia
Corresponding Author: msaadmr@um.edu.my

Received: 06/09/2025 Accepted: 08/13/2025 Published: 09/20/2025

Abstract
As English massive open online courses (MOOCs) continue to expand in China, learners benefit from flexible access to language learning but also face risks to their psychological well-being due to limited interaction and high autonomy. This study aims to examine the extent to which self-regulated learning strategies (SRLS) predict five dimensions of well-being: Positive Emotion, Engagement, Relationships, Meaning, and Accomplishment, among 566 Chinese university students learning English through MOOCs. Guided by Seligman's PERMA model and social cognitive theory, the study examines how SRLS predict well-being and also explores learners' SRLS and well-being levels, as well as whether these levels differ by gender and academic year. A quantitative design was employed, and data were collected via online surveys and analyzed using structural equation modeling (SEM). Participants reported moderate levels of SRLS use and moderately high levels of well-being, with the highest scores in study environment management and Accomplishment. SEM revealed that SRLS significantly and positively predicted all five PERMA dimensions, especially Accomplishment and Engagement. However, gender and academic year did not show any significant influence on SRLS use or well-being outcomes. The findings show that SRLS contribute to psychological well-being and flourishing, and offer valuable guidance for creating English MOOCs that are both emotionally supportive and autonomy-oriented.

Keywords: English massive open online courses, self-regulated learning strategies, psychological well-being, positive emotion, engagement, relationships, meaning, accomplishment

How to cite: Yue, D. & Saad, M. R. M. (2025). Self-Regulated Learning Strategies as Predictors of PERMA Well-being in MOOC-Based English Language Learning. *Arab World English Journal*, *16*(3):26-44. https://dx.doi.org/10.24093/awej/vol16no3.2

Introduction

In recent years, MOOC-based English learning has become increasingly widespread in China due to its flexibility and accessibility (Yang & Li, 2024). However, its highly autonomous and self-paced nature often limits opportunities for real-time interaction and emotional support, which may lead to learner disengagement and emotional fatigue **(**Butler & Kern, 2016; Shkembi et al., 2024**)**. These concerns have brought psychological well-being to the forefront of online language education.

Positive psychology has been increasingly integrated into second language learning, giving rise to the concept of positive education in school settings (Cheng & Chen, 2021; Chen & Tang, 2024). Researchers have increasingly recognized emotional well-being as essential for fostering resilience, reducing anxiety, and sustaining motivation in second language (L2) contexts (Alqarni, 2022; MacIntyre & Gregersen, 2012; MacIntyre et al., 2019). Reflecting the broader positive turn in applied linguistics, recent scholarship has emphasized the value of incorporating positive psychology into foreign language education.

The PERMA model proposed by Seligman (2011) offers a five-dimensional perspective on well-being, encompassing Positive Emotion, Engagement, Relationships, Meaning, and Accomplishment. These elements are regarded as key pathways to psychological flourishing and have gained increasing attention in second language education research. Empirical research has consistently supported the model's effectiveness in improving learner outcomes (Oades et al., 2011; Turner & Thielking, 2019).

General education has widely adopted the PERMA model over the years (Hoare et al., 2017). Despite this, researchers in China have seldom addressed how the PERMA model applies to second language instruction. Notably, He (2021) investigated the relationship between emotional well-being and second language development in the context of learning Chinese. In the EFL domain, One of the earliest applications of the PERMA framework in English language instruction is found in Cheng and Chen's (2021) study. Nonetheless, the classroom-based nature of their study limits its applicability to digital learning formats such as MOOCs.

Applications of the PERMA model in online language learning are even more scarce. Although studies like Kiikeri et al. (2024) have integrated PERMA into digital education in Finland, few have focused specifically on language learners. In particular, research targeting Chinese EFL learners in MOOC-based English learning environments remains largely unexplored. These gaps highlight the need to explore how psychological well-being through the lens of PERMA can be supported in online English learning among Chinese university students.

In addition to well-being frameworks, SRLS including goal setting, monitoring, and help-seeking, have been widely acknowledged for promoting both academic achievement and emotional adaptation in digital learning contexts (Ly & Ihme, 2024; Shen & Teng, 2025; Zimmerman, 2000). Importantly, SRLS have been shown to enhance psychological well-being in EFL contexts (Derakhshan & Fathi, 2024). For example, goal setting can enhance learners' sense of accomplishment and support positive emotional experiences. In addition, teachers who integrate self-regulation strategies into EFL classrooms help students engage more meaningfully with feedback and improve their writing performance (Yang & Zhang, 2023). These connections suggest that SRLS may significantly contribute to learners' well-being, as outlined in the PERMA framework. However, how overall SRLS relates to specific well-being dimensions in MOOC-based English learning remains unclear.

Furthermore, demographic variables such as gender and academic level may shape learners' self-regulation and well-being (Bai & Wang, 2020; Dunn et al., 2014), but their effects in online EFL settings have received limited attention.

Thus, this study aims to examine the role of SRLS in predicting the well-being of Chinese university students learning English through MOOCs, drawing on Seligman's PERMA model and social cognitive theory. It also explores students' levels of SRLS and well-being, and whether these vary by gender and academic year.

The study put forward the following research questions to achieve research objectives: (1) What are the levels of SRLS and PERMA well-being among Chinese EFL learners in MOOCs? (2) Do SRLS and PERMA well-being levels differ by gender and academic year? (3) Do SRLS predict the five dimensions of PERMA well-being?

Literature Review

PERMA Well-being and SRLS in Online English Learning

Psychological well-being has become increasingly important in foreign language education, particularly in online settings where learners often face emotional fatigue and reduced motivation (MacIntyre & Gregersen, 2012). The PERMA model (Seligman, 2011) offers a structured approach to understanding well-being. Studies have shown its relevance in language learning: enjoyment and hope aid in overcoming challenges (Fredrickson, 2001), engagement fosters persistence (Kim, Song, & Lee, 2021), and meaning and achievement contribute to motivation (Sun & Wang, 2020). Many recent studies have focused on general or teacher education settings (Leng & Zhang, 2024). In contrast, Cheng and Chen (2021) conducted a study in a university EFL context. Their 14-week PERMA-based intervention significantly improved students' English listening and speaking skills. Li (2024) also found that engagement, meaning, and accomplishment predicted L2 grit dimensions, though neither grit nor PERMA significantly predicted exam scores. Nevertheless, studies exploring PERMA in online English learning remain limited.

Self-regulated learning strategies (SRLS) refer to "an active, constructive process whereby learners set goals for their learning and then attempt to monitor, regulate, and control their cognition, motivation, and behavior, guided and constrained by their goals and the contextual features in the environment" (Pintrich, 2000, p. 453). SRLS are especially important in online education environments such as MOOCs, where learners often lack direct teacher guidance and must rely on their planning and regulation to succeed. These strategies include goal setting, time management, self-monitoring, and help-seeking, and are linked to higher engagement and performance (Wang et al., 2023). In MOOC contexts, further refinements include self-evaluation, strategic planning, elaboration, persistence, and study environment management (Jansen et al., 2022; Kizilcec et al., 2017). However, many learners still struggle to sustain regulation in such settings (Vilkova, 2022). Recent research distinguishes between personal strategies (internal cognitive and motivational processes) and social strategies (peer and teacher interaction) (Wan et al., 2012). Given the typically solitary nature of MOOCs, learners tend to rely more on personal strategies, though discussion forums may offer some opportunities for social regulation.

Gender and Grade Level Differences in SRL Strategies and Psychological Well-being

Gender and grade level are frequently examined variables in studies of self-regulated learning and psychological well-being. Some international research suggests that female students

tend to report higher levels of both enjoyment and anxiety in academic settings (Dewaele & MacIntyre, 2014), but findings from Chinese EFL contexts are mixed. For example, Jiang and Dewaele (2019) and Shen and Teng (2025) reported no significant gender differences in emotional well-being among Chinese learners, including in online environments.

Regarding SRL, prior studies have indicated that female learners may use strategies like time management and help-seeking more frequently (Salahshour, Sharifi, & Salahshour, 2013). However, research from Chinese contexts has yielded inconsistent results. Shen and Bai (2024) found no gender-based differences in SRL in traditional classrooms, and Dinh and Phuong (2024) reported similar findings in MOOCs.

Grade-level differences have received less attention, with some studies suggesting that senior students develop more effective SRL strategies over time (Zhang, Gu, & Hu, 2008). Nevertheless, others have found no significant differences across academic years. For example, Shen and Teng (2025) and Ying, Shao, and Cui (2023) observed that both SRL and well-being levels remained relatively stable across grades in online English courses, possibly due to standardized curricula and limited interaction in MOOC environments.

Linking Self-regulated Learning and PERMA Well-being

Hypothesis Developed

Effective self-regulation is essential for sustaining both learning motivation and psychological well-being in highly autonomous environments like MOOCs (Rodríguez et al., 2022). The following sections explain how SRLS contribute to learner well-being across the five dimensions of the PERMA model.

Positive emotions such as enjoyment and hope often arise from learners' sense of control and achievement. Learners often use strategies like goal setting, time management, and self-monitoring to reduce anxiety and build confidence in self-paced learning (Fredrickson, 2001; Zhang & Fathi, 2024). When learners set their own goals and persist in completing tasks, they are more likely to experience higher emotional well-being and reduced stress (Bryce et al., 2020). Planning, closely tied to goal setting, also contributes to maintaining positive emotions over time.

H1: SRLS significantly and positively predict positive emotion

SRLS assist students to sustain engagement by making learners concentrate on their tasks, organize their schedules efficiently, and keep track of how their learning is progressing. These skills are especially important in autonomous MOOC environments (Teng & Zhang, 2022). Among Chinese EFL learners, SRL has been found to predict behavioral, emotional, and cognitive engagement. Especially help-seeking and time management show more powerful effects (Lan & Hew, 2020; Shen & Teng, 2025). These findings underscore the importance of incorporating planning, metacognitive guidance, and peer interaction into MOOC design to promote deeper and more sustained learner engagement.

H2: SRLS significantly and positively predict engagement.

Although MOOCs limit face-to-face interaction (Farikah et al., 2023), SRLS can foster social connectedness through strategic peer engagement. Help-seeking, collaborative planning, and constructive feedback are key behaviors that encourage active participation in forums and group tasks, leading to stronger feelings of belonging and interpersonal support. According to self-determination theory, such experiences fulfill learners' need for relatedness, which is critical for motivation and well-being (Derakhshan & Fathi, 2024; Ryan & Deci, 2000).

H3: SRLS significantly and positively predict relationships.

A strong sense of meaning often arises when learners perceive their English study as aligned with personal values or long-term goals. SRLS facilitate this alignment through goal setting, self-reflection, and purpose-driven planning. These strategies prompt learners to connect daily efforts with larger aspirations, reinforcing internal motivation and meaning over time (Seligman, 2011; Steger et al., 2006). In MOOC settings, where learning is largely self-regulated, such reflection may be particularly important for sustaining motivation and psychological growth.
H4: SRLS significantly and positively predict meaning.

Finally, the dimension of accomplishment is closely linked to SRL behaviors such as self-evaluation, strategic feedback use, and reflective review. These practices help learners track their development and acknowledge success, which strengthens their sense of efficacy and motivation (Sun & Wang, 2020). Building on these theoretical foundations, this study proposes a conceptual model to be empirically tested through structural equation modeling.
H5: SRLS significantly and positively predict accomplishment.

Self-regulated learning strategies (SRLS) contribute to accomplishment by fostering goal-directed behaviors, sustained effort, and self-monitoring. Zimmerman (2000) emphasized that learners who set clear goals, plan strategically, and monitor their performance tend to achieve better academic outcomes and greater satisfaction. Bai and Wang (2020) also found that EFL learners with high self-efficacy and intrinsic value made more effective use of SRLS, leading to improved language outcomes and stronger confidence. Similarly, Sun and Wang (2020) reported that students using self-regulated writing strategies performed better and felt a stronger sense of achievement. These findings suggest that SRLS not only support academic success but also promote a deeper, more subjective sense of accomplishment, aligning with the PERMA model (Seligman, 2011).

Conceptual Framework

Based on the previous research, this study proposes a conceptual framework that integrates SRLS and psychological well-being using the PERMA model, aiming to examine their predictive relationships in the context of English MOOC learning (see Figure 1).

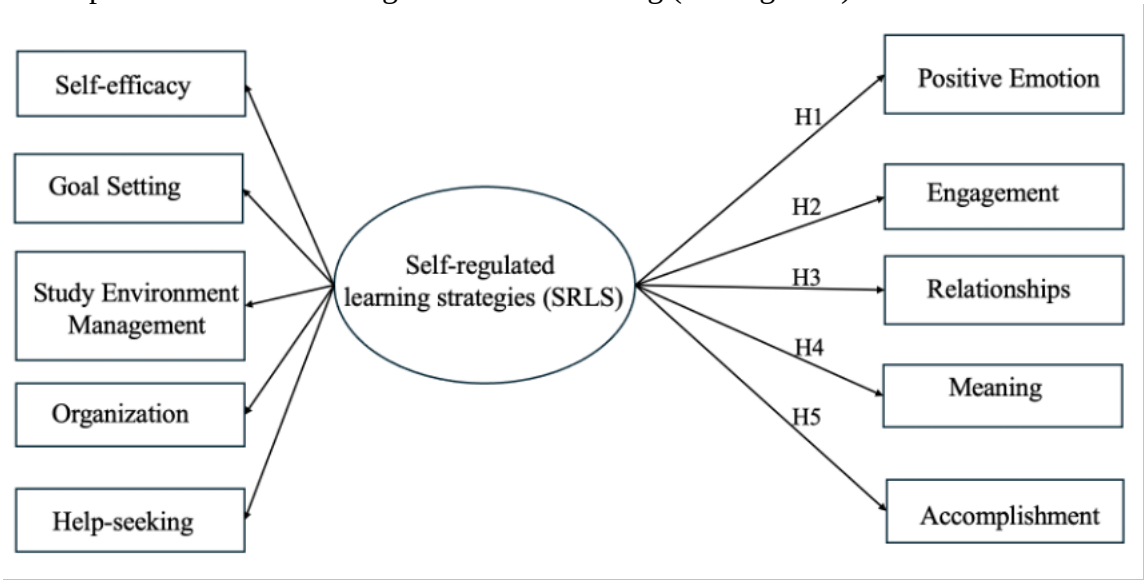


Figure 1. Conceptual framework illustrating the hypothesized relationships between SRLS and the five dimensions of PERMA well-being.

Method

Quantitative research design used to examine the relationship between self-regulated learning strategies (SRLS) and psychological well-being. The data were collected by using a cross-sectional survey from Chinese university students enrolled in a MOOC-based English course.

Participants

This study involved 566 undergraduate students from two public universities located in Northwest China. All participants were non-English majors who had prior experience learning English through MOOC platforms designed to prepare learners for standardized English proficiency tests. Participants were selected through convenience sampling. Among the 566 students, 408 were female (72.1%) and 158 were male (27.9%). In terms of academic year, the sample included 166 freshmen (29.3%), 225 sophomores (39.8%), 117 juniors (20.7%), and 58 seniors (10.2%). Participants were invited to fill an online questionnaire distributed via Wenjuanxing (a widely used online survey tool in China) during scheduled English class sessions.

Instruments

Self-Regulated Learning Questionnaire (MOOC-SRLQ)

To assess learners' use of SRLS in online English language courses, this study employed the Massive Open Online Course Self-Regulated Learning Questionnaire (MOOC-SRLQ) developed by Maldonado-Mahauad et al. (2020). Grounded in Pintrich's (2000) self-regulated learning theory, the instrument categorizes learners' strategic behaviors into five core dimensions: self-efficacy (six items), goal setting (five items), study environment management (three items), organization (four items), and help-seeking (four items). These dimensions reflect the three main categories of SRL strategies commonly used in empirical research: cognitive, metacognitive, and resource management strategies. Although initially designed for general MOOC learners, minor contextual adaptations were made to fit the current study. Specifically, the term "MOOC" was replaced with "English MOOC course" in all items to enhance relevance for Chinese EFL learners, while preserving the original meaning and structure of each item. Each items were assessed on a five-point Likert scale starting from one ("Not at all true for me") to five ("Very true for me"), with higher scores indicating stronger self-regulated learning behavior. The instrument demonstrated high internal consistency across all subscales in the present study.

PERMA Well-being Questionnaire

This study used the PERMA Well-being Questionnaire (Kiikeri, Uusiautti, & Purtilo-Nieminen, 2024) to measure learners' psychological well-being in CET-focused English MOOCs. The questionnaire is based on Seligman's (2011) PERMA theory. The instrument evaluates five core dimensions of well-being: Positive Emotion (five items), Engagement (five items), Positive Relationships (five items), Meaning (five items), and Accomplishment (six items). These dimensions reflect the holistic nature of well-being and represent a learner's emotional, motivational, and relational experiences in learning contexts. Although initially designed for general well-being assessment, minor contextual adaptations were made to fit the current study. Specifically, the wording of certain items was adapted to reflect the context of "English MOOC" to enhance contextual relevance for Chinese EFL learners, while retaining the original meaning and structure of each item. Each item was evaluated by participants on a four-point Likert scale ranging from one ("Describes me very poorly") to four ("Describes me very well"), with higher

scores indicating stronger well-being outcomes. The scale showed satisfactory to excellent internal reliability across all subscales.

Both questionnaires were translated into Chinese using the blind back-translation method (Brislin, 1970) to ensure semantic equivalence between the English and Chinese versions.

Table 1. *Descriptive statistics and correlations (N = 566).*

	1	2	3	4	5	6	7	8	9	10	11	12
SE	1											
GS	.625**	1										
SEM	.544**	.552**	1									
O	.463**	.481**	.410**	1								
HS	.584**	.562**	.524**	.488**	1							
SRLS total	.850**	.831**	.737**	.705**	.800**	1						
P	.211**	.142**	.232**	.154**	.232**	.243**	1					
E	.268**	.208**	.265**	.233**	.334**	.329**	.536**	1				
R	.179**	.163**	.171**	.188**	.221**	.232**	.426**	.552**	1			
M	.232**	.170**	.205**	.145**	.237**	.252**	.414**	.458**	.482**	1		
A	.288**	.268**	.245**	.219**	.378**	.356**	.476**	.526**	.468**	.525**	1	
PERMA total	.307**	.248**	.292**	.245**	.366**	.368**	.753**	.799**	.763**	.736**	.785**	1
Mean	3.189	3.297	3.663	3.106	3.381	3.298	2.736	2.94	2.684	2.85	2.905	2.826
SD	0.762	0.77	0.853	0.773	0.851	0.63	0.733	0.665	0.696	0.615	0.598	0.506
Cronbach's α	0.891	0.893	0.852	0.879	0.894	0.941	0.904	0.906	0.926	0.892	0.911	0.946

Notes**:** SE = Self-Efficacy; GS = Goal Setting; SEM = Study Environment Management; O = Organization; HS = Help-Seeking; P = Positive Emotion; E = Engagement; R = Relationships; M = Meaning; A = Accomplishment; SRLS = Self-Regulated Learning Strategies; PERMA = PERMA well-being.
($p < .01$.)

Research Procedure

To fulfill the aims of this study, an online questionnaire was distributed via Wenjuanxing with support from English instructors at two universities. Before participation, students were informed of the study's purpose, assured of confidentiality, and provided electronic informed consent. Ethical approval was obtained from the Universiti Malaya Research Ethics Committee. The survey took approximately 10–15 minutes to complete and included validated measures of SRLS and PERMA well-being. All responses were collected anonymously and exported into SPSS for data cleaning and analysis.

Data Analysis

Refer to data analysis, first, internal consistency (Cronbach's alpha) and confirmatory factor analyses (CFA) were conducted using AMOS 26.0 (Arbuckle, 2019) to assess the reliability and validity of the SRLS and PERMA scales. Second, descriptive statistics and MANOVA were used to examine gender and grade-level differences in SRLS and PERMA dimensions. Third, Pearson correlations were computed to explore relationships between SRLS and PERMA variables. Finally, SEM was used to test the hypothesized model, in which SRLS predicted PERMA well-being, which in turn mediated its effect on English MOOC learning outcomes.

Results

Reliability of The Measures

This study used the PERMA well-being scale and the SRLS questionnaire, both of which demonstrated satisfactory to high internal reliability. The Cronbach's alpha values for the five PERMA subscales ranged from .892 to .946, indicating excellent internal consistency. Similarly, the five SRLS subscales showed good reliability, with alpha values between .852 and .894 (see Table 1). The overall Cronbach's alpha of the total SRLS scale reached .941, further supporting the consistency of the instrument. These results suggest that all items within each subscale measured the intended constructs with adequate internal consistency.

Confirmatory Factor Aanalysis (CFA)

CFA were conducted to evaluate the factorial validity of the measurement models. The first-order model of SRLS, comprising five latent factors—Self-Efficacy, Goal Setting, Study Environment Management, Organization, and Help-Seeking—showed acceptable fit: $\chi^2/df = 2.73$, RMSEA = .055, 90% CI [.053–.058], CFI = .902, TLI = .897, SRMR = .065, GFI = .800. All factor loadings exceeded .50 and were statistically significant. A second-order CFA was then performed as shown in Figure 2, specifying SRLS as a higher-order construct reflected by the five dimensions. The model demonstrated good fit: $\chi^2(204) = 509.50$, $\chi^2/df = 2.50$, RMSEA = .051, 90% CI [.046–.057], CFI = .961, TLI = .956, GFI = .925, AGFI = .907. Inter-factor correlations ranged from .48 to .69, supporting the second-order structure.

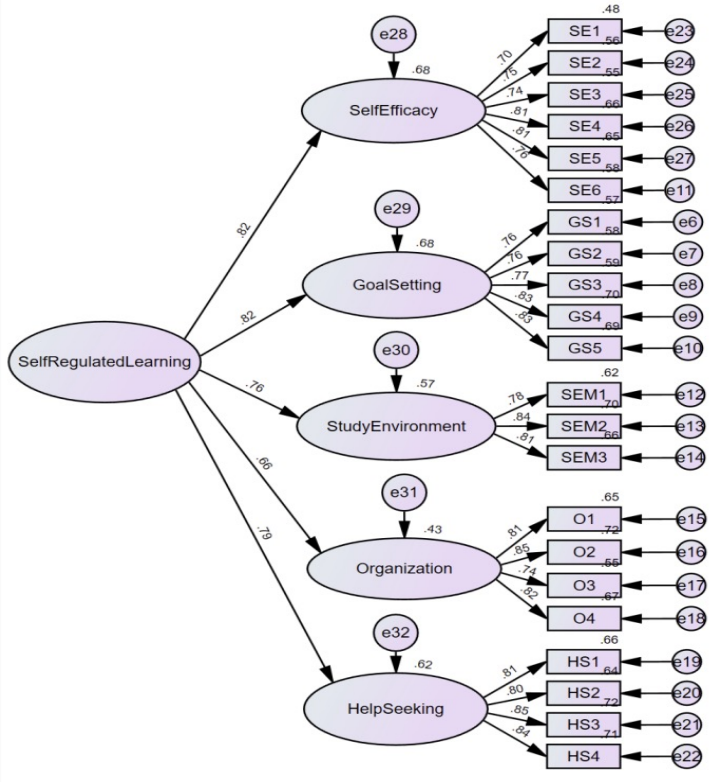


Figure 2. Confirmatory Factor Analysis Model of SRLS
Note: SE = self-efficacy; GS = goal setting; SEM = study environment management; O = organization; HS = help-seeking.

The CFA for the PERMA well-being construct in Figure 3, modeled with five first-order factors—Positive Emotion, Engagement, Relationships, Meaning, and Accomplishment—also demonstrated acceptable fit: $\chi^2/df = 2.97$, RMSEA = .059, 90% CI [.055–.064], CFI = .946, TLI = .939, GFI = .886, SRMR = .053. These findings indicate that both the SRLS and PERMA measurement models are valid and reliable for use in subsequent structural equation modeling.

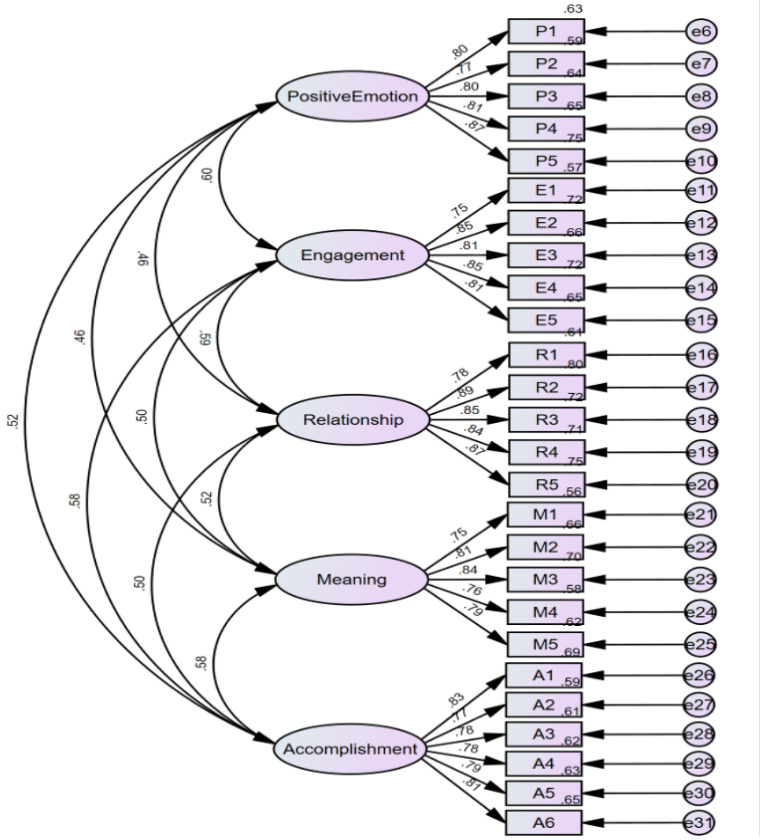


Figure 3. Confirmatory Factor Analysis Model of PERMA Dimensions
Note: P = positive emotion; E = engagement; R = relationship; M = meaning; A = accomplishment.

Overall Level of SRL Strategy Use And PERMA Well-being Level

Following Pimentel's (2010) five-point Likert scale, scores ranging from 3.41 to 5.00 represent high levels, 2.61 to 3.40 indicate moderate levels, and 1.81 to 2.60 reflect low levels. For the four-point PERMA scale, Alico and Guimba (2015) classify 3.25–4.00 as high, 2.50–3.24 as moderately high, 1.75–2.49 as moderately low, and 1.00–1.74 as low.

SRLS was measured on a five-point scale. Participants reported a medium level of SRLS use (M = 3.30, SD = 0.63), falling between "Somewhat true" and "Quite true" (see Table 1). Among the subscales, Study Environment Management had the highest mean score (M = 3.66), followed by Help-Seeking (M = 3.38) and Goal Setting (M = 3.30). Self-Efficacy (M = 3.19) and Organization (M = 3.11) were slightly lower but still moderate.

PERMA was assessed on a four-point scale. Participants showed a moderately high overall well-being level (M = 2.83, SD = 0.51). Engagement (M = 2.94), Accomplishment (M = 2.91), and Meaning (M = 2.85) were highest, while Positive Emotion (M = 2.74) and Relationships (M = 2.68) were relatively lower, approaching the mid-range.

These results suggest moderate use of SRLS and moderately high PERMA well-being. This analysis addresses RQ1, which examined learners' overall levels of SRLS and PERMA well-being.

Differences Across Genders And Grade Levels

As regards gender and grade level differences, ANOVA results showed that neither gender nor grade level had significant effects on the total SRLS and PERMA scores. For SRLS, no significant gender differences were observed, $F(1, 564) = 0.039$, $p = .845$, nor were there significant grade level differences, $F(3, 562) = 0.577$, $p = .630$. Similarly, for PERMA well-being, gender differences were not significant, $F(1, 564) = 0.008$, $p = .927$, and grade level differences were also not significant, $F(3, 562) = 1.015$, $p = .386$. These results suggest that both male and female students across different academic years reported comparable levels of SRLS use and PERMA well-being in MOOC-based English learning.

Table 2. *ANOVA results for the effects of gender and grade level on SRLS and PERMA*

Independent Variable	Dependent Variable	*F*	Sig.	Partial eta squared
Gender	SRLS	0.039	.845	.000
Grade level	SRLS	0.577	.630	.003
Gender	PERMA	0.008	.927	.000
Grade level	PERMA	1.015	.386	.005

These results suggest that both male and female students across different academic years reported comparable levels of SRLS use and PERMA well-being in MOOC-based English learning. These results directly address Research Question 2 (RQ2), indicating that no significant differences were found in students' SRL strategy use or PERMA well-being levels across gender or academic grade levels.

Relationships Between SRLS And PERMA Well-being

As illustrated in Table 1, Pearson correlation analysis identified statistically significant and positive associations between all five components of SRLS and the PERMA dimensions of well-being. The correlation coefficients ranged from $r = .142$ to .378 ($p < .01$), suggesting that learners who more frequently engaged in SRLS tended to report higher levels of well-being. Help-Seeking (HS) showed the strongest correlations with Accomplishment ($r = .378$), Engagement ($r = .334$), and Positive Emotion ($r = .232$), all significant at $p < .01$. Organization (O) showed relatively weaker correlations with the PERMA dimensions.

This study employed SEM to test the predictive effects of SRLS on PERMA. As shown in Figure 4, the model demonstrated an acceptable fit to the data: $\chi^2(1066) = 2892.18$, $\chi^2/df = 2.71$, RMSEA = .055, 90% CI [.053, .057], CFI = .903, TLI = .898.

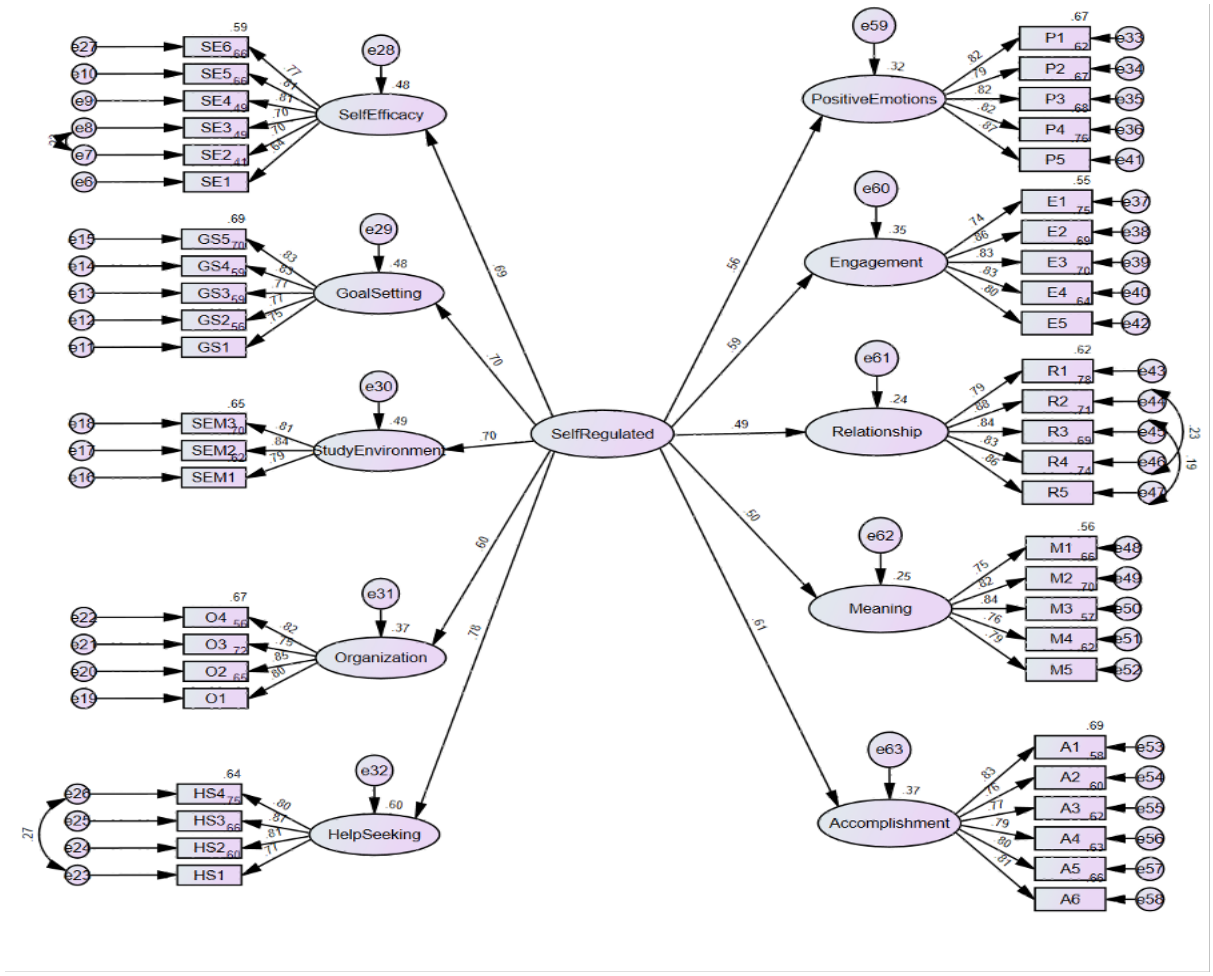


Figure 4. Structural Equation Model (SEM) Showing the Predictive Effects of SRLS on PERMA Well-being Dimensions

Path analysis revealed that SRLS significantly and positively predicted all five dimensions of PERMA well-being: Positive Emotion (β = .441, $p < .001$), Engagement (β = .541, $p < .001$), Relationships (β = .439, $p < .001$), Meaning (β = .462, $p < .001$), and Accomplishment (β = .566, $p < .001$), confirming the hypothesized positive influence of SRLS on PERMA well-being in the context of English MOOCs.

Table 3. *Summary of hypothesis testing results for SRLS → PERMA well-being dimensions*

Path	β (Standardized)	*p*-value	Hypothesis	Supported?
SRLS → Positive Emotion	0.441	< .001	H1	Yes
SRLS → Engagement	0.541	< .001	H2	Yes
SRLS → Relationships	0.439	< .001	H3	Yes
SRLS → Meaning	0.462	< .001	H4	Yes
SRLS → Accomplishment	0.566	< .001	H5	Yes

These results provide empirical support for Research Question 3, demonstrating that SRLS exert significant positive effects on Positive Emotion, Engagement, Relationships, Meaning, and Accomplishment. All five proposed hypotheses were statistically confirmed.

Discussion

Overall levels of SRLS use and PERMA well-being

This study found that Chinese university students reported a moderate level of SRLS use in MOOC-based English courses. The result is consistent with prior findings in traditional classrooms (Shen et al., 2023). However, differences were found in the use of specific strategies. They found that persistence and self-evaluation were the most frequently used strategies. In contrast, the current study identified that study environment management was used most commonly, followed by help-seeking and goal setting. This pattern aligns with findings from Dinh and Phuong (2024) in MOOC contexts. This difference may be attributed to the autonomous nature of MOOCs, which require learners to manage their study environments and rely more on digital interaction. The high use of environment management reflects efforts to stay focused in flexible but unstructured settings. And help-seeking is common because online tools like forums and peer support replace real-time teacher feedback. (Dai, Lin, & Liu, 2024).

Participants also reported a moderately high level of overall PERMA well-being. Accomplishment and engagement emerged as the most prominent dimensions. The results imply that learners were deeply immersed in their studies and felt a genuine sense of achievement. In contrast, positive emotion and relationships scored lower, suggesting limited emotional satisfaction and weaker interpersonal support, likely due to the solitary, asynchronous nature of MOOCs. This pattern aligns with previous research showing that online learners often maintain strong task focus but struggle with emotional fulfillment and social connectedness (Lan & Saad, 2024; Miranda et al., 2024). Challenges such as language anxiety and self-doubt (Wang & Zhang, 2021) may further reduce emotional well-being. Compared to findings from face-to-face settings, Li (2024) found that learners reported higher levels of emotional and relational well-being in face-to-face settings. In contrast, the current results emphasize the role of learning modality and learner context. The reduced real-time interaction and sense of community in MOOCs may explain the relatively lower levels of emotional and social well-being observed in this study.

Differences across genders and grade levels

This study found no significant gender or grade-level differences in SRLS or PERMA well-being among Chinese university students in MOOC-based English learning. These findings align with recent research suggesting that gender effects on SRL are generally minimal in online environments (Shen & Bai, 2024; Sun & Wang, 2020). Although earlier classroom-based studies noted that female students often used more planning and time management strategies (Bai, Shen, & Mei, 2020), such differences appear less evident in digital contexts.

Similarly, there is no significant gender differences were observed in emotional well-being, supporting prior findings that male and female Chinese EFL learners report comparable levels of academic emotions, such as enjoyment and anxiety (Jiang & Dewaele, 2019; Li, 2022). This study contrasts with international studies, which often report more substantial emotional variation among females (Dewaele & MacIntyre, 2014). The discrepancy may reflect cultural influences: in collectivist systems like China's, standardized expectations and the depersonalized nature of MOOCs may suppress gender-based emotional differences. Korlat et al. (2021) found higher engagement among Austrian female students in more socially differentiated learning systems.

Likewise, no grade-level differences were found in SRLS or well-being, which is consistent with studies showing that Chinese undergraduates across academic years engage with similar curricula and assessments (Shen et al., 2023; Ying et al., 2023). Emotional development

may vary with age, but such differences may be muted within the relatively narrow span of university education.

In sum, gender and academic year appear to exert limited influence on SRLS and PERMA well-being in MOOC settings. Future research could explore how learner traits such as motivation, digital competence, or self-efficacy interact with demographic variables in shaping learning and well-being outcomes.

Relationships between SRLS and PERMA well-being

The findings indicate that higher levels of SRLS are significantly associated with better outcomes across all five PERMA dimensions. This is consistent with social cognitive theory, which highlights the role of self-regulation in fostering persistence, positive emotion, and a sense of accomplishment (Bandura, 1991). Among the five dimensions, Accomplishment indicated the strongest association with SRLS. This may be attributed to the strong emphasis on test performance and achievement-focused attitudes commonly observed among Chinese EFL learners. In self-paced MOOC environments, students who regulate their goals, time, and progress may experience a stronger sense of competence and success. Engagement was also strongly predicted by SRLS, underscoring the role of self-regulation in sustaining psychological immersion and focus, even in large-scale, low-interaction settings. This supports prior research highlighting the motivational function of SRL in online contexts (Shen & Teng, 2025; Tang, 2024). Meaning and Positive Emotion were positively linked to SRLS as well. Learners who plan intentionally and reflect on their goals are more likely to find personal relevance in their tasks and manage stress effectively, consistent with Fredrickson's broaden-and-build theory (2001). While the relationship factor was still statistically relevant, it showed the weakest link among all variables. This may be due to the asynchronous and socially restricted nature of MOOCs. In such environments, SRLS often struggle to promote meaningful peer interaction unless collaborative elements are deliberately built in.

Previous research has mainly addressed the cognitive and behavioral aspects of SRLS, whereas this study highlights its emotional and psychological significance. Overall, SRLS serve dual functions in the context of MOOCs: supporting goal-directed engagement and promoting emotional resilience. Future research may further explore how specific SRLS components influence each PERMA dimension and how socially enriched course designs can enhance learners' relational well-being.

Overall, these findings highlight the role of SRLS as a significant predictor of learner well-being across various dimensions in English-language MOOCs. These findings provide empirical support for the study's hypotheses. Specifically, all five proposed hypotheses were confirmed that SRLS significantly and positively predicted positive emotion, engagement, relationships, meaning, and accomplishment. This reinforces the theoretical assumption that SRLS are not only cognitive and behavioral enablers, but also important predictors of learners' emotional and psychological well-being in online language learning environments.

Implications And Directions for Future Research

This study advances our understanding of how SRLS influence the well-being of Chinese EFL learners within MOOC-based English learning environments, using the PERMA model as a theoretical lens. Theoretically, the research bridges self-regulated learning and positive psychology by demonstrating how cognitive and motivational strategies contribute to

psychological well-being in digital language education. It extends the PERMA framework to autonomous online contexts and highlights SRLS as a key mechanism for sustaining learner flourishing beyond academic performance. Practically, the findings underscore the need to design MOOC environments that structurally support SRLS at multiple levels. This involves more than offering isolated strategy prompts. Instead, platforms should embed SRLS-enhancing features directly into the course structure, including goal-setting interfaces, real-time reflection checkpoints, progress dashboards, and feedback systems responsive to learners' emotional needs. These design features should function synergistically to foster learner autonomy, emotional engagement, and task persistence. Moreover, developers and instructors should consider balancing cognitive demands with social and affective supports. While MOOCs often emphasize content delivery, greater attention is needed to support learners' emotional and motivational regulation, especially in the absence of real-time teacher and peer interaction. This could be achieved by aligning SRLS features with personalized learning paths, adaptive challenges, and scaffolded peer interaction modules.

Future research should consider longitudinal or experimental designs to explore the dynamic relationship between SRLS and well-being. Integrating multimodal data such as system logs and screen activity could also offer insights beyond self-reported measures (Ifenthaler & Yau, 2020). Further studies may examine potential mediators such as self-efficacy and emotional regulation, and moderators such as digital literacy and instructional style that influence the SRLS–PERMA relationship. Cross-cultural investigations could clarify contextual influences. Lastly, intervention studies evaluating SRLS training programs in MOOC settings may help enhance both learner well-being and academic persistence, offering practical implications for online course design and learner support.

Conclusion

This study examined how SRLS predict PERMA well-being among Chinese EFL learners in MOOC-based English courses. Results showed moderate SRLS use and moderately high well-being, with the highest scores in study environment management and Accomplishment. The analysis revealed no significant differences across gender or academic year. Structural equation modeling confirmed that SRLS significantly predicted all five PERMA dimensions, with the strongest effects on Accomplishment and Engagement. The results suggest that SRLS are crucial for learners' psychological well-being. They also highlight the need to combine cognitive and emotional support in English MOOCs. Designing platforms that cultivate SRLS while addressing learners' well-being can help promote more engaged, resilient, and successful EFL learners in digital contexts.

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 ChatGPT-4o only to assist with editing and improving language clarity and readability during the writing process. After using the tool, the authors reviewed and edited the content as needed and took full responsibility for the content of the publication.

About the Authors:
Danyang Yue is a PhD candidate in the Department of Language and Literacy Education, Faculty of Education, University of Malaya. Her research focuses on self-regulated learning strategies and positive psychology in English as a Second Language contexts, particularly in relation to students' English achievement in online learning platforms. ORCID: https://orcid.org/0009-0005-4406-2849

Mohd Rashid Bin Mohd Saad (PhD) is a Senior Lecturer in the Department of Language and Literacy Education, Faculty of Education, University of Malaya. He has over 25 years of teaching experience in higher education. His research focuses on TESL issues, trends, and emerging technologies for non-native EFL teachers. ORCID: https://orcid.org/0000-0002-1050-294X

Reference

Alico, J., & Guimba, W. (2015). Level and causes of pre-university students' English test anxiety: a case study on Mindanao State University. *Researchers World Journal of Arts, Science & Commerce*, *3*(1), 1-10. https://www.researchersworld.com/index.php/rworld/article/view/632

Alqarni, N. A. (2022). Investigating the Relationship between Emotional Well-being and Grit as Predictors of Saudi EFL Female Students' Foreign Language Achievement. *Arab World English Journal*, *13*(3), 498–511. https://dx.doi.org/10.24093/awej/vol13no3.32

Arbuckle, J. L. (2019). *IBM SPSS Amos 26 user's guide*. IBM Corporation.

Bai, B., & Wang, J. (2020). The role of growth mindset, self-efficacy, and intrinsic value in self-regulated learning and English language learning achievements. *Language Teaching Research*, *27*(1), 207–228. https://doi.org/10.1177/1362168820933190

Bai, B., Shen, B., & Mei, H. (2020). Hong Kong primary students' self-regulated writing strategy use: Influences of gender, writing proficiency, and grade level. *Studies in Educational Evaluation*, *65*, 100839. https://doi.org/10.1016/j.stueduc.2020.100839

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes*, *50*(2), 248–287. https://doi.org/10.1016/0749-5978(91)90022-L

Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, *1*(3), 185–216. https://doi.org/10.1177/135910457000100301

Bryce, C. I., Alexander, B. L., Fraser, A. M., & Fabes, R. A. (2020). Dimensions of hope in adolescence: Relations to academic functioning and well-being. *Psychology in the Schools*, *57*(2), 171–190. https://doi.org/10.1002/pits.22311

Butler, J., & Kern, M. L. (2016). The PERMA-Profiler: A brief multidimensional measure of flourishing. *International Journal of Wellbeing*, *6*(3), 1–48. https://doi.org/10.5502/ijw.v6i3.526

Chen, S., & Tang, L. (2024). EFL teachers' engagement: The impact of well-being and emotion regulation. *Heliyon*, *10*(5). https://doi.org/10.1016/j.heliyon.2024.e27338

Cheng, M. , & Chen, P. (2021). Applying PERMA to Develop College Students' English Listening and Speaking Proficiency in China. *International Journal of English Language and Literature Studies*, *10*(4), 333–350. https://doi.org/10.18488/journal.23.2021.104.333.350

Dai, Y., Lin, Z., & Liu, A. (2024). Facilitating Students' Adaptive Help-seeking and Peer Interactions through an Analytics-enhanced Forum in Engineering Design Education. *Procedia CIRP*, *128*(1), 292-297.

Derakhshan, A., & Fathi, J. (2024). The Interplay between Perceived Teacher Support, Self-regulation, and Psychological Well-being among EFL Students. *Iranian Journal of Language Teaching Research*, *12*(3), 113-138. https://www.magiran.com/p2798617

Dewaele, J., & MacIntyre, P. D. (2014). The Two Faces of Janus? Anxiety and Enjoyment in the Foreign language Classroom. *Studies in Second Language Learning and Teaching*, *4*(2), 237-274. https://doi.org/10.14746/ssllt.2014.4.2.5

Dinh, C., & Phuong, H. (2024). Examining student Characteristics, Self-Regulated Learning Strategies, and their perceived effects on satisfaction and academic performance in MOOCs. *The Electronic Journal of e-Learning*, *22*(8), 41–59. https://doi.org/10.34190/ejel.22.8.3446

Dunn, K. E., Rakes, G. C., & Rakes, T. A. (2014). Influence of academic self-regulation, critical thinking, and age on online graduate students' academic help-seeking. *Distance Education*, *35*(1), 75–89. https://doi.org/10.1080/01587919.2014.891426

Farikah, F., Mulyani, M., Astuty, A., & Mazid, S. (2023). Student Engagement in Virtual Learning: The Self-Determination Theory Perspective. *Arab World English Journal*, *14*(3), 182–198. https://doi.org/10.24093/awej/vol14no3.12

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, *56*(3), 218–226. https://doi.org/10.1037/0003-066X.56.3.218

He, F. (2021). *Learner well-being and its impact on second language acquisition of Chinese as a second language* (Doctoral dissertation, Minzu University of China). CNKI. https://doi.org/10.27667/d.cnki.gzymu.2021.000647

Hoare, E., Bott, D., and Robinson, J. (2017). Learn it, Live it, Teach it, Embed it: Implementing a whole school approach to foster positive mental health and wellbeing through positive education. *Intentional Journal of Wellneing*, *7*(3), 56–71. https://doi.org/10.5502/ijw.v7i3.645

Ifenthaler, D., & Yau, J. Y. (2020). Utilising learning analytics to support study success in higher education: a systematic review. *Educational Technology Research and Development*, *68*(4), 1961–1990. https://doi.org/10.1007/s11423-020-09788-z

Jansen, R. S., Van Leeuwen, A., Janssen, J., & Kester, L. (2022). Exploring the link between self-regulated learning and learner behaviour in a massive open online course. *Journal of Computer Assisted Learning*, *38*(4), 993–1004. https://doi.org/10.1111/jcal.12675

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

Kiikeri, P., Uusiautti, S., & Purtilo-Nieminen, S. (2024). Students' Thriving and Well-being in Online Learning Environments in Vocational Education and Training. *International Journal for Research in Vocational Education and Training*, *11*(1), 119–145. https://doi.org/10.13152/ijrvet.11.1.6

Kim, H. E., Song, H. D., & Lee, Y. C. (2021). The effect of academic self-efficacy, perceived usefulness, perceived ease of use, and acceptance attitude on learning persistence through

student engagement in MOOC. In *EdMedia+ Innovate Learning* (pp. 892-897). Association for the Advancement of Computing in Education (AACE).

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education, 104*, 18–33. https://doi.org/10.1016/j.compedu.2016.10.001

Korlat, S., Kollmayer, M., Holzer, J., Lüftenegger, M., Pelikan, E. R., Schober, B., & Spiel, C. (2021). Gender Differences in Digital Learning During COVID-19: Competence Beliefs, Intrinsic Value, Learning Engagement, and Perceived Teacher Support. *Frontiers in psychology*, *12*, 637776. https://doi.org/10.3389/fpsyg.2021.637776

Lan, M., and Hew, K. F. (2020). Examining learning engagement in MOOCs: a self-determination theoretical perspective using mixed method. *Int. J. Educ. Technol. Higher Educ. 17*, 1–24. https://doi.org/10.1186/s41239-020-0179-5

Lan, Y., Saad, M.R.B.(2025). Emotional Presence in the Community of Inquiry: Addition of PERMA in Online English Teaching and Learning. *Asia-Pacific Edu Res 34*(2), 723–737. https://doi.org/10.1007/s40299-024-00891-w

Leng, L., & Zhang, L. J. (2024). Exploring positive emotions and teaching Effectiveness in language Education: Insights from the PERMA model. *International Journal of Applied Linguistics*. https://doi.org/10.1111/ijal.12642

LI Geng-chun (2024). A Study on the Relationships among English Major Students' Well-being, L2 Grit, and English Language Achievement. *Journal of Qiqihar University, 10*, 115-119. https://doi.org/10.13971/j.cnki.cn23-1435/c.2024.10.022

Li, C. (2022). Foreign language learning boredom and enjoyment: The effects of learner variables and teacher variables. *Language Teaching Research*, *29*(4), 1499-1524. https://doi.org/10.1177/13621688221090324

Ly, N., & Ihme, M. (2024). A Regularized-Interface method as a unified formulation for simulations of High-Pressure multiphase flows. *Journal of Computational Physics*, *518*(6), 113310. https://doi.org/10.1016/j.jcp.2024.113310

MacIntyre, P. D., Gregersen, T., & Mercer, S. (2019). Setting an agenda for positive psychology in SLA: Theory, practice, and research. *Modern Language Journal, 103*(1), 262–274. https://doi.org/10.1111/modl.12544

MacIntyre, P. D., & Gregersen, T. (2012). Affect: The role of language anxiety and other emotions in language learning. In S. Mercer, S. Ryan, & M. Williams (Eds.), *Psychology for language learning: Insights from research, theory & practice* (pp. 103–118). Palgrave Macmillan. https://doi.org/10.1057/9781137032829_8

Maldonado-Mahauad, J., Pérez-Sanagustín, M., & Beyle, C. (2020). A questionnaire for measuring self-regulated learning in massive open online courses. In *Proceedings of the 2020 XLVI Latin American Computing Conference (CLEI)* (pp. 400–409). IEEE.

Miranda, J. O., Tulabut, O. A. P., De Ala, C. L., Gonzales, A., Embalsado, J. V. M., Estacio, L. V., et al. (2024). The PERMA model of well-being and student engagement amid Covid-19: A two-wave study among a sample of Filipino university students. *Psychology in the Schools, 61*(11), 4347–4360. https://doi.org/10.1002/pits.23286

Oades, L. G., Robinson, P., Green, S., & Spence, G. B. (2011). Towards a positive university. *The Journal of Positive Psychology*, *6*(6), 432–439. https://doi.org/10.1080/17439760.2011.634828

Pimentel, J. L. (2010). A note on the usage of Likert Scaling for research data analysis. *USM R&D Journal, 18*(2), 109-112.

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. 452–502). Academic Press.

Rodríguez, S., González-Suárez, R., Vieites, T., Piñeiro, I., & Díaz-Freire, F. M. (2022). Self-Regulation and Students Well-being: A Systematic Review 2010–2020. *Sustainability*, *14*(4), 2346. https://doi.org/10.3390/su14042346

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

Salahshour, F., Sharifi, M., & Salahshour, N. (2013). The relationship between language learning strategy use, language proficiency level, and learner gender. *Procedia - Social and Behavioral Sciences, 70*(1), 634–643. https://doi.org/10.1016/j.sbspro.2013.01.103

Seligman, M. E. P. (2011). *Flourish: A visionary new understanding of happiness and well-being.* Free Press.

Shen, B. & Bai, B. (2024). Chinese university students' self-regulated writing strategy use and EFL writing performance: influences of self-efficacy, gender, and major. *Applied Linguistics Review*, *15*(1), 161-188. https://doi.org/10.1515/applirev-2020-0103

Shen, B., Wang, Y., Yang, Y., & Yu, X. (2023). Relationships between Chinese university EFL learners' academic emotions and self-regulated learning strategies: A structural equation model. *Language Teaching Research*, *0*(0). https://doi.org/10.1177/13621688221144832

Shen, X., & Teng , L. S. (2025). Unveiling the effect of self-regulated learning strategies on learning engagement among Chinese EFL learners in the *online* learning environment. *Digital Applied Linguistics, 3*. https://doi.org/10.29140/dal.v3.102473

Shkembi, F., Sauku, V. H., Ibrahimi, S., & Ibrahimi, E. (2024). The psychological impact of online education on student well-being and motivation. *International Social Sciences and Education Journal*, *2*(2), 01–08. https://doi.org/10.61424/issej.v2i2.122

Steger, M. F., Frazier, P., Oishi, S., & Kaler, M. (2006). The meaning in life questionnaire: Assessing the presence of and search for meaning in life. Journal of Counseling Psychology, *53*(1), 80–93. https://doi.org/10.1037/0022-0167.53.1.80

Sun, T., & Wang, C. (2020). College students' writing self-efficacy and writing self-regulated learning strategies in learning English as a foreign language. *System*, *90*, 102221. https://doi.org/10.1016/j.system.2020.102221

Tang, H. (2024). Understanding self-regulated learning and learner performance in MOOCs. *Distance Education*, *45*(2), 244–259. https://doi.org/10.1080/01587919.2024.2338712

Teng, L. S., & Zhang, L. J. (2022). Can self-regulation be transferred to second/foreign language learning and teaching? Current status, controversies, and future directions. Applied Linguistics *43*(3), 587-595. ISSN 0142-6001.

Turner, K., & Thielking, M. (2019). Teacher wellbeing: Its effects on teaching practice and student learning. *Issues in Educational Research*, *29*(3), 938–960. http://www.iier.org.au/iier29/turner2.pdf

Vilkova, K. (2022). The promises and pitfalls of self-regulated learning interventions in MOOCs. *Technology, Knowledge and Learning: Learning mathematics, science and the*

arts in the context of digital technologies, *27*(3), 689–705. https://doi.org/10.1007/s10758-021-09580-9

Wan, Z., Compeau, D.R., & Haggerty, N. (2012). The Effects of Self-Regulated Learning Processes on E-Learning Outcomes in Organizational Settings. *Journal of Management Information Systems*, *29*(1), 307 - 340. https://doi.org/10.2753/MIS0742-1222290109

Wang, C., Zhu, S., & Zhang, H. (2023). Computer-assisted English learning: Uncovering the relationship between motivation and self-regulation. *Journal of Computer Assisted Learning*, *39*(6), 1860-1873. https://doi.org/10.1111/jcal.12846

Wang, X., & Zhang, W. (2021). Psychological Anxiety of College Students' Foreign Language Learning in Online Course. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.598992

Yang, L. F., & Zhang, L. J. (2023). Self-regulation and student engagement with feedback: The case of Chinese EFL student writers. *Journal of English for Academic Purposes*, *63*(8), 101226. https://doi.org/10.1016/j.jeap.2023.101226

Ying, S., Shao, J. R., & Cui, Q. (2023). A study on college students' self-regulated learning strategies in foreign language MOOCs. *Overseas English*, (06), 16–18.

Zhang, L. J., & Fathi, J. (2024). The mediating role of mindfulness and emotion regulation in the relationship between teacher self-compassion and work engagement among EFL teachers: A serial mediation model. *System*, *125*, 103433. https://doi.org/10.1016/j.system.2024.103433

Zhang, L. J., Gu, P. Y., & Hu, G. (2008). A cognitive perspective on Singaporean primary school pupils' use of reading strategies in learning to read in English. *The British journal of educational psychology*, *78*(2), 245–271. https://doi.org/10.1348/000709907X218179

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. *Handbook of self-regulation* (pp. 13–39). Academic Press. https://doi.org/10.1016/B978-012109890-2/50031-7

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp. 45-59
DOI: https://dx.doi.org/10.24093/awej/vol16no3.3

Students and Instructors' Views on the Use of Portfolios in Teaching Writing at the Tertiary Level in Saudi Arabia

Hussein Taha Assaggaf
Department of Foreign Languages
College of Arts and Humanities, Al-Baha University
Kingdom of Saudi Arabia
Email: hassggaf@bu.edu.sa

Received:06/17/2025 Accepted:08/16/2025 Published: 09/20/2025

Abstract

The portfolio is an effective alternative assessment technique that has recently been used in various English as a Foreign Language (EFL) context. This study aims to explore the views of both students and writing instructors in using portfolios in teaching writing for Saudi EFL undergraduates. Participants are both undergraduate students and writing instructors at the English department, Albaha University, Saudi Arabia. The main methods for data collection are a questionnaire, group discussions, and interviews. The findings showed overall positive views towards portfolio use. It has also been found that the portfolio was considered an effective tool to develop several writing aspects such as motivation, confidence and autonomy. The findings also revealed that students and instructors perceived a few obstacles to portfolio implementation, such as traditional learning practices and program grading systems. The study concludes with recommendations for better implementation of portfolios.

Keywords: Alternative assessment, English as a foreign language, portfolios, portfolio assessment, tertiary level in Saudi Arabia, writing

How to cite: Assaggaf, H. T. (2025). Students and Instructors' Views on the Use of Portfolios in Teaching Writing at the Tertiary Level in Saudi Arabia. *Arab World English Journal*, *16*(3):45-59. https://dx.doi.org/10.24093/awej/vol16no3.3

Introduction

Teaching writing in EFL contexts has long been a concern for researchers and educators. Research findings have disclosed that writing instruction in EFL contexts was lacking. These were attributed to numerous factors, including, reliance on outdated techniques (Assaggaf et al., 2012), old educational policies (Ezza, 2010), and first language ineffective writing instruction (Al-Khuweileh & Al Shoumali, 2000). Just like most EFL contexts, writing instruction in Saudi Arabia has heavily relied on conventional techniques and methodologies (Alkodimi & Al-Ahdal, 2021). Previous research has related that to several considerations such as the use of teacher-centered approaches (Alshawi & Alshumaimeri, 2017), lack of adequate assessment training (Alotabi, 2014), and lack of assessment skills (Al-Abdulkareem, 2004).

To construct better writing environments, alternative assessment techniques have been adopted globally at various levels. One of the techniques that has been highly recommended is portfolio. According to Fox (2017), portfolios have recently played a prominent role in writing "to mark a shift from product-oriented pedagogies to process-oriented ones"(p. 138). Fenwick & Parsons (1999) highlight three advantages of using portfolios in teaching writing. Firstly, portfolios boost students' confidence as learners can monitor their progress over time. Another significant benefit is their ability to foster better teacher-student conversations, as both the teacher and the student maintain long-term interaction on students' written assignments. Thirdly, portfolios enhance learning experience which ultimately leads to improved outcomes. Furthermore, Fox (2017) lists three evident benefits: learner awareness, autonomy, and reflection.

To examine these values, plenty of research on portfolios has been conducted in many EFL contexts (e.g., Apeadido et al., 2024 [Ghana]; Arslan & Gümüş, 2020 [Turkey]; Guo & Li, 2024 [China]; Kalra et al., 2017 [Thailand]; Lam, 2013 [Hong Kong]; Laraba & Boukhedimi, 2021 [Algeria]; Ma'arif et al., 2021 [Indonesia]; Quayyum, 2022 [Pakistan]). Some of these studies focus on portfolio assessment and writing skills development (e.g., Qayyum, 2022); some on students' perceptions (e.g., Laraba & Boukhedimi, 2021); and some others on AI-assisted portfolio assessment (e.g., Khasawneh et al., 2025). As regarding the Saudi context, even though portfolio assessment has previously been examined at Saudi schools (e.g., Alabdelwahab, 2002), very few have been explored at a Saudi university (Alhuwaydi, 2017). This study, therefore, seeks to close the current research gap and add to the existing knowledge in this area.

In spite of the revolutionary procedures that the Saudi educational authorities have been going through due to the Saudi Vision 2030, some have reported a shortage in the writing assessment (e.g., Alshawi & Alshumaimeri, 2017; Alhuwaydi, 2017). In fact, students at the university need to experience the latest assessment methods to help them develop self-regulated learners to promote the notion of lifelong learning (Van der Schaaf et al., 2012). This will ensure the optimal realization of the state's ambitious vision and facilitate the achievement of its goals. This study, therefore, intends to investigate the implementation of a portfolio as an alternative assessment technique to enhance the learning of writing at the tertiary level. The primary focus, however, is on reflecting both students' and instructors' voices in using portfolios in the teaching and learning of writing. Specifically, this study seeks to understand both students' and instructors' views on implementing portfolios in the teaching writing for Saudi EFL undergraduates. Hence, the study is guided by two research questions:1) What are students' views on using portfolios in teaching writing? and 2) What are instructors' views on using portfolios in teaching writing?

The findings of this study are expected to be significant for the writing instructors and policy makers at the tertiary level in the Saudi context and perhaps other EFL contexts.

Literature Review

Alternative Assessment

Over the years, assessment of writing courses has relied on conventional methods in which a student would respond to a teacher's question by doing specific timed tasks and receive immediate feedback. According to Mabry (1999), from the beginning of the nineties of the last century, teachers started to look for better methods. Later, a variety of new assessment methods were experienced. These included checklists, rating scales, rubrics, journals, logs, videotapes and audiotapes, portfolios, conferences, diaries, self-assessment, and peer-assessments. Such new techniques provide alternatives to the traditional techniques, which some educators started to criticize as being overwhelming to students (Klenowski, 2002). Indeed, there have been claims that learners are being over-examined (Kohn, 2000; Au, 2011). Hence, alternative techniques such as portfolios were believed to provide opportunities to reduce any imbalance caused by the traditional testing methods and to produce far better results. (Klenowski, 2002).

Portfolios in Writing Courses

A portfolio is a dynamic tool whose characteristics vary significantly based on its purpose, reasons and audience. This makes it hard to decide on specific definitions or guidelines for what constitutes a portfolio. One of the early definitions of portfolio is presented by Moya & O'Malley (1994). It states that a portfolio is "a collection of a student's work, experiences, exhibitions, self-ratings" (p.2). A portfolio is also defined as a file that contains "repositories of artifacts (e.g., reflections, works in progress, self- and peer- assessments, final products) assembled over time as evidence of development, learning, or capability." (Fox, 2017, p. 138). In other words, a portfolio is seen as the mechanism and steps that require students to maintain systematic records of their academic work in a folder or box to be assessed and commented on later by instructors and class peers (Brown, 2005). For classifying portfolio, the literature demonstrates a diversity of criteria which resulted in a wide range of portfolio types. For instance, Burke et al., (1994) classify portfolios as personal, academic, and professional. Danielson & Abrutyn (1997) classify portfolios into working, display, and assessment. Lam (2018) presented three types of portfolios: progress (growth), working (efforts), and showcase (achievement) portfolio.

Over time, research on a portfolio has identified various benefits to language learning in general and writing in particular. For instance, a portfolio has been considered a tool to enhance students' motivation (Al-Hidabi et al., 2019; Muho & Leka, 2021; Pitri, 2021). When students work constantly on their assignments and other documents to produce better drafts with the help of their instructors and peers, this makes them more motivated. Furthermore, a portfolio has also been found effective in developing students' self-reflection (Boumediene et al., 2018; Zubizarreta, 2009). By self-reflection practice, learners develop numerous skills such as developing more profound insights into their writing abilities as well as fostering independence, which is one important goal in teaching writing.

Given the significance of portfolio in teaching writing, many scholars establish principles to enhance the effectiveness of portfolio as a learning tool. For instance, Nunes (2004) outlined two fundamental principles. The first is to facilitate continuous interaction between the teacher and learners, and the second is to enhance documenting students' reflections to evaluate their experience in later stages. Indeed, writing instructors need to create efficient strategies that encourage deep communication between themselves and their students, and allow students to express their views on the new experience they come across.

Furthermore, Hamp-Lyons & Condon (2000) highlight several characteristics to contribute to best implementations of portfolio in writing. These are nine essential components of a well-designed portfolio: collection, range, context richness, delayed evaluation, selection, student-centered control, self-evaluation and reflection, growth along parameters, and development over time. The importance of collection, reflection, and selection were emphasized by Hamp-Lyons & Condon, particularly in creating a portfolio with a meaning and purpose. They also underlined that adding more of these components would improve the caliber and scope of assessment, transforming the portfolio into a more thorough evaluation tool.

Relevant Studies

Several research studies have been conducted on portfolio implementations for undergraduates in EFL contexts. For instance, Martinez-Lirola & Rubio (2009) conducted two studies at two universities in Spain to explore students' opinions of portfolio use. For data collection, they used two questionnaires: one for quantitative data and another for qualitative data. The findings showed that the use of portfolios was found beneficial mainly because it led to effort empowerment and daily basis work. The findings also revealed that the students recognized learning efficacy and competency development as strong characteristics of portfolio evaluation.

Similarly, Caner (2010) aimed at exploring the views of students in portfolio assessment among EFL students doing a writing course at Anadolu University in Turkey. Using interviews and questionnaires, the study reported optimistic attitudes towards portfolio assessment. Although participants found the portfolio a supportive instrument, they also believed that it had loaded them with extra duties while compiling it. Consequently, slightly over half of the subjects preferred not to be evaluated by portfolios, and most of them preferred to be assessed by traditional paper and pencil tests. In response to this, instructors were advised to provide extra time for the students' portfolio preparation processes.

In the same line, Lam (2013) conducted a small-scale classroom investigation that involved 16 EFL pre-university students in Hong Kong, with the data collection mainly concerned with student perceptions about their experiences with two portfolio systems, either a working portfolio or a showcase portfolio. The findings indicated that students from the showcase portfolio group were less enthusiastic about the effectiveness of portfolio assessment and enquired whether it could promote autonomy in writing. The working portfolio group was more receptive to the experience, considering that a feedback-rich environment in the working portfolio system could facilitate writing improvement.

In the same vein, Assaggaf & Bamahra (2016) investigated the views of Yemeni computer science undergraduates on using a portfolio as a tool to enhance report writing. To achieve that purpose, three instruments were employed: group discussions, written reflections, and a questionnaire. The findings showed that portfolio was found helpful in several learning aspects, including elements of writing, feedback and report writing elements.

Similarly, Laraba & Boukhedimi (2021) aimed to investigate the perceptions of using portfolio in a writing course at a public university in Algeria. For data collection, they used a content analysis of mixed methods for students' written reflections. The findings of the study showed positive perceptions towards portfolios, particularly in relation to motivation towards further progress and self-reliance.

By the same token, Qayyum (2022) investigated the use of portfolios among remedial writing skills undergraduates in an academic institution in Pakistan. The main purpose was to

obtain the insights, benefits and challenges of these portfolios using a thematic analysis of data collected from multiple sources. The results showed positive attitudes towards portfolios. In particular, it has been found that those portfolios enhanced the writing abilities of students by helping them produce paragraphs with better organization, clearer sentences and allowed them to develop a focus on critical thinking.

In a similar vein, Punyapratheep (2022) conducted a study that aimed to both develop criteria for portfolio assessment and investigate students' perceptions toward the use of portfolio in assessing writing abilities of Thai EFL undergraduates. To this end, he used three techniques namely, a questionnaire, a semi-structured interview, and a portfolio self-assessment form. The findings revealed a positive attitude towards the use of portfolios, and that the participants showed significant improvements in their writing abilities. The study also concluded with four criteria of portfolio assessment in assessing English writing ability.

In the same line, Apeadido et al., (2024) looked at the attitudes and perceptions of portfolio assessment among both lecturers and students in a university in Ghana. A five-point Likert questionnaire was used to collect data for a descriptive analysis. The results showed some discrepancies in both lecturers' and students' attitudes, as lecturers showed higher perception of portfolio assessment. The study concluded by highlighting the importance of addressing the differences in perceptions between lecturers and students for promoting a shared practice of portfolio assessment.

These studies have informed us about the use of portfolios among undergraduates in different EFL contexts. We have seen that the findings of these studies were positive towards portfolio use. However, when searching for similar studies in the Saudi context, very few could be identified. An early study was conducted by Alabdelwahab (2002). He conducted a case study that aimed to explore the attitudes of students, teachers and administrators in using portfolio among Saudi intermediate school learners. For data collection, he utilized interviews and observation. The findings uncovered that most of the learners enjoyed using the self-assessment portfolio, and that both the teachers and administrators showed positive attitudes towards the use of this technique in teaching English at Saudi intermediate schools.

At the university level, only two studies were identified. Alhuwaydi (2017) investigated the perceptions of writing instructors and students at Qassim University in Saudi Arabia. Two primary instruments were employed: multiple semi-structured interviews and a questionnaire. The results showed positive perceptions of both instructors and students. Even though the instructors believed portfolio assessment provided them with more comprehensive picture of students' writing abilities, they also revealed some obstacles to portfolio implementation assessment, including university grading system, and large class sizes. The results showed also that the students felt more confident, as portfolios helped them exhibit their best writing in an attractive fashion.

Similarly, Alam & Aktar (2019) conducted a pilot experimental study to implement formative portfolio assessment at King Khalid University. Their methodology included a test for measuring students' performance and an interview for eliciting students' views on the use of portfolio. The participants were forty entry-level students at the university receiving their preparatory English language lessons. The findings showed that the portfolio was an effective tool for enhancing students' writing abilities.

In light of the review above, it is evident that the portfolio has been extensively researched among undergraduates in numerous EFL contexts; however, studies in the Saudi context remain

limited. This study, therefore, seeks to address this gap by focusing on the views of both students and writing instructors.

Methodology

Study Context and Participants

The study was conducted while teaching a writing course in the English department, College of Science & Arts, Albaha University. In this undergraduate program, students needed to pass three writing courses: Writing one, Writing two, and Essay Writing. This study was conducted in a course entitled "Writing two", which is presented in the second semester of the first academic year. According to the course specifications, "Writing two" was intended to "develop students' ability to write a well-connected expository, narrative, descriptive, persuasive, informative multi-paragraphs that have sound sufficient supporting details of various types (e.g., facts, examples, quotations, etc.) in addition to using correct grammar, spelling, punctuation, capitalization, indentation etc." (Albaha University, College of Arts & Sciences, English Department, 2018). The participants were of two types: students and writing instructors. The student participants were twenty-two male students who attended the course "Writing two" during the second semester of the year 2022. Data was also collected from four teaching staff members in the same department. They were PhD holders, who have been teaching writing courses for a couple of semesters in the same department. Even though the instructors were not involved in the implementation of the portfolio, their experience in the same learning environment is considered invaluable for adding depth and strength to the findings from the student participants.

Procedures

From the start and during the introductory session of the course (Writing two) students were introduced to the implementation of portfolio as a new assessment technique in the course. For that purpose, students were instructed right from the beginning to maintain a file in which all documents used in the course should be retained. The documents include different drafts of short written answers, assignments, reflections, and answer sheets. For every assignment given, students were asked to share it with their peers and get feedback from them. They were then asked to write another draft and place it in the file to show later to the instructor. All drafts with comments should be placed in the portfolio. The students were also informed that the portfolio would be assigned 10% of the course overall grade, that is the marks identified as "assignments" in their course specification.

Data Collection and Analysis

For data collection, four techniques were utilized: student self-reflections, group discussions, a short questionnaire and unstructured interviews. Here are more details about each one.

Individual reflections: from the beginning, student participants were asked to write their reflections on their experience with portfolios at least two times during the course. The students were encouraged to write about what they like or dislike about their experience with portfolios and whatever they felt relevant to share with others. They were told to do that in English or even in Arabic (mother tongue), so that they feel more comfortable when they write.

A questionnaire: a short questionnaire, adapted from Assaggaf & Bamahra (2016) with slight modification, was administered to the students. The questionnaire contained ten items that

focused on eliciting students' perceptions on portfolio focusing on specific areas, including writing abilities revision, feedback, grammar. The administration was held for the student participants in the fourteenth week of the semester; after ensuring students have had whole experience with portfolio implementation.

Group discussions: Along with the questionnaire, two group discussions were held with the students: one towards the middle of the course (week eight), and the other by the end of the course. The discussions were held inside the classroom with most students present. An open group discussion was selected as a tool to encourage students to express their views openly when they are together. The first discussion focused on their experience and the good and bad aspects of portfolios. In addition to that, the second group discussion was also dedicated to getting further explanations on the themes and responses derived from the questionnaire's statistics.

Unstructured interviews: these were conducted with the writing instructors after the course was completed. The interviews were held individually with four writing instructors in the same department. Although they were unstructured, the interviews focused on three aspects, namely, their view on portfolio implementation for their students, the possible advantages of employing such a technique in that particular context, and the expected challenges that might arise.

For analyzing qualitative data from the interviews and group discussions, the researcher listened to the recordings several times and the main themes were formulated. Data from the reflections were also compared to the initial themes obtained from the discussions. Three themes first emerged: general views, positive views and negative views. For every theme, at least one excerpt is highlighted to give proper exemplification in the presentation of the findings. Data from questionnaires were analyzed by counting responses in every item and then calculating the percentages accordingly. For a better presentation, data from the students and the instructors were made separate from each other in both the analysis and the presentation. The items used in the analysis were labelled as: (GD1) for group discussion one, (GD2) for group discussion two, (WI1) for writing instructor one, and so on.

Results

Students' Views

In response to the first research question concerning students' views on implementing portfolios in teaching writing, this section presents the findings derived from the questionnaire, individual reflections and group discussions.

Students' Views on Writing Skills and Development

The quantitative data obtained from the questionnaire is summarized in Table 1 below.

Table 1. *Students' views on writing skills*

No.	Questionnaire item	Disagree	Not sure	Agree
1	Implementing portfolios in my course helped me improve my writing abilities.	9%(2)	18%(4)	73%(16)
2	Portfolio helped me write better sentences.	9%(2)	23%(5)	68%(15)
3	Portfolio helped me write better paragraphs.	18%(4)	18%(4)	64%(14)
4	Portfolio helped me review my previous work.	5%(1)	9%(2)	86%(19)
5	Portfolio encouraged me to share my writing with others.	9%(2)	9%(2)	82%(18)
6	Portfolio helped me get more feedback from my instructor.	14% (3)	14%(3)	72%(16)
7	Portfolio helped me get more feedback from my classmates.	18%(4)	18%(4)	64%(14)
8	Portfolio helped me write grammatically correct sentences.	18%(4)	22%(5)	59%(13)

9	Portfolio helped me with writing techniques such as punctuation marks.	14%(3)	23%(5)	63%(14)

The first questionnaire item concerns students' views on the portfolio in general. As can be seen in Table 1 above, (73%) of the students believe that portfolio use was found helpful in developing writing abilities. In the discussion, students confirmed this finding by asserting the importance of portfolio in developing writing abilities. One student said, "I think it is good and even very good because you can review your work from time to time"(GD1).

The next two questionnaire items (2 & 3) are about writing sentences and paragraphs. As can be seen in the table above, students believe that the use of portfolio helped them improve their writing of sentences as well as paragraphs, with (68%) for sentences and (64%) for paragraphs. Item 4 shows a significant response concerning revision with (86%). The majority of students believe that portfolio could help them revise their writing in a better way. Qualitative data from group discussions also adds to this view. Here is an excerpt: "one thing I liked about this technique is that you are encouraged to look back at something you wrote some time ago and give it more review. I think this is good because you may find something to add, or you may change something to make it better"(GD1).

Items 5, 6 and 7 disclose also positive views towards feedback. As can be seen in Table 1 above, (82%) of the participants believe that portfolio encouraged them to show their written work to others to get feedback. The participants believe also that portfolio helped them get feedback from the instructor (72%) and from their classmates (64%). The difference in feedback between the instructor and classmates might be attributed to the fact that some students feel more trust in the instructor's feedback rather than their peers (Cui et al., 2022).

The final two questionnaire items (8 and 9) disclose positive views towards the influence of portfolio on writing grammatically correct sentences (59%). Data also shows that (63%) of participants view the clear influence of portfolio in developing writing techniques, particularly punctuation. Though the findings here seem pretty positive, there are higher negative percentages (22% and 23%) as compared to the other items. This may suggest that students found portfolio the least helpful in grammar and accuracy. In the discussion, students attributed that to the fact that they were not so good at grammar and punctuation. One response stated: "actually, we weren't so perfect in grammar and that's why we were not sure about our feedback." (GD2) This may be attributed to the level of students, as they were still in the second semester of their English program.

Students' Views on Portfolios' Benefits and Drawbacks

Analysis of the qualitative data from the reflections and group discussions revealed further patterns regarding the positive and negative aspects of portfolios.

Benefits: The major beneficial aspects in the students' views are:

1) Motivation: Students believed that portfolio helped motivate them to do more writing. Working on one's own writing encourages students to do more writing,

2) Better experience: Students believed that previously, writing was not given enough time or attention. The common practice was for a student to complete the required task quickly and submit it immediately to the teacher for feedback. However, students now felt they had more time to revise and discuss their written tasks, which made them more experienced writers.

3) Self-correction: Students believed that portfolio use enabled them to find their errors and correct them by themselves.

4) Peer feedback: Students believed that portfolio use encouraged them to share their written assignments with each other and give and receive feedback. This was not so common in their previous practice.

Drawbacks: the major obstacles of portfolio use in the students' views are:

1. Long and demanding process: For students, the portfolio was a novel and demanding experience, differing from the usual traditional writing assessment tools. That experience was deemed demanding as they needed to keep all the documents in one file throughout the semester, as well as undergo the process of revising many times.
2. Lack of adequate incentives: Students expressed disappointment, as they found portfolios demanding and the marks assigned were only limited. For them the marks assigned were not enough compared to the effort required.
3. Reluctant sharing: Students emphasized that the practice of sharing their assignments with one another was not seriously treated in the previous courses and thus they tried to avoid it during this course. The underlying reasons appeared to come from concerns about language errors, as students seemed to believe these errors could convey negative impressions of them.

Instructors' Views

The second research question is to find out the perceptions of writing instructors on the use of portfolio in teaching writing for Saudi EFL undergraduates. The analysis of data collected from instructors has been categorized into three: views on using portfolio in general, strengths, and weaknesses.

a) Using portfolio as a tool in teaching writing: Data showed that the perception of the portfolio among instructors is highly favorable. Instructors expressed undeniable views that portfolio use is an invaluable tool for improving students' writing abilities in this context. Here are two excerpts: "No doubt that it is one of the useful methods for improving our students' writing abilities. I guess we need it a lot."[WI3]. "We need it as a modern technique that can help eradicate old and traditional techniques that are sometimes used here"[WI2]. As we can see, the instructors have a favorable view towards implementing this tool to develop students' writing abilities. We can observe from the second excerpt that portfolio is seen as a new technique that is hoped to replace other traditional techniques.

b) Strengths of portfolio implementation: Data revealed numerous benefits of using portfolios, which are summarized below.

1) Developing autonomy: instructors feel that portfolio is useful for developing autonomy. Here is an excerpt "The practice here is that students wait for the corrections to be given by the teacher only. If we use such a method, they'll try that first on their own, that's definitely going to be great!" [WI2].
2) Cooperative learning: the implementation of portfolios creates an atmosphere where students cooperate with one another, which hopefully results in better outcomes. One instructor asserted: "When students work together on their written work, that means we provide good atmosphere for them to help each other, I mean they can collaborate with one another before they go to the teacher]"[WI3].
3) Confidence and Motivation: when students work together to review and revise their own work and with others, this will likely make them feel motivated and even more confident. One instructor stated: "Now students spend more time with their classmates and with the

teacher, more than in the traditional way. This will absolutely make them more motivated and of course will develop confidence in what they do"[WI4].

4) Continuous revision: instructors believe that portfolio use will provide better opportunities for students to review their written assignments on their own as well as with the others over a period of time, which is longer than the conventional practice. One instructor stated: "Good that it allows a learner to review over a period of time"[WI4].
5) Enjoyable learning experience: instructors believe that portfolios would make writing learning more fun by breaking the dullness of a normal writing class. An instructor confirmed: "When students work together in class, I think this creates better learning context where learning becomes more enjoyable". [WI3].

c) Weaknesses of portfolio implementation: Even though instructors expressed optimistic attitudes towards the use of portfolios, they have also revealed some potential obstacles based on their knowledge and experience with this context. Here are the most significant ones.

1) Self-review reluctance: students are usually reluctant to continually revise their own written work. Here is an excerpt from the interviews: "From my experience with my students they tend to get bored of reviewing things several times and they usually wait for the corrections to be given by the instructor" [WI1].

The instructor here refers to the reluctant treatment of students with portfolio being a demanding task. Like in most traditional learning contexts, learners avoid doing more revision on their own and tend to prefer feedback to be given by teachers. This behavior, however, needs to be changed by practice. As mentioned above, one advantage of learning is autonomy which has numerous benefits to learners and learning as well.

2) Peer review avoided: students in this learning context try to avoid sharing their writing with other colleagues, as also found out by the students. Here's an excerpt from the interviews: "peer reviewing is not a common practice here, and leaners find it embarrassing to share their work or discuss it with their peers" [WI2].

3) Traditional practices: instructors emphasized that the learning in the department, particularly on the part of the students, is still generally traditional. Here is an excerpt: "students basically give focus to the grammatical and structural errors over other issues in writing" [WI1].

4) Grading system: the current course's grading distribution does not give even space for new techniques such as portfolio assessment, as only a few marks are assigned to assignments. Here is an excerpt from the data: "we have only 10 to 15% marks that can be used in evaluating such a technique. Other marks are for the other tests and exams" [WI3]. Such grading will likely hinder the implementation of this technique or any other non-traditional technique (Alhuwaydi, 2017; Alsadaawi, 2010).

Discussion

The main objective of this study is to elicit students' and instructors' views on portfolio implementation in teaching writing among Saudi EFL undergraduate students. The main finding here shows positive attitudes of both students and instructors on portfolio use. Data from both instructors and students confirmed the favorable use of portfolios in improving writing teaching and learning. This finding is in line with previous research (e.g., Alabdelwahab, 2002; Alam & Aktar, 2019; Apeadido et al., 2024; Qayyum, 2022). This finding, then, confirms the common consensus among learners and instructors from varied EFL contexts.

The study revealed also that the portfolio is seen of great value for improving writing skills and abilities. Findings from students showed that portfolio was seen helpful in many aspects, including enhancing writing experience, motivation, self-correction and feedback. On the other hand, instructors believed that portfolio was considered helpful in many respects such as confidence, motivation, cooperation and autonomy. This finding is in accordance with the findings by Alam & Aktar (2019), who reported that portfolio was seen useful in developing motivation, confidence, and autonomy among leaners at King Khaled University in Saudi Arabia. Furthermore, the findings are parallel with what was found by Laraba & Boukhedimi (2021), who reported influence of portfolio among EFL undergraduates in a university in Algeria, particularly in motivation and autonomy.

The findings from the questionnaire have revealed that students found portfolio helpful in developing writing skills such as writing sentences and paragraphs, grammar and feedback. This finding is in accordance with Qayyum (2022), who studied portfolio use among Pakistani students at a university in Pakistan. She found that portfolios helped students improve writing simple and straightforward sentences. In addition, the findings here are in line with Assaggaf & Bamahra (2016), who reported that portfolio was believed to be helpful in writing better sentences and paragraphs by computer undergraduates in Yemen.

The present study also revealed a few obstacles that were believed to hinder portfolio implementation in this Saudi context. A major obstacle that has been stated was the socio-cultural aspects that were reported in this context. Both students and instructors believed that student-student contact when dealing with written tasks was not smooth enough, as some students seemed reluctant to share their writing with their peers. Furthermore, some traditional practices were also reported, particularly in dealing with feedback, where the focus was geared towards form and structure. It was also reported that both students and instructors found the use of portfolios to load students with extra duties. Both students and instructors agreed that portfolios formed a burden on the students, particularly when they are not rewarded accordingly. This agrees with the findings of Caner (2010), who studied the portfolio in a Turkish university and found that most students preferred to be assessed by traditional paper and pencil tests, as they found portfolios more demanding.

The study reported also that the grading system, where only limited marks were assigned to the portfolio, was believed to hinder the use of this form of assessment. Instructors expressed their dissatisfaction with this and felt that it would hold back any innovation in implementing innovative assessment techniques. Students also reported that the effort they escorted in the portfolio was not rewarded sufficiently by the instructors. This finding is in line with Alhuwaydi (2017), who reported the university grading system as an obstacle in his study at Qassim University.

Conclusion and Recommendation

With respect to the underpinning findings above, a few recommendations can be suggested. More flexibility in the grading systems at Saudi universities needs to be considered to facilitate effective use of new assessment techniques such as portfolios. Indeed, we need to encourage the use of these assessment techniques in both universities and schools to achieve a broader impact. It is hoped that training students at early stages in schools will grant them greater practice and experience prior to university enrolment. This, in turn, should significantly foster advances in the teaching and learning of writing and other key learning abilities. Furthermore, given the socio-

cultural aspects highlighted in the study's findings, e-portfolio could effectively be utilized in some institutions as alternative to paper portfolios. This recommendation is supported by Alshawi & Alshumaimeri (2017) and Nasseif (2021), who both of whom argue that e-portfolios are more appropriate in numerous Saudi undergraduate writing classes.

For further research, it would be highly valuable to suggest studying the process of portfolio implementation. It would be useful to gain a full understanding of how portfolios are practically implemented. In addition, more research could be carried out comparing the use of portfolio and e-portfolio across different learning levels and stages to identify the best choices for each.

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: AI and AI-assisted technologies were not used.

About the Author: Hussein Taha Assaggaf is currently an assistant professor of applied linguistics at the Department of Foreign Languages, College of Arts & Humanities, Al-Baha University, Kingdom of Saudi Arabia. His research interests include, EFL writing, English for academic purposes, pragmatics and discourse analysis. ORCID ID: https://orcid.org/0000-0002-3991-5430

References

Alabdelwahab, S. Q. (2002). *Portfolio assessment: A qualitative investigation of portfolio self-assessment practices in an intermediate EFL classroom, Saudi Arabia* (Doctoral dissertation, The Ohio State University).

Al-Abdulkareem, S. A. (2004*). Investigating science teachers' beliefs about science teaching: Struggles in implementing science education reform in Saudi Arabia* (Doctoral dissertation, West Virginia University).

Alam, M. J., & Aktar, T. (2019). Assessment Challenges & Impact of Formative Portfolio Assessment (FPA) on EFL Learners' Writing Performance: A Case Study on the Preparatory English Language Course. *English Language Teaching, 12*(7), 161-172. https://doi.org/10.5539/elt.v12n7p161

Albaha University, College of Arts & Sciences, English Department. (2018). *Course Specifications*.

Al-Hidabi, D. A., Kabilan, M. K., & Shakfa, M. D. (2019). Using E-Portfolio to Enhance Student's Attitudes Towards Learning English at University College of Applied Sciences. *IOSR Journal of Research & Method in Education, 9*(5), 1-9. https://dx.doi.org/10.9790/0837-2410050107

Alhuwaydi, A. A. (2017). *Portfolio as an EFL writing assessment in a Saudi university context*. (Doctoral dissertation, The University of Memphis).

Al-Khuweileh, A. and Al Shoumali, A. (2000). Writing errors: a study of the writing ability of Arab learners of academic English and Arabic at university. *Language, Culture and Curriculum*: *13*(2), 174-183.

Alkodimi, K. A., & Al-Ahdal, A.A.M.H. (2021). Strategies of Teaching Writing at Saudi Tertiary

Level Institutions: Reality and Expectations. *Arab World English Journal, 12* (2) 399- 413. https://dx.doi.org/10.24093/awej/vol12no2.27

Alotabi, K. A. (2014). Student assessment strategies in Saudi Arabia: a case study of pre and post classroom practices. *Literacy Information and Computer Education Journal (LICEJ), 3*(1), 1786-1763.

Alsadaawi, A., (2010). Saudi national assessment of educational progress (SNAEP), *International Journal of Education Policy and Leadership, 5*(11), 1-14. https://doi.org/10.22230/ijepl.2010v5n11a202

Alshawi, A. T., & Alshumaimeri, Y. A. (2017). Teacher electronic portfolio and its relation to EFL student teacher performance and attitude. *International Journal of Education and Literacy Studies*, *5*(1), 42-54. http://dx.doi.org/10.7575/aiac.ijels.v.5n.1p.42

Apeadido, S., Dzomeku, D., & Nanor, E. (2024). Navigating the Portfolio Landscape: Examining Lecturers' and Students' Perspectives on Portfolio Assessment Practices. *Asian Journal of Education and Social Studies, 50*(5), 563-582. http://dx.doi.org/10.9734/AJESS/2024/v50i51386

Arslan, R. Ş., & Gümüş, S. N. (2020). The effect of portfolio-keeping on EFL young learners' writing achievement and their motivation towards writing. *Journal of Narrative and Language Studies*, *8*(14), 130-150.

Assaggaf, H. T. & Bamahra, Y. S. (2016). The effects of portfolio use in teaching report writing: EFLstudents' perspective. *International Journal of Applied Linguistics and English Literature*, *5*(3), 26-34. http://dx.doi.org/10.7575/aiac.ijalel.v.5n.3p.26

Assaggaf, H. T., Stapa, S. H., & Mustafa, J. (2012). Investigating teaching and learning practices of second language ESP writing at tertiary level in Yemen. *Arab World English Journal*, *3*(2), 48-75.

Au, W. (2011). Teaching under the new Taylorism: High-stakes testing and the standardization of the 21st century curriculum. *Journal of curriculum studies, 43*(1), 25-45. http://dx.doi.org/10.1080/00220272.2010.521261

Boumediene, H., Berrahal, F. K., & Hardji, M. B. (2018). Using portfolio assessment to enhance reflective writing in EFL classrooms. *Literacy Information and Computer Education Journal(LICEJ)*, *9*(3), 2994-3001.

Brown, D. J. (2005). *Testing in language programs: A comprehensive guide to English language assessment*. New York: McGraw-Hill Companies, Inc. https://doi.org/10.1191/0265532205lt306xx

Burke, K., Fogarty, R., & Belgrad, S. (1994). *The mindful school: The portfolio connection.* Palatine, IL: Skylight.

Caner, M. (2010). Students views on using portfolio assessment in EFL writing courses. *Anadolu University Journal of Social Sciences*, *10*(1), 223–236.

Cui, Y., Schunn, C. D., & Gai, X. (2022). Peer feedback and teacher feedback: a comparative study of revision effectiveness in writing instruction for EFL learners. *Higher Education Research & Development*, *41*(6), 1838-1854. https://doi.org/10.1080/07294360.2021.1969541

Danielson, C., & Abrutyn, L. (1997). *An Introduction to Using Portfolios in the Classroom.* Association for Supervision and Curriculum Development: USA.

Ezza, E. (2010). Arab EFL learners' writing dilemma at tertiary level. *English Language Teaching, 3*(4), 33-39.

Fenwick, T.J., & Parsons, J. (1999). A note on using portfolios to assess learning. *Canadian Social Studies*, 33(3), 90–92.

Fox, J. (2017). Using portfolios for assessment/alternative assessment. In E. Shohamy, I. G. Or, & S. May (Eds.), *Language testing and assessment, Encyclopedia of language and education* (3rd ed., pp. 135-148). Dordrecht, The Netherland: Springer. https://doi.org/10.1007/978-3-319-02261-1_9

Hamp – Lyons, L., & W. Condon. (2000). *Assessing the Portfolio: Principles for Practice, Theory, and Research*. Cresskill, NJ: Hampton Press.

Guo, H., & Li, D. (2024). The Effect of Electronic Portfolio-Based Writing Instruction on the Second Language Writing Performance and Writing Self-Efficacy of EFL Learners: A Mixed Methods Study. *SAGE Open, 14*(2), 1-18. https://dx.doi.org/10.1177/21582440241257697

Kalra, R., Sundrarajun, C., & Komintarachat, H. (2017). Using Portfolio as an Alternative Assessment Tool to Enhance Thai EFL Students' Writing skill. *Arab World English Journal*, *8*(4), 292-302. https://dx.doi.org/10.24093/awej/vol8no4.20

Khasawneh, M. A. S., Aladini, A., Assi, S. A., & Ajanil, B. (2025). Portfolio assessment in AI-enhanced learning environments: a pathway to emotion regulation, mindfulness, and language learning attitudes. *Language Testing in Asia, 15*(1), 5-15. https://doi.org/10.1186/s40468-025-00345-0

Klenowski, V. (2002). *Developing portfolios for learning and assessment: Processes and principles*. Routledge.

Kohn, A. (2000). *The case against standardized testing: Raising the scores, ruining the schools*. Portsmouth, NH: Heinemann.

Lam, R. (2013). Two portfolio systems: EFL students' perceptions of writing ability, text improvement, and feedback. *Assessing Writing*, *18*(2), 132-153.

Lam, R. (2018). *Portfolio assessment for the teaching and learning of writing*. Singapore: Springer.

Laraba, W., & Boukhedimi, Y. (2021). EFL Students' Perceptions of the Working Portfolio in Classroom Writing Assessment. *Alpha*, *8*(2), 477-490.

Ma'arif, A. S., Abdullah, F., Fatimah, A. S., & Hidayati, A. N. (2021). Portfolio-based assessment in English language learning: Highlighting the students' perceptions. *J-SHMIC: Journal of English for Academic*, *8*(1), 1-11.

Mabry, L. (1999). *Portfolios Plus: A Critical Guide to Alternative Assessment*. Corwin Press.

Martinez-Lirola, M., & Rubio, F. (2009). Students' beliefs about portfolio evaluation and its influence on their learning outcomes to develop EFL in a Spanish context. *IJES, International Journal of English Studies*, *9*(1), 91-111.

Moya, S. S., & O'Malley, J. M. (1994). A portfolio assessment model for ESL. *The Journal of Educational Issues of Language Minority Students, 13*(1), 13-36.

Muho, A., & Leka, K. (2021). Students' perceptions of portfolio as a motivating factor in learning English as a foreign language. *Journal of Educational and Social Research*, *11*(6), 47-56. https://doi.org/10.36941/JESR-2021-0127

Nasseif, H. (2021). Exploring E-Portfolio as a new technology tool in Saudi Arabian higher education: A case study. *Studies in Technology Enhanced Learning*, *1*(2), 479-500. *https://doi.org/10.21428/8c225f6e.0e590f93*

Nunes, D. (2004) Portfolios in the EFL classroom: disclosing an informed practice. *ELT Journal*,

58(4), 327-335.

Pitri, A. W. (2021). The Correlation between Portfolio Assessment and Students' Motivation in Learning English. *IOSR Journal of Research & Method in Education*, *11*(4), 56-59.

Punyapratheep, K. (2022). *The use of portfolio in assessing English writing ability of Thai undergraduate students*. (Doctoral dissertation, Chulalongkorn University). https://doi.org/10.58837/CHULA.THE.2022.144

Qayyum, R. (2022). Using Portfolios to Formatively Assess the Writing Skills of Undergraduate ESL/EFL Students. In: Ali Raza, N., Coombe, C. (eds) *English Language Teaching in Pakistan. English Language Teaching: Theory, Research and Pedagogy*. Springer, Singapore. https://doi.org/10.1007/978-981-16-7826-4_15

Van der Schaaf, M., Baartman, L., & Prins, F. (2012). Exploring the role of assessment criteria during teachers' collaborative judgement processes of students' portfolios. *Assessment & Evaluation in Higher Education*, *37*(7), 847-860.

Zubizarreta, J. (2009). *The learning portfolio: Reflective practice for improving student learning*. John Wiley & Sons.

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.60-88
DOI: https://dx.doi.org/10.24093/awej/vol16no3.4

Research Trends of Digital Game-based Learning for Technical Vocabulary Acquisition: A Systematic Decennial Literature Review (2014-2024)

Feng Tian
Centre for Instructional Technology & Multimedia, Universiti Sains Malaysia, Malaysia;
School of Foreign Languages and Cultures, Ningxia University, China
Corresponding Author: tianfeng@student.usm.my

Irfan Naufal Umar
Centre for Instructional Technology & Multimedia, Universiti Sains Malaysia, Malaysia
Email: irfan@usm.my

Received: 06/02/2025 Accepted: 08/03/2025 Published:09/20/2025

Abstract
The approach of digital game-based learning has gained prominence as an effective strategy for technical vocabulary acquisition in foreign language education. But its acquisition presents persistent challenges due to its specificity, contextual dependence, and limited exposure in general language input. In response, this systematic review investigates how digital game-based learning has been utilized to support technical vocabulary acquisition in foreign language education. The study explores trends in this field by examining the evolution of research from 2010 to 2024, the types of digital games employed, the technical domains targeted, the learner variables and outcomes investigated, and the gaps and emerging directions in existing literature. Following the PRISMA framework, 17 empirical studies were selected and analyzed. Findings reveal a marked increase in research since 2020, with most studies reporting positive effects on vocabulary performance, learner motivation, and attitudes. The reviewed literature highlights the effectiveness of game-based approaches, particularly in higher education contexts and English for specific purposes. However, notable limitations persist, including the dominance of English-language settings, a narrow focus on certain technical fields, small sample sizes, and a lack of longitudinal designs. This review underscores the pedagogical potential of integrating game-based learning into technical vocabulary instruction and calls for more theory-driven, diversified, and context-sensitive research. Future studies are encouraged to explore how game design features interact with cognitive and motivational learner variables and to expand the application of such approaches across varied languages, disciplines, and educational settings.

Keywords: Digital game-based learning, foreign language education, foreign language for specific purposes, systematic review, technical vocabulary

How to cite: Tian, F., & Umar, I. N. (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://dx.doi.org/10.24093/awej/vol16no3.4

Introduction

Technical vocabulary refers to the lexical terms that are predominantly used in professional texts or specific disciplinary fields, while remaining rare or entirely absent in general contexts (Chung & Nation, 2003). These terms are essential in Foreign Language for Specific Purposes (FLSP) contexts, where learners must comprehend and produce professional materials (Le & Miller, 2023). However, acquiring such vocabulary is challenging due to its low frequency, high specificity, and limited contextual exposure (Ardasheva & Tretter, 2017). To address this issue, Digital Game-Based Learning (DGBL) has emerged as a promising instructional approach. Serious games, games designed explicitly for educational purposes, offer immersive, interactive, and repeatable learning environments (Plass et al., 2020). Unlike commercial games that are adapted for learning, serious games are purposefully aligned with instructional goals and assessment criteria (Soyoof et al., 2024). Studies have shown that DGBL can increase learners' motivation, create rich contextual input, and support long-term vocabulary retention (e.g., Abdullah Alhebshi & Gamlo, 2022; Ali & Ghazali, 2016; Bendebbache & Azeroual, 2022). These effects can be interpreted through Self-Determination Theory, which emphasizes the role of autonomy and competence in sustaining engagement (Ryan et al., 2006). Additionally, the multimodal nature of games, including visual, auditory and textual input, can reduce extraneous cognitive load and promote dual coding, as proposed by Cognitive Load Theory (Frumuselu, 2018; Sweller, 1994).

DGBL has been successfully applied across multiple disciplines and language skills, including reading (Babaie, 2022), listening (Moniza Ray & Ajit, 2024), speaking (Pratiwi et al., 2024), and grammar (Castillo-Cuesta, 2020). Vocabulary learning, in particular, has been a common focus due to its incremental nature and need for repeated exposure (Schmitt & Schmitt, 2020). However, most studies focus on general vocabulary. Despite the growing importance of technical vocabulary in FLSP, few studies have systematically examined how DGBL can support its acquisition (Alharbi, 2023; Gómez-Ortiz et al., 2023; Shi et al., 2024).

To clarify how Digital Game-Based Vocabulary Learning (DGBVL) fits within the broader scholarly landscape, Figure One presents a conceptual model adapted from Hung et al. (2018), showing the relationship between DGBL, Digital Game-Based Language Learning (DGBLL), and general Game-Based Learning (GBL). This structure informs the scope and orientation of this review.

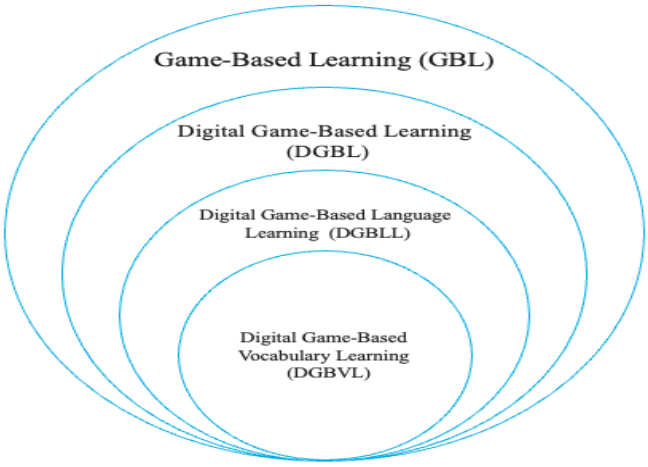


Figure 1. DGBVL and its relations to other research contexts (adapted from p. 92)

The aim of this study is to conduct a systematic review of empirical research on DGBL for technical vocabulary acquisition in foreign language contexts. The literature review synthesizes existing evidence, identifies dominant patterns and underexplored areas, and informs future research as well as practice. This review is significant in addressing the fragmented understanding of how DGBL supports technical vocabulary learning and in guiding the development of theory-informed, context-sensitive interventions in FLSP education.

Specifically, this study is guided by the following research questions:

(1) What are the trends in the application of DGBL for technical vocabulary acquisition from 2010 to 2024?
(2) What types of digital games have been used for technical vocabulary acquisition?
(3) In which technical fields has DGBL been applied to assist students in acquiring technical vocabulary?
(4) What variables have been considered in DGBL research on technical vocabulary acquisition, and what are the associated research outcomes?
(5) What are the research gaps and emerging trends in this field?

Literature review

Technical vocabulary

The definition of technical vocabulary remains imprecise in many existing studies, with the terms "technical vocabulary", "specialized vocabulary", and "academic vocabulary" often being used interchangeably. For instance, in their empirical studies, some researchers (Jia et al., 2023; Yi-ju et al., 2020) have referred to it as academic vocabulary in their empirical studies, while some others have categorized it as specialized vocabulary (Ali & Muhammad, 2019), professional vocabulary (Mustaeva et al., 2022) and discipline-specific vocabulary (Kohnke & Ting, 2021). Despite terminological variations, researchers generally agree that technical vocabulary occupies a substantial proportion of professional discourse, with coverage rates often exceeding 30%, demonstrating its central role in comprehending specialized content (e.g., Drayton & Coxhead, 2023; Lei & Liu, 2016).

Nation (2001) provided a more systematic classification of vocabulary based on its usage, dividing it into four categories: high-frequency vocabulary, low-frequency vocabulary, academic vocabulary, and technical vocabulary. High-frequency vocabulary refers to words that appear regularly in various texts and daily conversations. These are typically taught in English classes of general schools. Academic vocabulary, on the other hand, consists of words widely used across different disciplines but with limited occurrences in non-academic contexts (Beck et al., 2013; Othman, 2024). Technical vocabulary, in contrast, is tightly linked to specific disciplines and conveys specialized knowledge, as it often includes both subject-specific uses of general words and entirely field-exclusive terms (Chung & Nation, 2004; Liu & Lei, 2019). For example, the term "pilot" from high-frequency word lists generally refers to a person who operates the flying controls of an aircraft. However, in an academic context, it denotes a preliminary test or study conducted before a larger-scale study.

Unlike general or academic vocabulary, technical terms usually require deeper contextualization for proper understanding and use. They tend to lack synonymic alternatives, present high domain dependence, and require precise comprehension to avoid miscommunication. This creates higher learning demands for foreign language learners in technical disciplines. Several researchers argue that foreign language learners need to understand at least 95% of the words in a

text to achieve adequate comprehension and 98% for full comprehension (Coxhead & Demecheleer, 2018; Hsueh-Chao & Nation, 2000). A study by Chung and Nation (2003) revealed that technical terms account for 31.2% of the vocabulary in anatomy texts. The research of Lei and Liu (2016) found that technical vocabulary in medical profession (including high-frequency words with specialized medical meanings and low-frequency technical medical terms) constitutes 31.75% of tokens in a medical research article corpus and 30.44% in a medical textbook corpus. Similarly, Drayton and Coxhead (2023) reviewed several related studies and concluded that technical terms are used less frequently in spoken language than in written language with written language coverage around also 30%. The coverage rates are summarized in the Table One:

Table 1. *Coverage of Technical vocabulary in spoken and written corpora*

Technical Vocabulary List	Spoken Corpus	Written Corpus
Carpentry	10.69%	38.35%
Plumbing	11.59%	34.48%
Fabrication	9.18%	30.47%
Automotive	12.75%	37.44%

Note 1. Adopted from (Drayton & Coxhead, 2023, p. 54)

It is evident through these previous studies that technical articles typically contain around 30% technical terms, making these terms essential for understanding the content (Woodward-Kron, 2008). Therefore, general foreign language proficiency is insufficient for effective technical communication (Drayton & Coxhead, 2023).

The importance of technical vocabulary in language for specific purposes

Many researchers have consistently emphasized that vocabulary constitutes one of the most critical components of language acquisition since the lexical information plays a pivotal role in guiding learners toward accurate syntactic structures (e.g., Calvo-Ferrer, 2021; Teng et al., 2024). As some other researchers have asserted (e.g., Hestiana & Anita, 2022; Javis, 1998), without vocabulary, there is no meaning. Furthermore, research has shown a positive correlation between vocabulary size and learners' proficiency in oral and written communication, as well as their listening and reading comprehension (Aravind & Rajasekaran, 2021; Bulut & Yıldız, 2024), since vocabulary errors are often a source of misunderstandings which can disrupt the fluency of communication (Lin & Qian, 2019; Simonnet et al., 2025). Rivers (1983) argued that in the context of basic communicative competence, the acquisition of meaning relies more on vocabulary than on grammar. Empirical evidence suggests that learners are more prone to making vocabulary errors than grammatical ones (Song, 2018; Teng et al., 2024). All other factors being equal, learners with larger vocabularies demonstrate superior proficiency in a wider range of language skills compared to those with smaller vocabulary (Stewart et al., 2021). Moreover, encountering unfamiliar and novel vocabulary items poses a significant challenge in foreign language learning (Banu et al., 2025; Boyinbode, 2018), particularly when reading scientific or technical texts.

Binon and Verlinde (2003) emphasize that, regardless of the method or approach we adopt or prefer, vocabulary remains a key component that permeates the entire field, which is especially unique and indispensable in the context of FLSP. Although specialized vocabulary may not constitute the core content of FLSP, understanding such vocabulary is crucial, as it directly influences students' comprehension and production of academic texts. Technical vocabulary extends beyond being "simple words" as it embodies concepts and ideas unique to specific

disciplines (Barcroft, 2018). It also serves as a linguistic tool essential for articulating key practices within these fields (Chung & Nation, 2003).

The challenges in acquiring technical vocabulary

The acquisition of technical vocabulary poses a significant challenge for students across all academic subjects (Shi et al., 2024; Wright & Peltier, 2016). One of the primary reasons is that most language learners perceive learning new words as a monotonous and passive activity that requires deliberate attention to subtle nuances of meaning, which diminishes their interest in and the effectiveness of foreign language acquisition (Wu, 2018). This challenge is particularly pronounced when learners engage with scientific texts, where they often struggle with word recognition and reading fluency (Kaldenberg et al., 2015), since these texts require them to simultaneously process complex scientific concepts and unfamiliar language. These terms are specific to the subject and are characterized by a large number of abstract concepts as well as domain-specific meaning that may differ significantly from their usage in daily contexts (Carrier, 2013; Kamaşak et al., 2021). This cognitive overload can hinder students' ability to fully comprehend both the scientific content and language nuances (Castellano-Risco et al., 2020), thereby making the acquisition of technical vocabulary increasingly difficult and tedious. The second challenge lies in the limited teaching time in class with the result of insufficient opportunities for repetition, which is an essential factor for effective vocabulary retention. Without regular review and repeated practice, newly acquired vocabulary tends to gradually fade from memory (Zou et al., 2021). The third challenge in acquiring technical vocabulary for learners is the rapid emergence of new terms driven by advancements in various fields and technologies, which necessitates their ongoing efforts to learn and adapt to newly introduced terms (Liu & Lei, 2019).

From the perspective of language teacher, teaching technical vocabulary has also presented considerable challenges. First, many teachers of foreign languages usually lack a thorough comprehension of the specialized disciplines, which makes it difficult for them to provide relevant contexts that help learners understand the meanings and usage of technical terms (Shi et al., 2024). Second, technical vocabulary in professional fields can be highly specialized and complex, making it challenging for teachers to simplify and explain these concepts effectively to learners who are nor familiar with the subject matter (Miroshnychenko, 2024). Third, teachers often face the challenge of identifying appropriate resources to support the effective teaching of these terms, such as specialized textbooks, multimedia materials, or industry-specific content (Miroshnychenko, 2024). This issue is particularly critical given the rapid development of specialized fields, which makes it difficult to access resources that reflect the latest developments and trends. Finally, the limited time available in the curriculum presents another major obstacle for language educators (Huei et al., 2021). Balancing the teaching of technical vocabulary with other language skills, such as grammar, speaking, listening, and writing, within a constrained timeframe is a formidable task. Therefore, finding engaging and effective methods for teaching technical vocabulary has become a pressing necessity. These challenges, both from the perspective of learners and teachers, constitute the primary obstacles to the acquisition of technical vocabulary.

Trends of DGBL in vocabulary acquisition

As a crucial component of second language acquisition, vocabulary learning has been significantly influenced by the emergence of new technologies. Technological activities can

enhance the motivation of second language learners, offering increased exposure to the target language through oral and multimedia formats, as well as providing learners with more opportunities to engage with the target language using various technological tools (Al-Rashidi, 2025; Hao et al., 2021). Compared to traditional methods, the use of computer-assisted strategies, such as digital games, often has a positive impact on second language vocabulary acquisition. To enhance effective vocabulary acquisition through computer-assisted methods, educational researchers have employed a strategy known as DGBVL, which involves learning vocabulary through serious games (Chen et al., 2018; Othman, 2024). The DGBVL approach integrates gamified elements, such as competitive rules, rewards, feedback, and scoring, into vocabulary learning. This approach aims to improve both short-term and long-term vocabulary acquisition outcomes (Buil et al., 2018), thereby increasing motivation, academic performance, and engagement, while reducing anxiety, fostering satisfaction and autonomy, as well as encouraging learner interaction (Zou et al., 2021). Chang and Chang (2023) reviewed experimental studies on DGBL in foreign language vocabulary acquisition conducted between 2008 and 2020. As illustrated in Figure Two, the annual number of publications steadily increased during this period, with an average annual growth of 23.83%, and this is consistent with the finds of Zou et al. (2021). Based on a follow-up search in the Scopus database using the same searching keywords as those employed by Chang and Chang (2023), it revealed that the number of publications in 2021, 2022, 2023 and 2024 reached 28, 20, 33 and 34, respectively. This indicates a slight but consistent growth compared to the year before 2020.

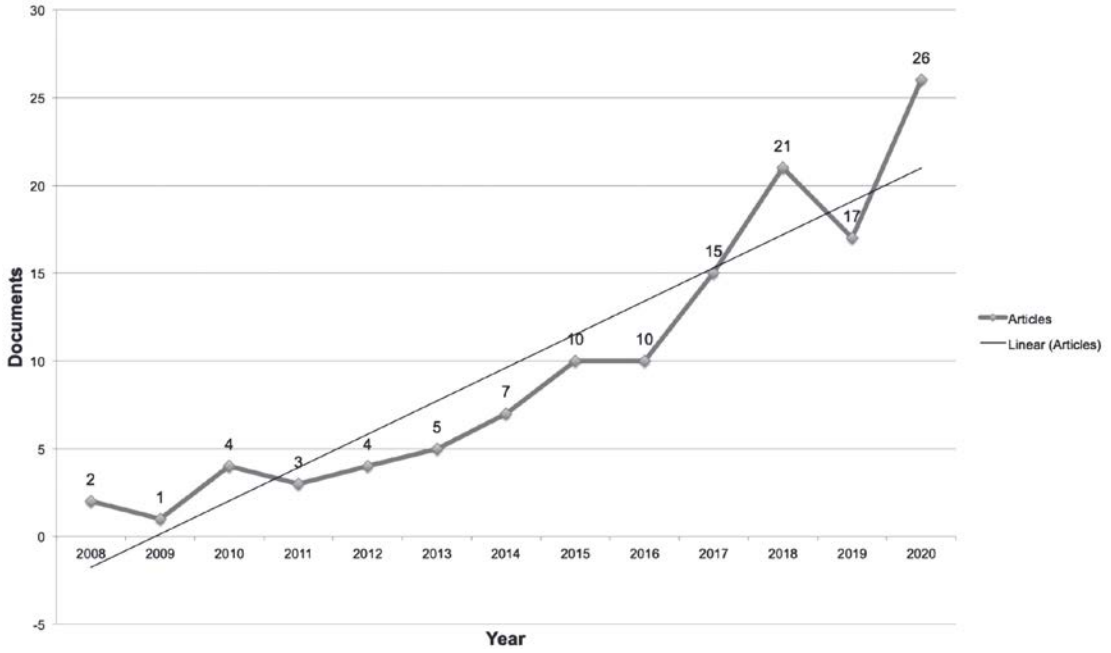


Figure 2. Annual scientific publications of DGBL in vocabulary acquisition (Chang & Chang, 2023, p. 11)

The study of Alfehaid (2019) on online vocabulary games revealed that 97% of teachers strongly agreed on the benefits of online games for vocabulary learning. According to a psychological analysis based on semi-structured interviews conducted by Fu et al. (2021) with 55 Chinese university students, 62% of participants reported that the DGBVL approach provided an enjoyable learning experience, enhanced their sense of autonomy, and further encouraged consistent and regular vocabulary learning behaviors. Additionally, the positive feedback

mechanisms within games minimized the negative impact of making mistakes, enabling learners to approach errors in a more constructive manner within a virtual environment, as well as allowing them to explore diverse learning paths and problem-solving strategies without experiencing the anxiety often associated with mistakes in traditional classroom settings (Vnucko & Klimova, 2023). Such learning activities, by fostering positive emotions in learners, can subsequently improve memory and support the retention of vocabulary. In a comprehensive review, Rasti-Behbahani (2021) identified seven recurring features that contribute to the effectiveness of DGBVL: high motivation, repeated exposure, timely feedback, immersive virtual environments, dual coding through multimodal input, interactivity, and learner control. These features align well with the cognitive and emotional needs of learners, particularly in vocabulary learning contexts that demand sustained engagement and contextualized practice.

Furthermore, numerous empirical studies have demonstrated the effectiveness and benefits of games in supporting vocabulary acquisition. In an empirical study involving 54 Spanish students, Calvo-Ferrer and Belda-Medina (2021) demonstrated that the intentional input provided by DGBVL methods significantly enhances vocabulary retention and facilitates incidental vocabulary acquisition. The researchers also highlighted the critical role of vocabulary repetition and contextual application within digital learning environments as key factors contributing to these outcomes. Similarly, Chen and Hsu (2020) conducted a study involving 66 non-English major students at a university in Taiwan, revealing that serious games are an effective tool for acquiring vocabulary and academic content. Their findings also demonstrated that the effectiveness of vocabulary learning is primarily attributed to repeated exposure to target vocabulary and the immersive nature of the game environment. Roohani and Vincheh (2023) investigated the effectiveness of various teaching strategies for phrasal verbs, including digital games, social media platforms (e.g., Telegram), and traditional classroom settings with 150 English learners from a language institute, demonstrating that all three methods enhanced learners' vocabulary acquisition, with the GBL approach having the most significant impact. Karimi and Nasouri (2024) designed a digital text-based game tailored for students at a private language institute in Iran. Their experimental study demonstrated that this innovative approach not only enhanced learner engagement and participation but also promoted a deeper state of flow, a psychological condition characterized by focused immersion and enjoyment in the task. Chowdhury et al. (2024) conducted an empirical investigation involving 13 Bangladeshi elementary students who utilized the Scratch programming platform to create their own games as a means of facilitating incidental vocabulary acquisition. The results emphasized the potential of DGBL environments to foster a low-anxiety, enjoyable learning atmosphere, which in turn enhanced leaners' intrinsic motivation and supported meaningful vocabulary development. The studies highlighted here represent only a small selection of the existing body of research. A comprehensive review of the current literature reveals that, based on various types of games, this approach exerts significant and multifaceted impacts on learners' subjective psychological states, such as engagement, motivation, and flow, as well as on objective dimensions as the academic performance. These effects have been consistently observed among students across diverse age groups, nationalities, and academic disciplines, underscoring the versatility and broad applicability of this instructional approach.

While innovative technologies have been widely applied in the teaching and learning of general vocabulary, technical vocabulary instruction still relies heavily on traditional pedagogical techniques. For example, Gaffas (2024) investigated the use of vocabulary notebooks for learning medical terminology, indicating that students who utilized vocabulary notebooks showed

significant advantages in long-term vocabulary retention compared to those who did not receive intervention. Some empirical studies (e.g., Xodabande et al., 2024; Yüksel et al., 2020) demonstrated that flashcards are more effective than traditional vocabulary lists in enhancing the flexibility of learning technical vocabulary anytime and anywhere, while significantly improving students' learning outcomes. Del Moral et al. (2024) incorporated group role-playing activities into traditional health science classrooms to teach medical terminology with the results that this approach significantly increased students' engagement and improved their vocabulary acquisition. Moreover, some researchers have integrated some other approaches such as metaverse-based flipped classrooms (Celik & Baturay, 2024), video-based lectures (Shi et al., 2024) into traditional classrooms to assist students in learning technical vocabulary. These studies indicate that the teaching and learning of technical vocabulary have gradually gained attention from both educators and researchers. However, the integration of advanced technologies such as DGBL, artificial intelligence, and adaptive learning systems into technical vocabulary instruction remains insufficiently explored. Given the proven effectiveness of these emerging technologies in general vocabulary acquisition, researchers argue that DGBL is particularly effective for teaching complex systems, such as scientific or social issues, by employing a "situated cognition" approach. This method could place learners in immersive environments that require the application of specific knowledge and skills to solve problems and achieve objectives on one hand; on another hand, it fosters exploration and experimentation within a safe and interactive settings for leaners (Pitarch, 2024). Moreover, some online games offer adaptive mechanisms that dynamically adjust the difficulty of language content based on learners' proficiency levels, thereby enabling personalized and self-paced progression and enhancing the overall impact of DGBL as a versatile and transformative pedagogical tool (Plass et al., 2015).

Despite growing interest in the application of DGBL for technical vocabulary acquisition and calls for synthesizing existing academic findings to advance this field, no comprehensive review has yet been conducted to consolidate the available evidence and analyze its implications. In particular, there is a lack of systematic understanding regarding the types of digital games employed for technical vocabulary instruction, the specific learner groups and disciplinary domains targeted, and the range of learning outcomes observed. Moreover, the connections between game design elements and learner variables, such as motivation, anxiety, and cognitive engagement, remain underexplored in the context of technical vocabulary learning. To address this gap, the present study aims to systematically review empirical research on the application of DGBVL in the technical aspect, providing an overview of the development of this research topic, the types of games employed, and the learner groups that can benefit from this approach.

Methodology

In this section, based on the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines, the specific review methods, including the search strategy, inclusion criteria, and exclusion criteria, will be discussed in detail to facilitate the achievement of the research objectives.

Identification

This study follows the PRISMA 2020 checklist (Figure Three) to guide the data retrieval and literature screening process, ensuring the quality of the systematic review (Page et al., 2021). The literature search spans from 2014 to December 2024 and utilizes six databases: Web of

Science (WoS), Scopus, ScienceDirect, Google Scholar, Springer, and Taylor & Francis. The initial search was conducted across five databases, excluding Google Scholar, by using the keywords "digital game-based learning" OR "game-based learning" OR "video game" OR "serious game" OR "educational game" OR "digital game" AND "technical vocabulary" OR "specialized vocabulary" OR "academic vocabulary" OR "professional vocabulary" OR "terminology". This search resulted in the identification of 284 publications. To ensure comprehensive coverage, a manual search by applying the similar keywords was carried out on Google Scholar, targeting publications not indexed in the other databases. This process generated an additional seven publications. In total, 291 publications were initially identified. Following a thorough duplicate screening process, nine duplicated articles were removed. Consequently, 282 unique publications for subsequent analyses.

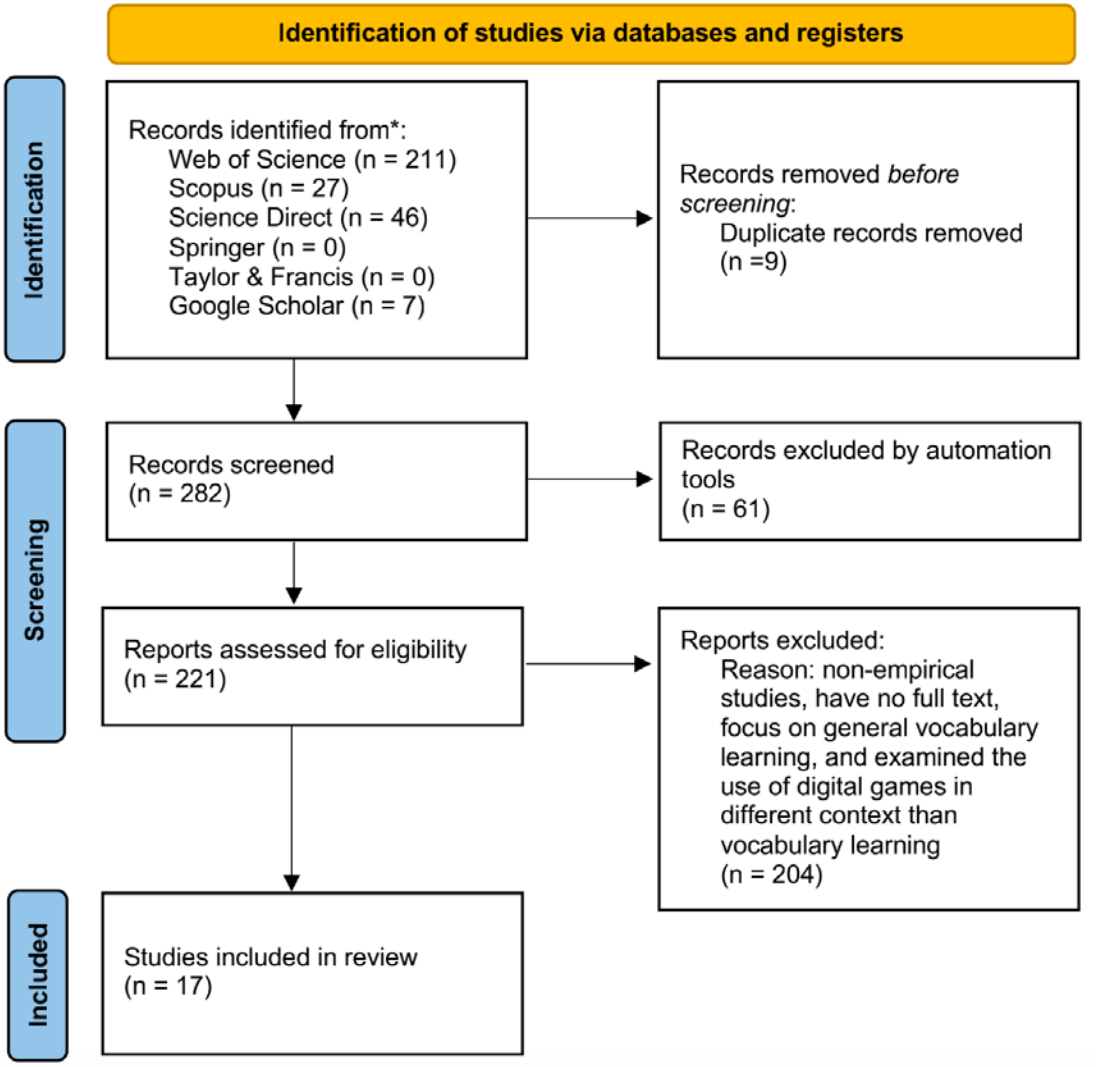


Figure 3. PRISMA flow diagram for screening and selecting studies

Screening and eligibility criteria

The screening phase aims to exclude publications that are not relevant to the objectives of this study. This process relies largely on automated filtering tools provided by the databases,

applying specific criteria such as publication year, type of publications, and language of the text (as outlined in Table 2). Although the concept of DGBL was first coined by Prensky in 2001, the current study focuses on analyzing research from the past decade (2014-2024) to account for the rapid advancements in educational technology. Notably, a review of earlier studies revealed that the volume of relevant research on DGBVL before 2014 was minimal (Vnucko & Klimova, 2023). To ensure comprehensive coverage of the topic, the inclusion criteria were broadened to encompass a range of publication types, including peer-reviewed journal articles, conference papers, and book chapters, which need to exceed five pages in length. The writing language is restricted to English to maintain consistency and accessibility in the analysis. Following the automated screening process, 61 publications were excluded as they did not meet the predefined criteria. This step resulted in a total of 221 publications deemed eligible for further evaluation in the subsequent phase of final eligibility.

The final stage of the systematic review process is eligibility determination, which involves a manual evaluation of the remaining publications from the previous screening phase. At this stage, the full texts and the abstracts are carefully examined to ensure alignment with the precise focus of the study. The primary inclusion criteria for this phase (as detailed in Table Two) require that eligible publications must present empirical research on the application of DGBL in technical vocabulary acquisition. Furthermore, only studies with accessible full texts were included. Specifically, studies were removed if they: (1) did not focus particularly on technical vocabulary acquisition, (2) were not empirical study, or (3) did not employ DGBL as a learning approach. This systematic manual evaluation ensured that only studies directly contributing to the research objectives and aligned with the scope of the review remained. As a result, a total of 204 publications were excluded from the analysis, leaving 17 studies that met the eligibility criteria and were deemed suitable for further review.

Table 2. *Inclusion and exclusion criteria*

Stage	Methods	Criteria	Inclusion	Exclusion
Screening	Automatic tools	Timespan	2014-2024	< 2014
		Sources	Article, conference paper, book chapter (more than 5 pages)	Thesis, books, reviews
		Language	English	Non-English
Eligibility	Manual evaluation	Research objective	Technical vocabulary acquisition	Other types of vocabulary or language learning
		Type of study	Empirical study	Non-empirical study
		Learning approach	DGBL	Non-DGBL

Results

This section presents the findings of 17 empirical studies published between 2014 and 2024 that investigated the application of DGBL in technical vocabulary acquisition, as summarized in Appendix One. The analysis highlights key aspects of these studies, including the distribution of publications over time, the types of games employed, the variables examined, and the outcomes reported.

Trends of the integration of DGBL in technical vocabulary from 2014 to 2024

Based on the data presented in Appendix A, a figure was generated (Figure Four) illustrating the annual publication trends for studies on the application of DGBL in technical vocabulary acquisition from 2014 to 2024. The figure illustrates a slow developmental pace in this research area prior to 2020, with an average annual publication rate of only one paper. Notably, no publications were recorded in 2018, reflecting minimal study activity during this period. However, a marked increase is observed from 2021 to 2024, with the average annual publication rate rising to 2.5 studies, which is 2.5 times increase compared to the period before 2020. The year 2024 stands out with four publications, representing the peak in this timeline. This upward trend suggests a growing interest in the integration of DGBL into technical vocabulary acquisition.

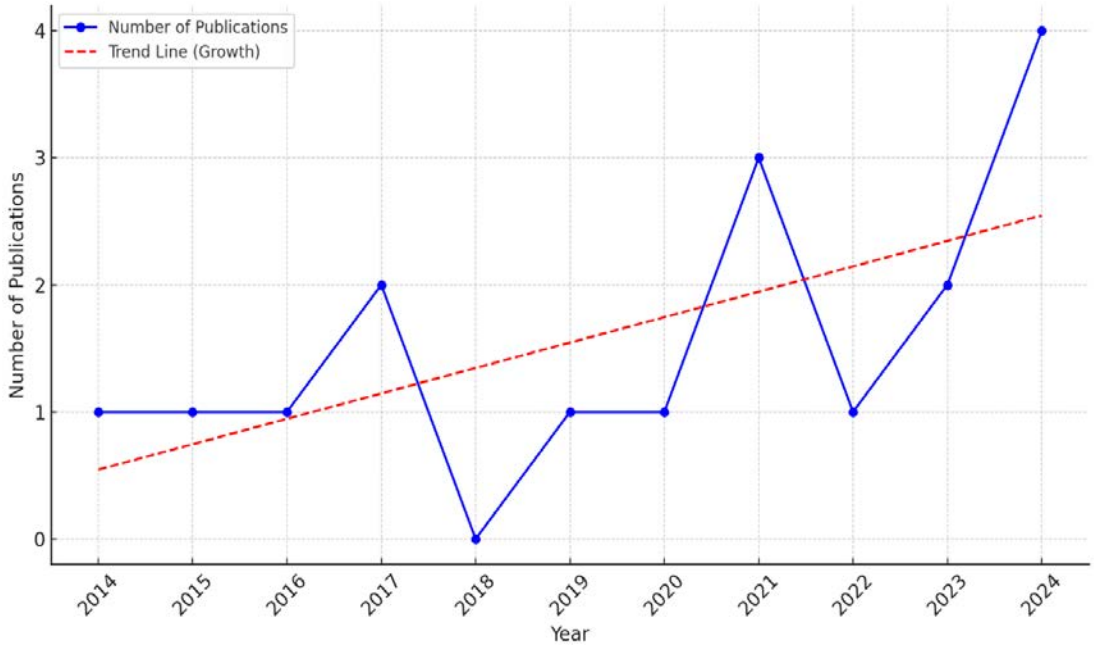


Figure 4. Annual scientific publications of DGBL in technical vocabulary acquisition

Compared to the application of DGBL in general vocabulary acquisition (as illustrated in Figure Five), it is evident that research on its application in technical vocabulary acquisition has lagged behind in terms of both growth rate and publication volume. The disparity demonstrates that the integration of DGBL into technical vocabulary learning has yet to achieve the same level of attention and implementation as its general vocabulary counterpart. However, starting from 2020, there has been an acceleration in research focused on the application of DGBL for both general and technical vocabulary acquisition. This progression highlights a growing recognition of the potential of DGBL approach to revolutionize language learning practices, driven by advancements in educational technology and gamified techniques, warranting further exploration and discussion.



Figure 5. Comparison of DGBL in technical and general vocabulary acquisition

Geographical distribution of publications

The geographical distribution of the 17 publications, as shown in Figure Six, is predominantly concentrated in Asia, Europe, and North America, reflecting significant regional differences in this research subject. Asia (N=12, 71%) emerges as the most active region, contributing the highest proportion of studies, followed by Europe (N=4, 24%). In North America, only the United States, has contributed to this field, with a single study (6%) published in 2015. Notably, China (N=5, 29%) stands out as the leading country in this field, producing the largest number of studies on the application of DGBL in technical vocabulary acquisition, followed by Malaysia (N=3, 18%). Turkey and Spain share the third position, each contributing two publications (12%). Other countries, including France, Ukraine, Kazakhstan, and Iran, have each produced a single study, matching the output of the United States. This distribution highlights a clear regional disparity, with Asia taking the current lead in advancing research on the employment of DGBL in technical vocabulary.

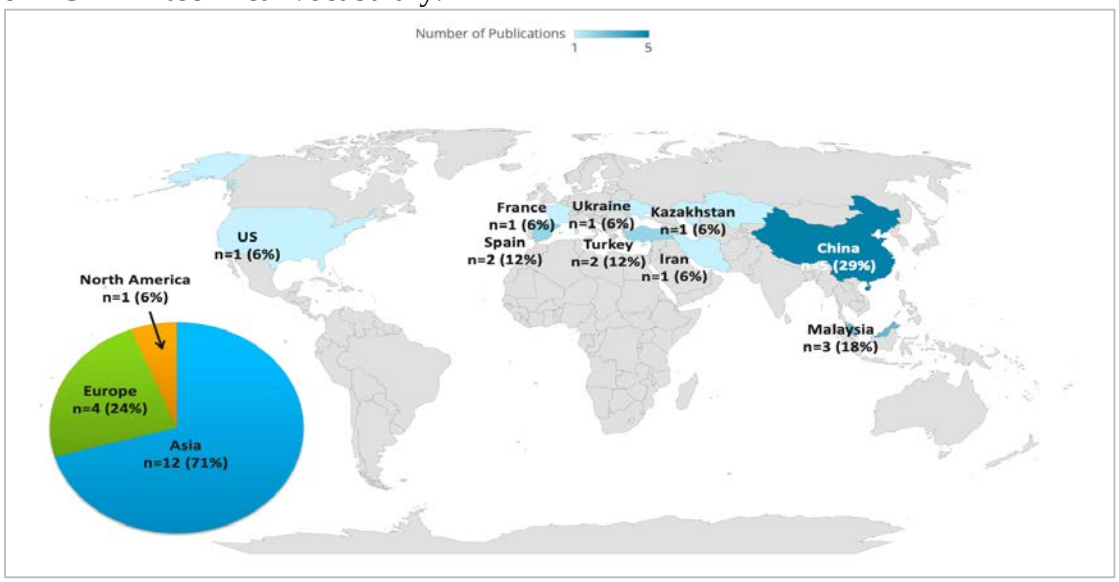


Figure 6. Geographical distribution of publications for DGBL in technical vocabulary acquisition

ISSN: 2229-9327

Target languages and games techniques

Among the 17 studies examined, the primary target language, as in much other research on language acquisition, is predominantly English. Sixteen out of the total studies (94%) investigate the application of DGBL in English technical vocabulary acquisition. The only exception is a study conducted by Neville (2015), which explores the employment of DGBL in learning German technical vocabulary. This finding aligns with the conclusion of other researchers (e.g., Hwang et al., 2023; Xu et al., 2020) in their literature review on DGBLL, which also highlight English as the most popular target language in this domain. This imbalance underscores that the current focus on English reflects its global prominence as a lingua franca, but it also emphasizes the limited attention given to other languages.

The types of games employed in these empirical studies (as illustrated in Figure Seven) can be analyzed from two perspectives: whether the games were specifically designed, and the forms of the games employed. First of all, among the 17 studies analyzed, four utilized games specifically developed for particular target audiences and specialized technical vocabulary. In the study of Wu et al. (2020), the designed program was based on the Bingo game prototype to facilitate medical vocabulary learning by strengthening the mapping between word forms and meanings. This research incorporated multimedia elements such as music, visuals, pronunciation, and animations to enhance engagement and effectiveness. The research of Kohnke and Ting (2021) employed a vocabulary learning app developed by Hong Kong Polytechnic University, tailored for students across seven professional fields, including textiles and fashion, design, business, nursing, engineering, tourism, as well as hospitality management. The app featured four difficulty levels for each discipline, with each level comprising 15 items. It also applied an adaptive system to help students build personalized vocabulary banks by revisiting unfamiliar or incorrectly answered words. Another study by Kaldarova et al. (2023) introduced a developed 2D video game to help primary students acquire computer science terminology. Lastly, Jia et al. (2023) designed an AR game based on flashcard and matching exercises to teach key digestive system terms. In contrast, the remaining 13 studies relied on existing platforms or commercially available games, reflecting a predominant preference for readily accessible resources over custom-designed solutions.

Second, in terms of game forms, only Jia et al. (2023) employed a developed AR game as the learning platform for the experimental group, comparing its effectiveness with two control groups, one using the online Quizlet and the other utilizing physical flashcards with puzzles. In the other 16 studies, 7 incorporated video games spanning 2D (Lu & Chang, 2016), 3D (Pitarch, 2024), and Kinect-based (Urun et al., 2017) forms, while 9 employed specialized vocabulary games, such as Quizlet (Bakhmat & Chaikovska, 2021) and Kahoot! (García-Pinar, 2024), as well as other tools like chatbots (Hsu & Chen, 2024). This demonstrates that, despite the relatively small number of studies on DGBL in technical vocabulary acquisition, the types of games applied are comparatively diverse.

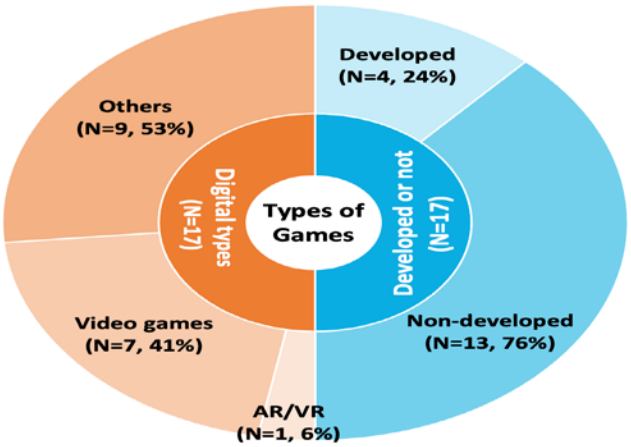


Figure 7. The types of games employed in reviewed studies

The game mechanics employed across the reviewed studies can be broadly categorized into five types: competitive, exploratory, collaborative, adaptive/self-paced, and feedback-based. Competitive games, such as matching or puzzle-based games (e.g., Jia et al., 2023; Chen et al., 2020), were often associated with enhanced motivation and improved retention. Exploratory mechanics, as found in 3D and narrative-driven games (e.g., Pitarch, 2024), promoted immersion and engagement. A few studies (e.g., Urun et al., 2017; García-Pinar, 2024) adopted collaborative elements, encouraging peer interaction and enhancing social motivation. Others integrated adaptive features (e.g., Kohnke & Ting, 2021) to provide personalized learning paths, while several studies emphasized immediate and low-stress feedback (e.g., Wu et al., 2020) to create emotionally supportive environments. This diversity of game mechanics suggests that DGBL approaches to technical vocabulary acquisition are not monolithic but vary according to learning goals, technological resources, and instructional contexts.

Research fields of technical vocabulary

Of those 17 studies reviewed, the range of domains involved in the implementation of DGBL for technical vocabulary acquisition remains relatively narrow as shown in Figure Eight. The majority of studies focus on engineering vocabulary (N=7), followed by medical (including nursing) terminology (N=5), with commercial vocabulary ranking third with three studies. Other fields, such as recycling and waste management, military, aviation, textiles and clothing, design, hotel management, as well as computer science, are each representing by only one research. The percentage of studies is not indicated because, as mentioned above, seven distinct professional disciplines were covered in the research of Kohnke and Ting (2021). Notably, many commonly academic fields (e.g., physics, chemistry, and biology) remain unexplored in this field. Moreover, among these studies, only the research of Kaldarova et al. (2023), which focused on computer science vocabulary acquisition, targeted primary school students. All other studies involved participants from higher education. This indicates that the scope of existing research remains relatively narrow, both in terms of the demographic groups studied and technical fields addressed.

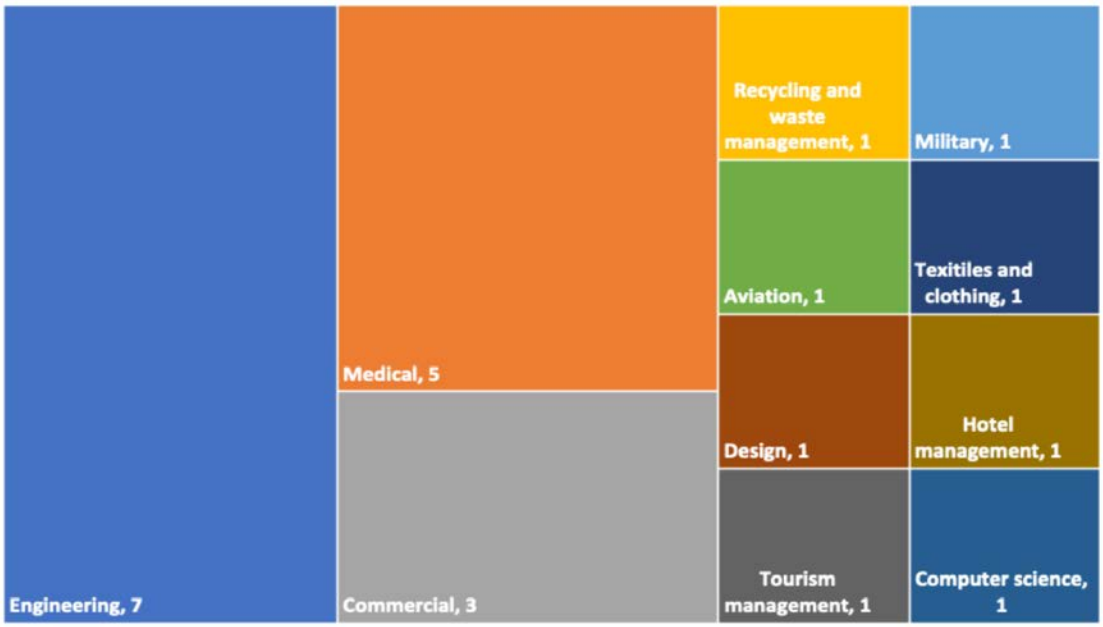


Figure 8. Fields of technical vocabulary in reviewed studies

Research outcomes

Figure Nine illustrates the variables examined in the reviewed studies. Notably, performance is the most frequently studies variables, with 14 out of 17 research addressing vocabulary acquisition results. Among these, only the study of Jia et al. (2023) explored the impact of DGBL on both short-term and long-term vocabulary retention, providing rare yet valuable insight into the sustained effect of DGBL on technical vocabulary learning. However, the limited inclusion of delayed post-tests across other studies suggests that long-term retention remains an underexplored area in this field. The second most commonly examined variables are students' attitudes toward this learning method and improvement of their motivation. However, no studies have specifically investigated the effects of DGBL on intrinsic and extrinsic motivation in technical vocabulary acquisition. Moreover, studies on factors related to acceptance rank forth, with two studies addressing this variable. Other research variables, such as cultural immersion, participation, metacognitive awareness, engagement, and satisfaction, are each represented by a single study.

Regarding research outcomes, three studies reported either inconclusive or negative results. For instance, in the study of Neville (2015), the primary aim was to determine whether immersion in a 3D-DGBLL environment enhanced participants' ability to produce written L2 narratives. However, because the 3D game was not specifically designed to reflect local contexts, the experimental group reported a lower sense of cultural immersion compared to the text-based control group, which was closely tied to local culture. Furthermore, the study of Ali and Muhammad (2019) found no significant differences in perceived usefulness or attitudes between male and female students when using games to learn technical vocabulary. Meanwhile, due to the short duration of the intervention (30-40 minutes per session), students displayed only moderate enthusiasm toward this learning method. Despite these findings, the remaining 14 studies demonstrated overwhelmingly positive outcomes. These studies confirmed that DGBL has a significant motivational and facilitative effect on students' acquisition of technical vocabulary, as well as on their learning motivation, attitudes, acceptance, satisfaction, and engagement. These results underscore the potential of DGBL as an effective tool for enhancing technical vocabulary learning across various dimensions. Nevertheless, it is important to note that only one out of the 17 studies (Kaldarova et al., 2023) involved K-12 participants, highlighting a marked absence of research focused on younger learners. Given the increasing integration of digital tools in school-level education, further studies targeting primary and secondary learners are urgently needed to validate the generalizability and scalability of DGBL in diverse educational contexts.

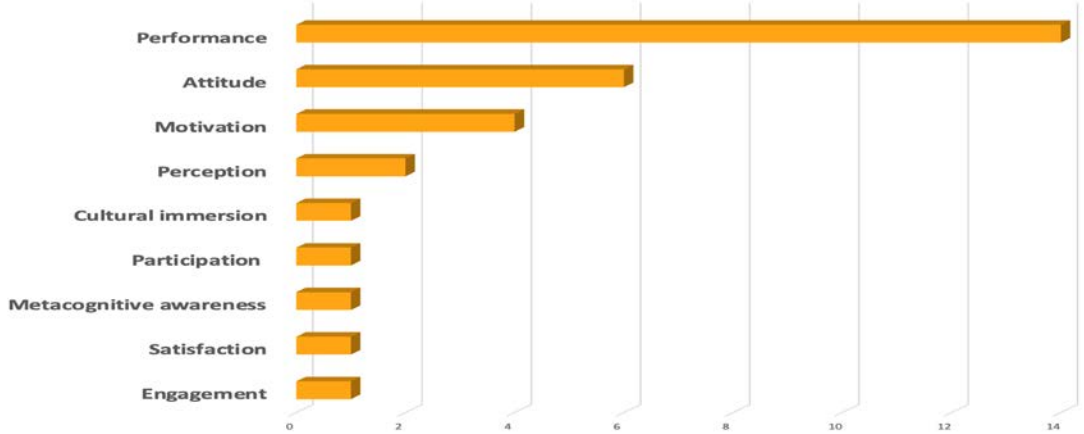


Figure 9. Research variables studied in reviewed publications

Discussion

This review synthesizes the findings from 17 empirical studies on applying DGBL for technical vocabulary acquisition. The analysis effectively addresses the first four research questions posed, offering a comprehensive overview of the field. The following discussion of these findings will highlight key research patterns and gaps within the existing body of research, providing valuable insights that can inform future investigations and educational practices.

Research trends and implications

The reviewed studies reveal a modest yet steady growth in DGBL research for technical vocabulary acquisition since 2020. However, this research remains predominantly concentrated in specific technical fields, such as engineering and medicine. While these domains are undeniably important and have yielded valuable insights, other disciplines, such as the humanities and commonly studied natural sciences (e.g., agriculture, architecture, and chemistry), have received insufficient attention. This imbalance highlights the narrow scope of current research and emphasizes the need to expand DGBL applications to a broader range of technical vocabulary domains to meet the diverse vocabulary acquisition needs of learners across various fields. Additionally, the majority of studies target population from higher education, with only the study of Kaldarova et al. (2023) being the single example focused on primary school students. This indicates a significant absence of K-12 learner populations in current research, which limits our understanding of how DGBL may benefit younger or secondary-level learners in acquiring subject-specific vocabulary. This finding aligns with the conclusions drawn by Vnucko and Klimova (2023) in their systematic review of DGBVL studies. This limited focus on population demographics constrains our understanding of the adaptability and effectiveness of DGBL across different age groups and educational levels.

Next, the analysis reveals that DGBL for technical vocabulary relies on off-the-shelf video games or educational platforms, such as Quizlet and Kahoot!, with limited use of custom-designed games. Among the 17 reviewed studies, only 4 employed games specifically tailored to learners' needs. Given the specialized nature of technical vocabulary, this reliance on generic platforms may compromise the effectiveness of DGBL in addressing learners' unique requirements. For instance, the research of Neville (2015) highlights this limitation, demonstrating that students in a non-customized 3D-DGBLL environment, which lacked strong cultural connections to the local context, experienced weaker cultural immersion compared to a control group using culturally relevant text-based materials. This finding underscores the importance of developing DGBL tools that align with learners' cultural and contextual realities. This is the same conclusion that Zhang et al. (2024) reached in their study, which also confirmed that purposefully developed games can be more effective in promoting academic development than commercial products since the solid theoretical foundation and precisely tailored learning objectives can serve students more accurately. In addition to form and customization, the reviewed studies also suggest variation in the game mechanics employed. While some studies utilized competitive formats (e.g., Kahoot!, quiz-based AR games), others adopted collaborative tasks (e.g., chatbot, paired challenges). These distinctions in mechanics appear to influence learning outcomes: competition is frequently linked to increased motivation and focus (Chen et al., 2023), while collaboration fosters communication skills and learner autonomy (Yang et al., 2025). However, most studies did not explicitly examine these relationships, and few reported comparative evidence across game mechanics.

Third, the linguistic scope of DGBL research predominantly focus on English as a second language, with limited attention given to multilingual learners who need to acquire technical vocabulary across multiple languages. This trend aligns with findings from general vocabulary acquisition studies (Hung et al., 2018; Xu et al., 2020), which similarly emphasizes a lack of diversity in linguistic contexts. This linguistic homogeneity also limits theoretical generalizability across diverse learner populations. According to Sweller's Cognitive Load Theory, additional language switching and unfamiliar technical terms may increase extraneous cognitive load (Bahari, 2023). Future DGBL research should investigate how multilingual game-based environments can be designed to reduce such load and support meaningful vocabulary transfer.

The distribution of variables examined in the analyzed studies highlights a strong emphasis on learning outcomes, with 14 out of 17 studies focusing on vocabulary acquisition. The results consistently demonstrate the effectiveness of DGBL in facilitating technical vocabulary learning. For instance, Soyoof et al. (2024) reported that students learning technical vocabulary through DGBL showed an average vocabulary knowledge improvement of 70.87%, compared to only 15.32% in the control group. Similarly, Lelardeux et al. (2017) found that DGBL could enhance vocabulary acquisition across an entire class. Some researchers (e.g., Abrams & Walsh, 2014; Chowdhury et al., 2024) attribute this effectiveness to the nature of digital games. While traditional textbook learning often exposes students only to the orthographic form of words, games offer a richer, multisensory experience by presenting vocabulary in both visual and auditory contexts (Simonnet et al., 2025). This immersive exposure facilitates a deeper understanding of word usage, likely due to the combination of meaningful contextualization and multimedia diversity, which significantly improve vocabulary retention during gameplay (Chen & Hsu, 2020). This finding is supported by Ramezanali and Faez (2019), who demonstrated that using dual annotations, combining L2 definitions with video aminations, is more effective for vocabulary learning and retention than relying on single annotations with L2 definitions. Likewise, Sundqvist (2019) confirmed the impact of interactive digital gaming environments in facilitating the mastery of new and particularly challenging vocabulary. The dynamic interactions within these games offer learners repeated encounters and contextual applications, which are essential for grasping intricate terms. Furthermore, specific psychological factors such as intrinsic and extrinsic motivation learning confidence, and cognitive load have received minimal attention in existing research.

Finally, the findings from the reviewed literature emphasize numerous advantages of DGBL for learners, aligning with the results of many studies on general vocabulary acquisition. Specifically, DGBL has been shown to aid learners in acquiring foreign language vocabulary through engaging gameplay (Kazu & Kuvvetli, 2023), enhance their motivation (Zhang et al., 2024) and engagement (Albayrak et al., 2022), reduce their cognitive load (Liu, 2024) and anxiety (Al-Aosail et al., 2024), establish their autonomous learning behavior (Fu et al., 2021), etc. However, a few studies indicate that the impact of DGBL on technical vocabulary acquisition may not always be significant. For instance, the study of Jia et al. (2023) revealed no notable differences between the AR-GBL group and the paper-based GBL group in terms of immediate vocabulary learning outcomes, although the AR-GBL group demonstrated superior in long-term retention. This finding is consistent with Lai and Chang (2021), who reported no significant differences between AR and flashcards in facilitating English vocabulary acquisition among first-grade students. Jia et al. (2023) attributed this to the potential additional cognitive load introduced by using digital devices and applications. Likewise, the research of Ebrahimzadeh (2017) echoed this perspective, suggesting that the high interactivity of games may impose an extra cognitive burden

on learners, which could contribute to the forgetting of some vocabulary. These mixed findings suggest that while DGBL holds promise as an effective tool for vocabulary learning, its design and implementation must carefully balance engagement and cognitive demands. Meanwhile, researcher have observed a lack of uniformity in the test forms used to evaluate technical vocabulary acquisition in the reviewed studies. This observation is consistent with the study of Vnucko and Klimova (2023), who noted the absence of standardized tools for assessing vocabulary in the digital game environment. This may draw from the complexity of technical vocabulary originating from multiple disciplines. Additionally, the findings underscore the limited attention given to the long-term retention of technical vocabulary. However, retaining and recalling vocabulary over extended periods is essential for effective and fluent communication (Masood, 2024). Among all reviewed studies, only Jia et al. (2023) provided a direct comparison between short-term and long-term retention outcomes, revealing the significant advantage of DGBL in supporting delayed recall. This gap calls for more longitudinal research that not only measures immediate post-test outcomes but also examines how learners retain specialized vocabulary weeks or months after initial learning. Many empirical studies have confirmed that multimodal multimedia input can significantly enhance long-term vocabulary retention (e.g., Teng, 2023; Teng & Zhang, 2023). This presents an important opportunity for future research to examine whether DGBL approaches yield comparable benefits for the enduring acquisition of technical vocabulary.

Suggestions for future studies

Given the relatively limited number of studies in this field, there are numerous possibilities for future research to delve deeper into various aspects of DGBL for technical vocabulary acquisition. First, from a multilingual and multidisciplinary perspective, the research could be extended to languages other than English. Such efforts would not only promote the research for linguistic diversity but also address the unique needs of learners in specialized linguistic environments. Simultaneously, the scope of research could be expanded to include additional disciplines such as chemistry, biology, and physics. These efforts would facilitate a better understanding of the adaptability and effectiveness of DGBL across diverse academic and professional contexts.

Then, although many existing studies focus on the immediate benefits of vocabulary acquisition, relatively little is known about the long-term retention of technical vocabulary or its practical application in academic or professional contexts. Future research could investigate, through longitudinal research, how learners retain technical vocabulary over time and how they utilize it in the reality, such as writing academic papers, delivering presentations, or engaging in professional communication.

In terms of game types, research on AR/VR-based games and developed games are very limited. Future educators could develop games which meet the needs of learners and the requirement of technical terms, incorporating professional and cultural elements. Moreover, future directions could explore the use of AI-driven games (Polyzi & Moussiades, 2023) to support personalized learning paths for technical vocabulary acquisition. Additionally, the potential of AR, VR, and AI could be explored to create adaptive, interactive, and immersive DGBL environments. Researchers could also investigate the interdisciplinary transferability of these tools, examining how games designed for one discipline could be adapted to others.

Finally, future research could also combine quantitative and qualitative research methods to understand the application of DGBL in technical vocabulary acquisition from multiple research

perspectives. Broader participant demographics, including young learners, adult professionals, and leaners from nontraditional educational settings, should also be considered to explore the adaptability of DGBL across different population. It is also worth exploring how teachers can provide instructional support for students in digital teaching environments and different teaching methods when using this technology.

Limitations

The study may have certain limitations in its literature selection process. First, several articles did not explicitly use terms related to technical vocabulary in their titles or keywords, such as the research of Dinçer and Dinçer (2021) who use "aviation English" as keywords. This created challenges in identifying relevant studies and may have led to the omission of some pertinent publications. Second, the evaluated publications were sourced from six databases: WoS, Scopus, ScienceDirect, Google Scholar, Springer, and Taylor & Francis. But unfortunately, there are very few journals dedicated to research on DGBL for technical vocabulary acquisition, and the inclusion criteria excluded master's and doctoral theses. Future research could address these limitations by broadening the search scope to include additional databases and publication types. Moreover, of those 17 reviewed studies, very few explicitly mentioned their theoretical foundations. Consequently, this study did not analyze the theoretical frameworks underlying the research. This aligns with the observation of Bodily et al. (2019) that articles in instructional design and technology journals rarely discuss learning theories or conceptual frameworks. Despite these limitations, they should not diminish the significance of this systematic review. It provides a comprehensive summary of the overall development of DGBL research in the field of technical vocabulary acquisition, offering valuable insights for future investigations and practical applications.

Conclusion

This study systematically reviews research on DGBL for technical vocabulary acquisition published between 2014 and 2024, yielding several key insights. The primary aim was to provide a comprehensive overview of current practices, learner populations, game types, and measured outcomes related to the implementation of DGBL in technical vocabulary instruction. First, most of these studies target English as the primary language and focus on population from higher education, with a significant concentration of research conducted in Asia. Second, the research platforms utilized predominantly consist of existing video games or online learning tools, with relatively few studies dedicated to games developed specifically for targeted learner population. Additionally, the research emphasis has primarily been on learning performance, attitudes, and motivation. However, studies investigating the long-term retention of technical vocabulary, intrinsic and extrinsic motivation, as well as affective, social, and metacognitive learning strategies, remain limited. These areas, critical to a holistic understanding of GGBL, have largely been overlooked. Fourth, the scope of technical fields explored is relatively narrow, with most studies concentrating on engineering, medicine, and business, leaving many other disciplines underrepresented.

Despite these limitations, the majority of the reviewed studies confirm the effectiveness of DGBL in facilitating technical vocabulary acquisition, though a few report less significant findings. Encouragingly, the number of publications in this field has been steadily increasing since 2020. Research has also evolved from employing basic computer-based games to incorporating

mobile platforms and diversified VR technologies, signaling a growing recognition of the potential of DGBL in foreign language education. Many researchers (e.g., Su & Zou, 2022; Zhang, 2022) similarly affirm that technological advancements have transformed language education, offering opportunities to enhance the efficiency of language learning and teaching. Accordingly, future research is recommended to explore the potential of DGBL among a wider range of learner groups, including vocational trainees and young learners in primary and secondary education. Such efforts would help assess the scalability and adaptability of DGBL across different educational levels, further advancing the field and supporting inclusive language learning practices. Exploring AI-driven, personalized game-based learning paths may also offer promising directions for enhancing learner engagement and long-term vocabulary retention.

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

Feng Tian is currently a PhD candidate at the Centre for Instructional Technology and Multimedia, Universiti Sains Malaysia. Her primary research interests focus on the application of digital technologies in foreign language education, particularly in teaching and learning technical vocabulary in French. ORCID: https://orcid.org/0009-0006-1905-3921

Irfan Naufal Umar is a professor at the Centre for Instructional Technology and Multimedia, Universiti Sains Malaysia, and the Director of the University Academic Quality Centre at the same university. His primary research interests include the application of educational technology and e-learning. ORCID: https://orcid.org/0000-0003-1719-5679

Reference

Abdullah Alhebshi, A., & Gamlo, N. (2022). The effects of Mobile Game-Based Learning on Saudi EFL foundation year students' vocabulary acquisition. *Arab World English Journal*, *13*(1), 408–425. https://doi.org/10.24093/awej/vol13no1.27

Abrams, S. S., & Walsh, S. (2014). Gamified vocabulary: Online resources and enriched language learning. *Journal of Adolescent & Adult Literacy*, *58*(1), 49–58. https://doi.org/10.1002/jaal.315

Al-Aosail, A. B. F., Lamouchi, A., Ibrahim, S. A., Gafar, M. G., & Hafez, W. E. (2024). Digital games for learning vocabulary: The effects on foreign language anxiety, buoyancy, online engagement, and vocabulary learning. *CALL-EJ*, *25*(4), 27–50.

Albayrak, M., Chen, J., & Williams, P. J. (2022). Exploring ESL students' perceived engagement in and experience of content vocabulary learning through virtual reality games.

In *Digital Games in Language Learning* (pp. 109–135). https://doi.org/10.4324/9781003240075-7

Alfehaid, A. (2019). Online English language learning activities and academic achievement: Experiences of first year students and their teachers. *Pertanika Journal of Social Sciences and Humanities*, *27*(3), 1557–1572.

Alharbi, K. (2023). Enhancing technical vocabulary acquisition: A multimedia strategic approach for English for specific purposes during COVID-19. *Interactive Learning Environments*, *0*(0), 1–19. https://doi.org/10.1080/10494820.2023.2205902

Ali, Z., & Ghazali, M. A. I. M. (2016). Learning technical vocabulary through a mobile app: English language teachers' perspectives. *International Journal of Language Education and Applied Linguistics*. https://doi.org/10.15282/ijleal.v4.487

Ali, Z., & Muhammad, N. N. (2019). Gender differences in perceived usefulness and attitudes towards VocBlast in learning specialized vocabulary. In M. Y. Mohamad Noor, B. E. Ahmad, M. R. Ismail, H. Hashim, & M. A. Abdullah Baharum (Eds.), *Proceedings of the Regional Conference on Science, Technology and Social Sciences (RCSTSS 2016)* (pp. 237–245). Springer. https://doi.org/10.1007/978-981-13-0203-9_22

Ali, Z., Sinnadurai, S. K., Hamzah, F., Azmuddin, R. A., & Sabuan, N. S. (2014). Ten interesting facts about OnVac: A tool to learn engineering and technology vocabulary. *International Journal of Language Education and Applied Linguistics*. https://doi.org/10.15282/ijleal.v1.417

Al-Rashidi, A. H. (2025). Does digital task-based instruction make a difference in EFL university students' motivation in Saudi Arabia? An Active theory perspective. *Learning and Motivation*, *90*, 102115. https://doi.org/10.1016/j.lmot.2025.102115

Aravind, B. R., & Rajasekaran, V. (2021). Exploring dysphasia learners' vocabulary acquisition through the cognitive theory of multimedia learning: An experimental study. *International Journal of Emerging Technologies in Learning (iJET)*, *16*(12), 263. https://doi.org/10.3991/ijet.v16i12.22173

Ardasheva, Y., & Tretter, T. R. (2017). Developing science-specific, technical vocabulary of high school newcomer English learners. *International Journal of Bilingual Education and Bilingualism*, *20*(3), 252–271. https://doi.org/10.1080/13670050.2015.1042356

Babaie, H. (2022). On digital game-based reading strategy training: Implications for L2 reading ability development. *Computer-Assisted Language Learning Electronic Journal*, *23*(3), Article 3.

Bahari, A. (2023). Challenges and affordances of cognitive load management in technology-assisted language learning: A systematic review. *International Journal of Human–Computer Interaction*, *39*(1), 85–100. https://doi.org/10.1080/10447318.2021.2019957

Bakhmat, N., & Chaikovska, O. (2021). Web-based professional vocabulary learning with the high school students. In *International Conference on Economics, Law and Education Research (ELER 2021)* (pp. 318-324). Atlantis Press. https://doi.org/10.2991/aebmr.k.210320.054

Banu, A. R., Srabonty, S. N., Islam, M., Tisha, A. P., & Islam, U. S. (2025). Barricades to reading expertise among students of secondary high school of Bangladesh: Major challenges and insights. *Asian Journal of Social Sciences and Legal Studies*, *7*(2), 308–315.

Barcroft, J. (2018). *Vocabulary in Language Teaching*. Routledge.

Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing Words to Life (Second Edition): Robust Vocabulary Instruction* (2nd Edition). The Guilford Press.

Bendebbache, A., & Azeroual, D. (2022). *The effectiveness of using the mobile game simcity in teaching technical vocabulary to urban planning students*. http://localhost:8080/xmlui/handle/123456789/13584

Binon, J., & Verlinde, S. (2003). How FSP shaped itself in the teaching of FLE: From teacher training at university to the design of a dictionary of business French [Comment le FOS s'est profilé dans la didactique du FLE : de la formation d'enseignants à l'université à la conception d'un dictionnaire du français des affaires]. *Les Cahiers de l'ASDIFLE*, *14*, 17–42.

Bodily, R., Leary, H., & West, R. E. (2019). Research trends in instructional design and technology journals. *British Journal of Educational Technology*, *50*(1), 64–79. https://doi.org/10.1111/bjet.12712

Boyinbode, O. (2018). Development of a gamification based English vocabulary mobile learning system. *International Journal of Computer Science and Mobile Computing*, *7*(8), 183–191.

Buil, I., Catalán, S., & Martínez, E. (2018). Exploring students' flow experiences in business simulation games. *Journal of Computer Assisted Learning*, *34*(2), 183–192. https://doi.org/10.1111/jcal.12237

Bulut, A., & Yıldız, M. (2024). The impact of computer-assisted and direct strategy teaching on reading comprehension. *Journal of Computer Assisted Learning*, *40*(6), 3369–3389. https://doi.org/10.1111/jcal.13068

Calvo-Ferrer, J. R. (2021). Effectiveness of type of feedback and frequency on digital game-based L2 vocabulary acquisition. *International Journal of Game-Based Learning (IJGBL)*, *11*(3), 38–55. https://doi.org/10.4018/IJGBL.2021070103

Calvo-Ferrer, J. R., & Belda-Medina, J. (2021). The effect of multiplayer video games on incidental and intentional L2 vocabulary learning: The case of among us. *Multimodal Technologies and Interaction*, *5*(12), Article 12. https://doi.org/10.3390/mti5120080

Carrier, S. J. (2013). Elementary preservice teachers' science vocabulary: Knowledge and application. *Journal of Science Teacher Education*, *24*(2), 405–425. https://doi.org/10.1007/s10972-012-9270-7

Castellano-Risco, I., Alejo-González, R., & Piquer-Píriz, A. M. (2020). The development of receptive vocabulary in CLIL vs EFL: Is the learning context the main variable? *System*, *91*, 102263. https://doi.org/10.1016/j.system.2020.102263

Castillo-Cuesta, L. (2020). Using digital games for enhancing EFL grammar and vocabulary in higher education. *International Journal of Emerging Technologies in Learning (iJET)*, *15*(20), 116. https://doi.org/10.3991/ijet.v15i20.16159

Celik, F., & Baturay, M. H. (2024). The effect of metaverse on L2 vocabulary learning, retention, student engagement, presence, and community feeling. *BMC PSYCHOLOGY*, *12*(1), 58. https://doi.org/10.1186/s40359-024-01549-4

Chang, J. A., & Chang, C. S. (2023). Games applied to vocabulary learning: A bibliometric analysis on two decades of research. *Journal of Educational Media & Library Sciences*, *60*(1). https://doi.org/10.6120/JoEMLS.202303_60(1).0023.RS.AE

Chen, C.-H., Liu, J.-H., & Shou, W.-C. (2023). How competition in a game-based science learning environment influences students' learning achievement, flow experience, and learning behavioral patterns. *Journal of Educational Technology & Society*, *21*(2), 164-176.

Chen, H.-J. H., & Hsu, H.-L. (2020). The impact of a serious game on vocabulary and content learning. *Computer Assisted Language Learning*, *33*(7), 811–832. https://doi.org/10.1080/09588221.2019.1593197

Chen, M.-H., Tseng, W.-T., & Hsiao, T.-Y. (2018). The effectiveness of digital game-based vocabulary learning: A framework-based view of meta-analysis. *British Journal of Educational Technology*, *49*(1), 69–77. https://doi.org/10.1111/bjet.12526

Chowdhury, M., Dixon, L., Kuo, L.-J., Donaldson, J. P., Eslami, Z., Viruru, R., & Luo, W. (2024). Digital game-based language learning for vocabulary development. *Computers and Education Open*, *6*, 100160. https://doi.org/10.1016/j.caeo.2024.100160

Chung T. M., & Nation P. (2003). Technical vocabulary in specialised texts. *Reading in a Foreign Language*, *15*(2), 103–116.

Chung, T. M., & Nation, P. (2004). Identifying technical vocabulary. *System*, *32*(2), 251–263. https://doi.org/10.1016/j.system.2003.11.008

Coxhead, A., & Demecheleer, M. (2018). Investigating the technical vocabulary of Plumbing. *English for Specific Purposes*, *51*, 84–97. https://doi.org/10.1016/j.esp.2018.03.006

Del Moral, B. L. M., VanPutte, C. L., & McCracken, B. A. (2024). The use of role-play in the learning of medical terminology for online and face-to-face courses. *Advances in Physiology Education*, *48*(3), 578–587. https://doi.org/10.1152/advan.00273.2023

Dinçer, N., & Dinçer, R. (2021). The effect of a serious game on aviation vocabulary acquisition. *International Journal of Serious Games*, *8*(4), Article 4. https://doi.org/10.17083/ijsg.v8i4.464

Drayton, J., & Coxhead, A. (2023). The development, evaluation and application of an aviation radiotelephony specialised technical vocabulary list. *English for Specific Purposes*, *69*, 51–66. https://doi.org/10.1016/j.esp.2022.10.001

Ebrahimzadeh, M. (2017). Readers, players, and watchers: EFL students' vocabulary acquisition through digital video games. *English Language Teaching*, *10*, 1. https://doi.org/10.5539/elt.v10n2p1

Frumuselu A. D. (2018). The implications of cognitive load theory and exposure to subtitles in English foreign language (EFL). *Translation and Translanguaging in Multilingual Contexts*, *4*(1), 55–76. https://doi.org/10.1075/ttmc.00004.fru

Fu, Y., Zhang, L., Zhao, S., & Chen, Y. (2021). Perceptions of non-English major college students on learning English vocabulary with gamified apps. *International Journal of Emerging Technologies in Learning (iJET)*, *16*(18), 268–276. https://doi.org/10.3991/ijet.v16i18.24125

Gaffas, Z. M. (2024). Learning medical terminology in an ESP medical course: Vocabulary notebooks versus word lists. *Australian Review of Applied Linguistics*. https://doi.org/10.1075/aral.24070.gaf

García-Pinar, A. (2024). Throtel, throttel, trottel or trotel. Engineering undergraduates' perception of the use of Kahoot! To review course content. *Teaching English as a Second or Foreign Language--TESL-EJ*, *28*. https://doi.org/10.55593/ej.28111int

Gómez-Ortiz, M.-J., Domínguez Romero, E., & Bobkina, J. (2023). Instagram as a learning tool to improve technical vocabulary for sports science students. *Journal of Hospitality, Leisure, Sport & Tourism Education*, *32*, 100416. https://doi.org/10.1016/j.jhlste.2022.100416

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

Hestiana, M., & Anita, A. (2022). The role of movie subtitles to improve students' vocabulary. *Journal of English Language Teaching and Learning*, *3*(1), 46–53. https://doi.org/10.33365/jeltl.v3i1.1715

Hsu, M.-H., & Chen, Y.-H. (2024). Personalized Medical Terminology Learning Game: Guess the Term. *Games for Health Journal*, *13*(2), 84–92. https://doi.org/10.1089/g4h.2023.0054

Hsueh-Chao, M. H., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, *13*(1), 403–430.

Huei, L. S., Yunus, M. M., & Hashim, H. (2021). Strategy to improve English vocabulary achievement during COVID-19 epidemic. Does Quizizz help? *Journal of Education and E-Learning Research*, *8*(2), 135–142. https://doi.org/10.20448/journal.509.2021.82.135.142

Hung, H.-T., Yang, J. C., Hwang, G.-J., Chu, H.-C., & Wang, C.-C. (2018). A scoping review of research on digital game-based language learning. *Computers & Education*, *126*, 89–104. https://doi.org/10.1016/j.compedu.2018.07.001

Hwang, G.-J., Chen, P.-Y., Chu, S.-T., Chuang, W.-H., Juan, C.-Y., & Chen, H.-Y. (2023). Game-based language learning in technological contexts: An integrated systematic review and bibliometric analysis. *International Journal of Online Pedagogy and Course Design (IJOPCD)*, *13*(1), 1–25. https://doi.org/10.4018/IJOPCD.316184

Javis, S. (1998). Conceptual transfer in the interlanguage lexicon. *Bloomington, IN: Indiana University Linguistics Club*.

Jia, W., Pack, A., Guan, Y., Zhang, L., & Zou, B. (2023). The influence of game-based learning media on academic English vocabulary learning in the EFL context. *Computer Assisted Language Learning*, 1–25. https://doi.org/10.1080/09588221.2023.2276800

Kaldarova, B., Omarov, B., Zhaidakbayeva, L., Tursynbayev, A., Beissenova, G., Kurmanbayev, B., & Anarbayev, A. (2023). Applying game-based learning to a primary school class in computer science terminology learning. *Frontiers in Education*, *8*. https://doi.org/10.3389/feduc.2023.1100275

Kaldenberg, E. R., Watt, S. J., & Therrien, W. J. (2015). Reading instruction in science for students with learning disabilities: A meta-analysis. *Learning Disability Quarterly*, *38*(3), 160–173. https://doi.org/10.1177/0731948714550204

Kamaşak, R., Sahan, K., & Rose, H. (2021). Academic language-related challenges at an English-medium university. *Journal of English for Academic Purposes*, *49*, 100945. https://doi.org/10.1016/j.jeap.2020.100945

Karimi, M. N., & Nasouri, A. (2024). EFL learners' flow experience and incidental vocabulary learning during text-based game tasks: The moderating role of working memory capacity. *System*, *124*, 103398. https://doi.org/10.1016/j.system.2024.103398

Kazu, İ. Y., & Kuvvetli, M. (2023). A triangulation method on the effectiveness of digital game-based language learning for vocabulary acquisition. *Education and Information Technologies*, *28*(10), 13541–13567. https://doi.org/10.1007/s10639-023-11756-y

Kohnke, L., & Ting, A. (2021). ESL students' perceptions of mobile applications for discipline-specific vocabulary acquisition for academic purposes. *Knowledge Management and E-Learning*, *13*(1), 102–117. Scopus. https://doi.org/10.34105/j.kmel.2021.13.006

Lai, J.-Y., & Chang, L.-T. (2021). Impacts of augmented reality apps on first graders' motivation and performance in English vocabulary learning. *Sage Open*, *11*(4), 21582440211047549. https://doi.org/10.1177/21582440211047549

Le, C. N. N., & Miller, J. (2023). A core meaning-based analysis of English semi-technical vocabulary in the medical field. *English for Specific Purposes*, *70*, 252–266. https://doi.org/10.1016/j.esp.2023.01.006

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. https://doi.org/10.1016/j.jeap.2016.01.008

Lelardeux, C. P., Galaup, M., Panzoli, D., Gilles, P., Minguella-Canela, J., Yassine-Diab, N., & Lagarrigue, P. (2017). Improving mechanical engineering vocabulary through the use of a game. *Procedia Manufacturing*, *13*, 1432–1439. https://doi.org/10.1016/j.promfg.2017.09.098

Lin, L. H. F., & Qian, D. D. (2019). The relationship between vocabulary knowledge and language proficiency. In *The Routledge Handbook of Vocabulary Studies*, 66-80. Routledge.

Liu, D. (2024). The effects of segmentation on cognitive load, vocabulary learning and retention, and reading comprehension in a multimedia learning environment. *BMC Psychology*, *12*(1). https://doi.org/10.1186/s40359-023-01489-5

Liu, D., & Lei, L. (2019). Technical Vocabulary. In *The Routledge handbook of vocabulary studies*, 111–124. Routledge.

Lu, F.-C., & Chang, B. (2016). Role-play game-enhanced English for a specific-purpose vocabulary-acquisition framework. *Educational Technology & Society*, *19*(2), 367–377.

Masood S. (2024). Cementing vocabulary in long-term memory: The influence of word association strategies. *International Journal of Literature and Languages*, *4*(12).

Miroshnychenko, I. (2024). Technical vocabulary in foreign language acquisition: Seeking effective methods for university students. *Collection of Scientific Papers «ΛΌГОΣ»*, Zurich, Switzerland. https://doi.org/10.36074/logos-24.05.2024.058

Moniza Ray, S. P., & Ajit, I. (2024). Digital game-based language learning: The impact of story-driven game Life is Strange 1 on language learners' listening skills. *Journal of Language Teaching and Research*, *15*(2), 448–457. https://doi.org/10.17507/jltr.1502.13

Mustaeva, G. S., Saidivalieva, B. S., & Ataeva, G. B. (2022). Features of professional vocabulary and its difference from terms. In *AIP Conference Proceedings* (Vol. 2432, No. 1, p. 060017). AIP Publishing LLC. https://doi.org/10.1063/5.0089598

Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge University Press.

Neville, D. O. (2015). The story in the mind: The effect of 3D gameplay on the structuring of written L2 narratives. *RECALL*, *27*(1), 21–37. https://doi.org/10.1017/S0958344014000160

Othman, J. (2024). Academic vocabulary learning in EMI classrooms: Challenges and strategies. *Arab World English Journal*, *15*, 3–18. https://doi.org/10.24093/awej/vol15no2.1

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

Pitarch, R. C. (2024). The engineer: Enhancing bachelor students' English technical engineering vocabulary proficiency with a serious video game. *Revista de Educación (Madrid)*, *405*, 293–310. https://doi.org/10.4438/1988-592X-RE-2024-405-636

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of Game-Based Learning. *Educational Psychologist*, *50*(4), 258–283. https://doi.org/10.1080/00461520.2015.1122533

Plass, J. L., Mayer, R. E., & Homer, B. D. (Eds.). (2020). *Handbook of Game-Based Learning*. The MIT Press.

Polyzi, P., & Moussiades, L. (2023). An artificial vocabulary learning assistant. *Education and Information Technologies*, *28*(12), 16431–16455. Scopus. https://doi.org/10.1007/s10639-023-11810-9

Pratiwi, T. L., Togatorop, S. A., & Marsevani, M. (2024). Enhancing students' speaking skills with Wordwall game-based learning: Classroom action research. *Journal of English Teaching, Applied Linguistics and Literatures (JETALL)*, *7*(2), Article 2. https://doi.org/10.20527/jetall.v7i2.19682

Ramezanali, N., & Faez, F. (2019). Vocabulary learning and retention through multimedia glossing. *Language Learning & Technology*, *23*(2), 105–124.

Rasti-Behbahani, A. (2021). Why digital games can be advantageous in vocabulary learning. *Theory and Practice in Language Studies*, *11*(2). https://doi.org/10.17507/tpls.1102.01

Rivers, W. M. (1983). *Communicating Naturally in a Second Language: Theory and Practice in Language Teaching* (6th ed.). Cambridge University Press.

Roohani, A., & Heidari Vincheh, M. (2023). Effect of game-based, social media, and classroom-based instruction on the learning of phrasal verbs. *Computer Assisted Language Learning*, *36*(3), 375–399. https://doi.org/10.1080/09588221.2021.1929325

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A Self-Determination Theory approach. *Motivation and Emotion*, *30*(4), 344–360. https://doi.org/10.1007/s11031-006-9051-8

Schmitt, N., & Schmitt, D. (2020). *Vocabulary in Language Teaching*. Cambridge University Press.

Shi, D., Révész, A., & Pellicer-Sánchez, A. (2024). The effects of task repetition on the processing and acquisition of technical vocabulary through video-lecture-based tasks: A mixed-methods study. *Language Learning*, *n/a*(n/a). https://doi.org/10.1111/lang.12679

Simonnet, E., Loiseau, M., & Lavoué, É. (2025). A systematic literature review of technology-assisted vocabulary learning. *Journal of Computer Assisted Learning*, *41*(1), e13096. https://doi.org/10.1111/jcal.13096

Song, S. (2018). Second language acquisition theories. In S. Song, *Second Language Acquisition as a Mode-Switching Process*, 9–36. Palgrave Macmillan UK. https://doi.org/10.1057/978-1-137-52436-2_2

Soyoof, A., Reynolds, B. L., Shadiev, R., & Vazquez-Calvo, B. (2024). A mixed-methods study of the incidental acquisition of foreign language vocabulary and healthcare knowledge through serious game play. *Computer Assisted Language Learning*, *37*(1–2), 27–60. https://doi.org/10.1080/09588221.2021.2021242

Stewart, J., Stoeckel, T., McLean, S., Nation, P., & Pinchbeck, G. G. (2021). What the research shows about written receptive vocabulary testing: A reply to webb. *Studies in Second Language Acquisition*, *43*(2), 462–471. https://doi.org/10.1017/S0272263121000437

Su, F., & Zou, D. (2022). Technology-enhanced collaborative language learning: Theoretical foundations, technologies, and implications. *Computer Assisted Language Learning*, *35*(8), 1754–1788. https://doi.org/10.1080/09588221.2020.1831545

Sundqvist, P. (2019). Commercial-off-the-shelf games in the digital wild and L2 learner vocabulary. *Language Learning & Technology*, *23*(1), 87–113. https://doi.org/10.10125/44674

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5

Tan, W., Supian, N., & Cheah, K. S. (2022). Game-based learning in improving English vocabulary and detecting metacognitive awareness among English for specific purposes undergraduates. *2nd International Conference on Social Science, Humanities, Education and Society Development* (ICONS 2021), Tegal, Indonesia. https://doi.org/10.2991/assehr.k.220101.021

Teng, M. F. (2023). The effectiveness of multimedia input on vocabulary learning and retention. *Innovation in Language Learning and Teaching*, *17*(3), 738–754.

Teng, M. F., Mizumoto, A., & Takeuchi, O. (2024). Understanding growth mindset, self-regulated vocabulary learning, and vocabulary knowledge. *System*, *122*, 103255. https://doi.org/10.1016/j.system.2024.103255

Teng, M. F., & Zhang, D. (2023). The associations between working memory and the effects of multimedia input on L2 vocabulary learning. *International Review of Applied Linguistics in Language Teaching*, *61*(3), 1021–1049. https://doi.org/10.1515/iral-2021-0130

Urun, M. F., Aksoy, H., & Comez, R. (2017). Supporting foreign language vocabulary learning through Kinect-based gaming. *International Journal of Game-Based Learning (IJGBL)*, *7*(1), 20–35. https://doi.org/10.4018/IJGBL.2017010102

Vnucko, G., & Klimova, B. (2023). Exploring the potential of digital game-based vocabulary learning: A systematic review. *Systems*, *11*(2), Article 2. https://doi.org/10.3390/systems11020057

Woodward-Kron, R. (2008). More than just jargon – The nature and role of specialist language in learning disciplinary knowledge. *Journal of English for Academic Purposes*, *7*(4), 234–249. https://doi.org/10.1016/j.jeap.2008.10.004

Wright, T. S., & Peltier, M. R. (2016). Supports for vocabulary instruction in early language and literacy methods textbooks. *Reading & Writing Quarterly*, *32*(6), 527–549. https://doi.org/10.1080/10573569.2015.1031925

Wu, T.-T. (2018). Improving the effectiveness of English vocabulary review by integrating ARCS with mobile game-based learning. *Journal of Computer Assisted Learning*, *34*(3), 315–323. https://doi.org/10.1111/jcal.12244

Wu, Y., Hong, J.-C., & Huang, P. S.-B. (2020). Using a game-based mobile application to learn medical academic vocabulary: Learning effects and learners' perceptions. *Taiwan International ESP Journal*, *11*(1). https://doi.org/10.6706/TIESPJ.202008_11(1).0001

Xodabande, I., Atai, M. R., & Hashemi, M. R. (2024). Exploring the effectiveness of mobile assisted learning with digital flashcards in enhancing long-term retention of technical vocabulary among university students. *Journal of Computers in Education*. https://doi.org/10.1007/s40692-024-00347-6

Xu, Z., Chen, Z., Eutsler, L., Geng, Z., & Kogut, A. (2020). A scoping review of digital game-based technology on English language learning. *Educational Technology Research and Development*, *68*(3), 877–904. https://doi.org/10.1007/s11423-019-09702-2

Yang, S., Lodge, J. M., & Brooks, C. (2025). Capturing multi-layered regulated learning in collaboration. *Metacognition and Learning*, *20*(1), 2. https://doi.org/10.1007/s11409-024-09409-7

Yi-ju, Jong-Chao Hong, & Paul Sin-Bao Huang. (2020). Using a game-based mobile application to learn medical academic vocabulary: Learning effects and learners' perceptions. *Taiwan International ESP Journal*, *11*(1). https://doi.org/10.6706/TIESPJ.202008_11(1).0001

Yüksel, H. G., Mercanoğlu, H. G., & Yılmaz, M. B. (2020). Digital flashcards vs. Wordlists for learning technical vocabulary. *Computer Assisted Language Learning*, *35*(8), 2001–2017. https://doi.org/10.1080/09588221.2020.1854312

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*, 1–43. https://doi.org/10.1080/09588221.2024.2344555

Zhang, W. (2022). The role of technology-based education and teacher professional development in English as a foreign language classes. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.910315

Zou, B., Huang, L., Ma, W., & Qiu, Y. (2021). Evaluation of the effectiveness of EFL online teaching during the COVID-19 pandemic. *SAGE Open*, *11*(4), 21582440211054491. https://doi.org/10.1177/21582440211054491

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

Appendices

Appendix A

Reviewed publications of DGBL for technical vocabulary acquisition

No.	Sources	Target language/Context	Sample	Professional field of voc
1	Ali et al. (2014)	(English) Online vocabulary game	32 students in a Malaysia public university	Engineering and technolo
2	Neville (2015)	(German) 3D digital game	42 students in a university of Southeast United States	German recycling and w management systems
3	Lu & Chang (2016)	(English) Digital role-play game	80 students in northern Taiwan vocational high school	Different workplace env (restaurant owner for ex
4	Urun et al. (2017)	(English) Kinect-based game	52 students in a state university in Turkey	Military vocabulary
5	Lelardeux et al. (2017)	(English) Digital anagram-like games	9 students in a French university	Engineering vocabulary operations and machine
6	Ali & Muhammad (2019)	(English) Mobile app game	129 students in Malaysia university	Engineering vocabulary
7	Wu et al. (2020)	(English) Mobile app game (custom-designed)	30 students from a College of Medicine in Taiwan	Medical vocabulary

No.	Sources	Target language/Context	Sample	Professional field of voc
8	Dinçer & Dinçer (2021)	(English) Digital serious video game	30 students in a state university in Turkey	Aviation vocabulary
9	Bakhmat & Chaikovska (2021)	(English) Online game	96 students in a state university in Podolia	Engineering vocabulary
10	Kohnke & Ting (2021)	(English) Mobile app game (custom-designed)	16 students in a Hong Kong university.	Technical vocabulary in and clothing, design, bus engineering, tourism ma hotel management
11	Tan et al. (2022)	(English) Digital video game	65 students in a private university in Malaysia	Commercial vocabulary
12	Kaldarova et al. (2023)	(English) Mobile app game (custom-designed)	50 primary students in Kazakhstan	Computer science vocab
13	Jia et al. (2023)	(English) AR-based digital game, Quizlet, physical flash card and puzzle (custom-designed AR game)	90 university students in China	Digestive vocabulary
14	Soyoof et al. (2024)	(English) Digital serious video game	160 university students in Iran	Nursing vocabulary
15	Pitarch (2024)	(English) Digital serious video game	64 university students in Valencia	Engineering vocabulary
16	Hsu & Chen (2024)	(English) Online chatbot with digital flash card game	48 nursing students in Taiwan	Medical vocabulary
17	García-Pinar (2024)	(English) Online game	36 university students in Spain	Engineering vocabulary

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.89-112
DOI: https://dx.doi.org/10.24093/awej/vol16no3.5

Innovative Linguistic Strategies of EFL University Students' Performance in Grammar Teaching

Raghad Khalaf Ulaywi Sahaw AlDulaimi
Department of English Language, College of Education
Al-Mustansiriyah University, Baghdad, Iraq
Email: raghadkhalaf@uomustansiriyah.edu.iq

Received:06/09/2025 Accepted: 09/01/2025 Published: 09/20/2025

Abstract

Teaching grammatical strategies is crucial for English as a Foreign Language (EFL) learners to apply correct language structures effectively. This study investigates the influence of innovative grammar teaching strategies—specifically type error and token error approaches—on students' grammatical accuracy compared with traditional methods. Conducted at Imam Jafar Al-Sadiq University with a sample of 100 EFL students from Iraqi private colleges, the research employed a non-randomized control group posttest experimental design. The main aim is to examine how new grammatical strategies affect students' performance across diverse grammatical contexts. The significance of the study lies in providing evidence-based approaches that enhance grammar teaching practices, encourage effective feedback use, and improve student outcomes. The central research question is: How do type error and token error strategies influence the grammatical performance of EFL learners compared with traditional grammar instruction? Participants were divided into three groups: one taught through type error strategies, one through token error strategies, and a control group taught traditionally. A grammar test was used to measure performance across the groups. Findings indicate that innovative strategies, particularly type error approaches, significantly improve students' grammar production and awareness, highlighting their value as practical solutions for overcoming common difficulties in grammar learning. This study contributes to educational innovation in EFL contexts and establishes a foundation for future research on effective grammar teaching methodologies.

Keywords: Educational innovation, EFL instruction, grammar teaching, linguistic strategies, student performance

How to cite: AlDulaimi, R.K.U. (2025). Innovative Linguistic Strategies of EFL University Students' Performance in Grammar Teaching. *Arab World English Journal, 16*(3): .89-112. https://dx.doi.org/10.24093/awej/vol16no3.5

Introduction

Pedagogically, this study highlights the significant shift in the way English grammar strategies are taught and learned by EFL students. Although grammar has traditionally occupied a central place in language instruction, conventional approaches often emphasized rote memorization over meaningful use. With the reemergence of communicative language teaching, grammar is now recognized as a tool that supports accuracy, fluency, and overall communicative competence (Kellaghan & Greaney, 2001).

Three principles guide effective grammar instruction. First, grammar should reflect authentic language use and illustrate how structures create and negotiate meaning in context. Second, grammar must be integrated with the four language skills—speaking, listening, reading, and writing—rather than taught in isolation. Third, grammar learning should be embedded in culturally and academically meaningful tasks that promote both communicative competence and higher-order thinking skills (Raghad, 2012). These orientations move away from grammar as static knowledge and toward grammar as a resource for interaction, awareness, and problem-solving.

The challenge lies in the contradiction between two schools of thought. Advocates of the grammar-less communicative approach argue that grammar should be minimized, as errors self-correct through meaningful exposure. By contrast, the grammar-integrated communicative approach insists that explicit strategies are essential for improving accuracy and comprehension (Janadhvla, 2018). This study positions itself within the latter framework, while also responding to unresolved challenges in the literature: namely, how to provide learners with explicit grammar feedback without discouraging communicative practice.

To address this gap, this study introduces type error and token error strategies as innovative approaches. Type error strategies help learners identify recurring error patterns at the structural level, while token error strategies focus on individual instances of errors in performance. These approaches are grounded in grammar learning strategies theory, which emphasizes conscious noticing and corrective feedback as means of improving accuracy, and Krashen's comprehensible input hypothesis, which stresses that grammar instruction is most effective when learners can notice gaps in their language use while receiving meaningful input. By combining explicit focus on errors with communicative practice, type and token strategies aim to overcome limitations of both traditional grammar drills and purely communicative exposure.

This study therefore investigates the effect of type error and token error strategies on Iraqi EFL college students' grammar achievement. It hypothesizes that:

- There is no statistically significant difference in grammar performance between students taught through type error strategies and those taught traditionally.
- There is no statistically significant difference in grammar performance between students taught through token error strategies and those taught traditionally.

Literature Review

Educationally, teaching strategies occupy a pedagogic role in classroom teaching. They are one of the indispensable columns in the instructional system planned for EFL learners. Educators need to create a favorable learning environment for students by adopting modern teaching approaches that introduce new ideas, concepts, and strategies to meet both students' needs and lesson objectives (Ascione, 2024:33).Teacher in the classroom adopts such grammar strategies to remedy such grammatical errors committed by EFL learners and to ensure students with active

learning. In addition, educators design such grammatical strategies to develop recent perspectives for EFL learners to adopt 21st-century innovative teaching strategies (Dania, 2024, p. 56).

In literature, many innovative academics and theoreticians on grammar teaching strategies have fulfilled in ELT texts, like Syafryadin, (2022), Qingfang (2023), Pawlak (2023), Wang et al., (2022), and Mathur (2022) argued that grammar strategies occupy a significant reconstruction structural strategy as a futuristic direction in the classroom. Learners employ these strategies as a future orientation method to facilitate the application of multidisciplinary theory in the classroom. This research examines the impact of two instructional strategies, type-error and token–error, on teaching grammar to EFL learners, with the aim of enhancing their grammatical performance. The primary function is to create an individual-directed learner who can use grammar strategies fluently. Consequently, this research paper explains the role of grammatical strategies in developing successful learners with classroom grammar instructions.

Innovative Grammar Strategies

Pawlak (2024) confirms that English teachers ultimately prone to stimulate their students with a complete encyclopedia of grammar strategies through grammar teaching class (p. 174). They are actions and concepts that students intentionally use to make language learning and language use easier, more effective, more efficient, and more enjoyable. Different assumptions and ideas influence many grammar steps that vary based on the concept of grammatical structures. They are essential and key components of effective classroom management skills. According to Scott (2000), grammatical strategies highlight the importance of carefully studying grammatical principles to organize language elements properly and effectively (p. 63). According to Bayon (2015), the grammar strategy is controversial as it emphasizes exposing students to grammatical procedures in a meaningful and thorough manner so they may learn this strategy of the target language as naturally as possible (p. 77).

Grammar strategies clarify the steps, ideas, and structures for describing, analyzing and accessing the lesson in the class. Type- error and token- error are structural strategies used to identify and remedy linguistic problems made by students. They are essential for precision and appropriate language use. They serve to encourage students on their structural journey to competence. Possessing a solid understanding of grammatical structures is indicative of a well-educated person. Learners may comprehend sentence patterns by using grammatical rules. The primary focus of sentence patterns or structural elements is on type and token error strategies (Raghad, 2021, p. 7). Because of the fact that grammar strategies refer to procedures used by teachers to reinforce students' linguistic structures in language, the aim is to increase students' participation in daily lessons, as college students seek effective grammar strategies that enable them to explore new and flexible ideas, and structures (Kurtz, 2023: 77). Consequently, these strategies help identify and address weaknesses in grammatical structures. They consider innovative because they concentrate on successful and integrated development not only of grammar but also essential receptive and productive skills, like listening, speaking, reading, and writing. Grammar strategies are crucial for developing grammatically correct language forms when structuring language. To prepare students for communicative tasks, grammar strategies play a favorable role in language instruction. The study supports the assertion of Doff (2000) that students can convey structural meaning through constructing phrases, clauses, and sentences (Ur, 2003, p. 54).

Procedurally Alqahtani (2022) reveals the affective role of teaching strategies for EFL students (p. 27). These strategies provide them with the basic steps to acquire practical knowledge to enrich their thinking abilities. As a result, students will become innovator of grammatical strategies.

Type-Error Strategy Components

The only grammatical or structural errors that EFL learners produce when structuring the English language are the primary focus of this research. Several studies have sought to relate the use of type error strategy to L2 problems, specially subject-verb agreement, the use of definite articles, and time-tense mistakes. They are all elaborated in this section as type errors strategy components rather than the second-kind components of errors strategy (Elzamarabyan, 2020, p. 43).

Fotheringham (2024) indicates that students can improve their grammatical competence by using the type-error strategy, which is effective for successful grammar teaching (p. 34). It signifies as follows:

- Using authentic writing sentence patterns for everyday situations.
- Making learning process clear through anchor schemes and students assignments focused on structural concepts.
- Providing students with diverse presentations for different structural attitudes.
- Having students innovate nonverbal representation of structural circumstances (pictures, stories, gesticulations).
- Using related graphic affairs to motivate students in structural task.
- Encouraging participants to engage in various declarative, cognitive, and imaginative activities to expose them to diverse contexts, including the construction of grammatically correct sentences.
- Utilizing the vigor of digital apparatus for training and rehearsal.
- Teaching and modelling any grammar motif, think aloud to model grammar instruction.

Creating a successful learning atmosphere is when educators and students achieve more realization of structures or components of type error strategies that combine this strategy with lesson curriculum and classroom content. Students can exchange and participate in prior knowledge, broaden their structures by grammar extension, get innovative structures, create structures, build correct patterns, and achieve successful grammar strategy (Muthmainuah, 2022, p. 13). The components of this strategy are as follows:

Subject-Verb Agreement Strategy

Al-Jarrah et al. (2021) explain that the English language has a problem with subject- verb concord method, in contrast to several European tongues (p. 144). English language verbs explicitly specify for concord in the Present Tense and exclusively when the doer of the action is "third- person singular." This theme activates subject-verb concord in English appears plain. Yet, faults strategy is common in college students' work (Alahmadi, 2019, pp. 48-55). As is common in this research, grammatical errors in language chunks are denoted by an asterisk, as follows:

Subject-verb agreement errors strategy (Note: 1)

a. She *leave since the iron master's sister's bounteousness.
b. The hawker *come down the summon of the ironmaster.
c. *Mr. David *think of himself as a potent hidden person.

ISSN: 2229-9327

d. *when another people views the looted dibs.
e. people who *lives a palatial existence."

Simpson (2018) indicates that the agreement mistake in statements (1a–c) occurs when general Present Tense form of a verb uses with a third-person singular subject (pp. 77). The third-person singular Present Tense structure of a verb applies to no third-person singular subject in sentences (1d–1e), which constitutes the agreement error. Despite the differences, if a student gives wide directions, all the faults in (1) may fix.

Subject-verb agreement error – rectification Strategy (Note: 2)

"S" or "es " should be appended after to a verb when it is in the present tense and contains a third-person singular subject. The bare form of a verb should use if it is in the present tense and has another subject. The following merits are treating subject-verb concord mistakes as type errors. It allows the student the appropriate time to increase the repair strategy made at the location of one error to another comparable one, while saving the teacher some "correction time." These additions undoubtedly demand the student's time and attention, but they are worthwhile as they encourage autonomous learning (Yousif, 2023, p. 126).

Indefinite articles strategy

The definite "the" and the indefinite "a"/ "an" are the two types of articles used in English. The unnecessary use of ellipsis and the insertion of articles when these articles where they are not needed can easily interpret as strategic errors. This strategy encourages students to visualize and assess their responses. Type errors also include blunders that involving one indefinite article instead of the other. Therefore, it leads students to fundamental ideas and procedures for better future forecasts with appropriate teaching methods and active manipulation of learning tasks (Raghad, 2021: 9).

Errors of replacement including " a"/ " an" strategy (Note: 3)

a. *an elated day
b. *an unity of nations
c. *an ubiquitous reality
d. *an yester week story
e. *an European voyage
f. *a obedient person
g. *an secure circle

Pak (2016) illustrates that the sentences above show these errors strategies usually happen before words that start with the letters " (e) u" and "y." (3a-e). Students are more likely to employ an in these cases because a must use structurally or grammatically before words with vowel-initial letters (p. 62). However, the need is one of sound rather than spelling: For a word to use before it, a vowel sound must be the first letter (as opposed to a vowel letter). Moreover, Harb (2014:87) suggests that utterances that go after *an* in (3a-e) all include vowel letters at the beginning. However the foremost of the letters stands for the gliding sound [j], which is not a vowel sound (Padgett, 2008, p. 80). As a result, the indefinite article " a" must use before words that start with a consonant sound. Since (3f, g) sequentially uses *a* before vowel and *an* in front of a consonant, the error strategy is more visible.

If students give the following two-point lesson simply once, all the mistakes strategy in (3) may fix promptly.

Errors involving indefinite articles – rectification Strategy (Note: 4)

ISSN: 2229-9327

A. Prior utterances starting with vowel sounds," an " must be applied; otherwise, "a" ought to be applied.
b. The letters " (e) u" and "u" do not symbolize vowel sounds. So, "a" has to be used prior to words starting with the previous letters.

Similar to the previously mentioned faults in the subject-verb concord strategy, considering replacement errors involving indefinite articles as a type error strategy reduces the amount of "rectification time" required by the teacher. Providing students with time for correction enables them to apply the recommended solution to similar recurring errors, thereby promoting autonomous learning (Raghad, 2023, p. 725).

Time and Tense Strategy

Raghad (2020) suggests that the modal auxiliary verbs "will " and "shall " use to express future time in English (p. 12). However, English language considers to have only two main syntactic tenses: the past and the present.

Time-Related Issues (Note: 5)

a. *Last month, Tom and Julia are running to the seashore, they view an incident. [Past Time – Present Tense]
b. *Last month, I travel with my colleagues and I suggest taking a walk with them. [Past Time – Present Tense]
c. *Once upon a time, when there were no mobile devices, people spent good time with their loved ones. [Past Time – Present Tense]
d. *Was it a vacation this day for the tornado? [Present Time – Past]
e. *Eggplants were overpriced these days. [Present Time – Past Tense]

All the phrases found in Time -Tense difficulties have adverbial components that may identify the present or past tense of time. However, these sentences do not use the proper tense to convey time (Aslan, 2021, p. 55).

Chang (2021) draws on the following three-point strategies may use to fix issues with Tense and Time:

Time-tense problems: Reparation Strategy (Note: 6)

Tenses ought to rely on the markers of time:

- Past Tense with Past Time markers,
- Present Tense with Present time markers; as well as
- Future tense with future time markers.

Like the other types of error strategies listed above, the strategies outlined in the previous section need only to give once. It facilitates "correction time" for the teacher. To put what they just learned into practice, the students will then ask to apply them to further errors of the same kind. This study attempts to illustrate how grammar learning strategies support the development of grammatical performance among EFL college students, in the light of the positive results of previous research (Borkovska, 2024, p. 58). It focuses on EFL College students who often seek innovative skills by engaging with applying the grammatical learning approaches. A variety of structurally designed learning instruments, including tests, questionnaires, and grammar quizzes, use to promote students' grammatical performance. Through the structured use of grammar-reinforcing questions and insightful feedback, EFL college students encourage to apply grammatical rules both accurately and systematically. This can assist them independently recognize patterns and nuances related to subject connected with subject areas. These results stem

from relevant pedagogical requirements and the effectiveness of the instruments, both of which align with the actual needs outlined in the ELT college course curriculum.

Token- Error Strategy Components

According to Ur (1996), grammar token error strategy refers to a system of grammatical structures that govern how utterances are constructed from acceptable structural elements within a given language (p. 85). It sometimes draws on the structures and meanings that the statements own within the language as a whole. This strategy is considered one of the essential analytical competencies required for successful implementation of 21st century learning approaches. Additionally, Rivers focuses on viewing these structures in class as the framework language uses to function. She further notes that attempting to use language without applying token strategies is comparable to expecting a chicken to move without its bones-fundamentally impossible and structurally unsupported. She opposes using didactic explanations of grammatical rules to teach token strategy. Rivers teaches grammar strategies by giving students' exercises and tasks that enhance them to follow the rules, which helps them learn the structures, build performance memories, and incorporate the information into their syntactic networks (Barlow, 2019, p. 99).

Another relevant researcher who focuses on components of token strategies is Celce Murcia (1988, p. 90), argues that these teaching strategies ought to be rule-focused rather than meaning -focused. In addition, the study of Savignon (1991, p. 25), concentrates on the need for grammar and structure to construct grammar strategies. He explains that the components of token strategies are not particularly effective when learned as a decontextualized, sentence- level system in a second language rather than being taught as a goal unto itself, these strategies are tools and resources to use in comprehending and generating verbal and graphic discourse. Prepositional errors, irregular Past Tense endings, incorrect use of reflexive pronouns, and question tags- among other common issues- consider more appropriately classified as components of token error strategies components rather than type error strategies.

Irregular Past Tense

Fokert (2019) explains that suffix "d" (debated, dated, decided, fated, etc.), or "ed" (as in played, faded, demanded) often added to the verb stem to indicate the past tense in English (p. 164). However, several English verbs, including steal, tell, sink, know, think, fall, choose, and stick, grow into the past tense naturally without any suffixation. EFL students regularly make mistakes like those listed in the following paragraphs while dealing with irregular past tense formation (7).

Errors in irregular past tense formation (Note: 7)

a. *The hawker stealed twenty valuable things from the merchant's dwelling [correct form: stole].
b. *The thief stoled jewelry of his neighbor and ran away [correct form: stole].
c. *The boat drowned in the sea earlier; it had each aid [correct form: drew].
d. *Sir. Bob teached me at the college since 2005 earlier pensioning. [Correct form: taught]
e. *When I was a child, I stucked imprints on a memoir provided with my mother [correct form: stuck].

Zarrinabadi (2023) clarifies that in (7a, c, and d) the forms of the verbs typically appear wrongly in the standard Past Tense structure (p, 347). The regular Past Tense finishing "ed" is excessively appended to the irregular past forms of verbs in (7b, e). Because the verbs in issue have varied Past Tense forms, as a token strategy, every fault in the statements in (7) ought to be

addressed individually. Therefore, token errors rather than type errors should use to tackle problems in irregular Past Tense creation. Raghad (2023) informs that even among verbs that produce past tenses erratically, there are other categories. For example, changing the letter "I" to the letter "a" makes the verbs drink, sink, etc (p. 725) form the past tense. In whatever way, the whole verb stems ending in "ing" or "ink" go through Past Tense evolution in that method, as shown by the examples of *bring* (Past: brought), *catch* (Past: caught), and *teach* (Past: taught). As a result, it is advised against using class references when correcting irregular past tense creation errors.

Prepositions (grammatical strategy)

Tetreault and Chodorow (2008) discuss that for researchers who are interested in investigating the structural or grammatical strategies made by EFL students, a preposition offers a new arena (p. 67). It is common to see errors in students' grammar writing that include using one preposition instead of another, such as those highlighted in (8).

Wrong prepositions (Note: 8)

a. *for one such time, Tom &Julie had the opportunity to see their relatives [right preposition: on].
b. *In vacations, they moved to an excellent place for bus [right preposition: with].
c. *Tania knocks the cat hard by the stick [right preposition: with].
d. *He went back to the house in foot, whereas some have taken the metro [right preposition: on].
e. *Reem and Diana have been singing since ten years [Right preposition: for].

Noor (2024) demonstrates that it is obvious that each wrong preposition in (8) needs to substitute for the right one (p.95). When a student can write the same preposition in many contexts, it becomes difficult to address the error with a single corrective lesson in these circumstances. Such an extension might produce unfavorable outcomes, causing the student to use a preposition that is inappropriate for the context in place of one that is missing. Let us say a preceptor enters the fault in (8a) and specifies the preposition "for" has to be replaced with "on". In sentences where "at" is already the correct preposition, such as "Their mates will not be at the function," the students may substitute 'on' for 'at." Therefore, incorrect prepositions in context are processed and corrected as token strategies rather than type issues. Prepositions that frequently use can also state in this way, as in the examples at the bottom (Bolliger, 2018, p. 59).

Abundant prepositions (Note: 9)

a. *The teacher discussed about the positive aspects of communication devices with the students in class.
b. *Students ought to prepare essential things for the journey.
c. *Always people have the confidence to faith in their abilities.

The preceptor must also expressly point out any superfluous prepositions in (9), just like they did for the wrong ones in (8). If not, the student might extrapolate from the instances in (9) to other circumstances where prepositions are particularly important. In such an instance, when the predicate in (9a) was "spoke" or "spoken," the preposition "about" required. Verb "discuss" did not come before *about* or any preposition; the teacher should elaborate to the student, because it means "to talk about." In conclusion, it is preferable to treat prepositional misuse as a token problem that opposes a typo (Raghad, 2023, p. 725).

Reflexive Pronouns (RPs) and Question Tags (QTs) *(Note: 10)*

Zhang (2010) explicates that pronouns and tags look like each other. These syntactic structures produce the grammatical elements that come before them in a phrase (p. 99). For

instance, in the proper noun *Tania*, herself is the prior (10a). In phrase (10b), the verb *has* uses earlier, however the negative form *has not* in the tag questions corresponds to the verb *has* used prior in the sentence. *Instances (Note: 10)*

a. Tania does not remember herself when she sings a romantic song.
b. She has succeeded the race, has not she?

Faults as in (11) and (12), occur in the production of EFL participants, despite the relevant association between reflexive pronouns (RPs) and question tags (QTs), a relationship that may seem naive.

Faults in the application of RPs (Note: 11)

a. *Tom herself lent me the box [right RP: himself]
b. *Cows himself were unclean organisms [right RP: themselves]
c. *Rana yourself gave the cattle grass [right RP: herself].

Students of English language specifically prefer the word "*isn't it*," and it uses even when the part of the sentence that comes before the tag does not include *is*, or a *noun* that substitutes *it* (12b-d) (Hashemi & Daneshfar, 2018, p. 90).

Faults in the application of QTs (Note: 12)

- *They were not traveling abroad years ago, did they? [proper QTs: were they]
- *The relatives can buy the palace, is not it? [proper QTs: cannot they]
- *John would have visited his aunt, is not it? [proper QTs would not he]
- *We write the introduction this day, is not it? [proper QTs: do not we]

Problems with the application of reflexive pronouns and question tags frequently refer to as token mistakes since a pronoun or question tag that is true to such cases probably might not be true to other cases. Specifically, the preceptor must always see the predecessor of a reflexive pronoun and, its equivalent in any question tag, to apply the correct structure of the final mentioned to the student (Litharge et al., 2020, p. 101).

Related Research

Through surveying numerous literary studies that have tackled some aspects of the problem of grammar teaching strategies for EFL students, Muhammad et al. study (2023), grammar teaching strategies to build EFL learners' abilities in grammar lessons, found out that grammar strategies concentrate on creating dynamic and positive learning environment to EFL learners to construct their abilities and personalities. The goal of their study is to concentrate on the structural approach that has created a successful grammatical learning environment for students in the class, and has currently stimulated students to support their recent communication tactics. The researchers applied a descriptive study. Findings revealed that teaching tactics conclude how words fit together to form structural sentences. They regard as one of the key components of linguistic structures that encourage students to produce language effectively and meaningfully as they structure their communication techniques. One effective strategy to improve students' communication is to discover and remedy grammatical errors (Nunan, 2002, p. 94).

The study of grammar teaching through digital stories was covered by Ozkaya &Dogan (2023). The goal of this qualitative study is to develop a model for teaching grammar through digital storytelling in grammar teaching course.4th grade Turkish teacher candidates involve in unstructured observation in the institution in the United States where the study was conducted. The results indicated that digital stories are structural tools that enhance innovative correction and

create grammatical questions such as tags. Also, it prepares digital stories for grammar topics, integrate these topics into grammar lesson plans applying different teaching materials.

The study of Borg (2024), an evaluation of grammar strategies, based on type error strategy. The purpose of this study is to evaluate the performance of grammar error strategies relying on human evaluation approach applied to grammatical structures. The study was conducted from October 2023 to February 2024 at several colleges in Japan. The researcher used a descriptive approach. He applied interviews and observation as techniques for collecting data. The analysis of full dataset shows that the grammar type error strategy plays a significant and effective role in teaching grammar lessons. The findings of the study indicate that type error strategy creates a positive learning environment for students in class, thus it becomes one of the basics of environmental education. In the classroom, students face structural errors when writing and structuring grammatical sentences, so this strategy has a diagnostic role to discover weaknesses in grammatical structural errors, and diagnosis appropriate solutions. Also, one of the core elements is to concentrate on the structural aspect of type error strategy efficiently.

The study of Alhasan (2024), investigates grammatical type error strategy among English language learners in their academic grammar writing instruction (p. 68). He applied a quasi-experimental posttest design to examine the impact of students' common type errors strategy they use in writing grammatical structures. (111) Grammatical structures answers collected from EFL college students. Also, an interview conducted. The research data was analyzed using Corder (1967) error analysis model and James (1998) framework of grammatical type errors. Moreover, Brown (2000) taxonomy adopted correcting scheme by classifying the types of grammatical errors. The ANOVA test showed that EFL students have structural problems with subject-verb agreement errors (rectification strategy), articles, and time tense errors. These are associated with the most frequent error kinds, which contain a specific error strategy. The study discovers that teachers can foster their students with correcting type error grammar strategies in the class to ensure successful pedagogical educational instructions.

Similarly, the study of Atkins (2023), identifies the effect of token error grammar strategies on EFL college students in grammar teaching lessons (p. 76). This study aims to stimulate students' ability to use grammatical structures on a daily basis through daily lectures. According to the structural perspective, token error strategy is to view as a framework that asks questions about the structure of grammatical items rather than the content of sentences in a discourse. The study indicates that historically, the students taught the strategy in the same time period of communicative language usage. Stern brought up the same issue, stating that the original focus of grammatical strategies is on linguistic or structural proficiency. Ninety college students took part in this study, both male and female. College students are divided into three groups. Three tests administered to all groups. The result of the study revealed that all the examinees in all groups achieved progress to a great extent and improvement.

One more significant researcher, Savignon study (1991), investigated the effect of token error strategy on ELT classes (p. 25). The study focused on the investigation of token error grammatical structures in ELT. Rather than taught as a goal unto itself, these grammatical structures are tools and resources used in comprehending and generating innovative methodologies. He depicts that token strategies components are not particularly useful when learned as a decontextualized sentence-level system in a second language. Prepositional errors, irregular Past Tense endings, the incorrect use of reflexive pronouns, and question tags - among the common faults strategies – consider more appropriately classified as components of token error

strategies rather than type error strategies. This research study revealed that token strategy has a significant and vital impact on the students' structures, specifically in ELT lessons.

Eckert study (2023), token-based strategies elicit the reward error and increase engagement of students explored essential academic success. The study aims to foster students to engage with token- based strategies through extrinsic rewards. In a quasi-experimental study design, 105 students participated in a classroom-based token strategy. All students gave a pretest. After 4 weeks of the experiment, they took the posttest. This study showed that token strategies excess students' engagement in grammar components, working by groups, and performing positive learning objectives with correct grammatical structures.

Methodology

This section outlines all the procedures applied to investigate how new grammatical strategies regarding type and token errors affect EFL students' performance in grammar teaching. It takes from the literature and theory. The outline includes the experimental design, population and sample, equalization of the experimental and control groups, research instrument, and teaching material for experimental study. A range of instructional activities and resources were used to apply the tow error-correction strategies, the type error strategy and the token error strategy. Students were regularly encouraged to reflect on their own mistakes and to negotiate proper forms in peer interactions through discussions. Sentence cards and comparison cards were included to assist students in recognizing and fixing lexical and grammatical mistakes in context, enabling them to see the distinctions between proper and improper forms. Students' ideas were arranged by using mind maps and innovative graphics to show the connections between words, categories and structures. These processes made it easier to fix type-level mistakes. Through collaborative practice opportunities offered by group activities and quizzes, students were able to identify token-level mistakes like spelling and word choice and reinforce proper forms through peer help and repetition. When combined, these methods made sure that the tow strategies were incorporated into dynamic and interesting teaching methods rather than being used separately.

Research Design

The study explored the effect of two grammar teaching strategies, namely type and token errors on the grammar performance of Iraqi EFL college students. To achieve the aim of the present study, the researcher applied Quasi-experimental design. A non-randomized control group post-test design used to investigate the hypotheses of the study. The whole experience achieved inside the classroom. Teaching content selected to contain correcting sentence errors in Stageberg's *An Introductory to English Grammar*. The experiment started on 1/10/2022 and ended on 1/2/2023. For five months, the teacher taught the control combination by traditional strategy. On the other hand, two empirical combinations received instructions to type and token errors strategies, which assessed by a grammar test to show the effect on the students' grammar teaching. Data from the present study collected.

Participants

In the first semester of the 2022 academic year, 850 college students, from the department of English, made up the study population, distributed on five Iraqi Colleges of Arts: Baghdad, Almustansiriyah, Babil, Wasit, and Kufa Universities. (100) Third-year college students from Imam Jafar Al Sadiq University were the study sample. The researcher herself taught the

participants for five months, from 1/10/2022 to 1/2/2023. The participants were divided into three groups: the first experimental group (section A) had (33) students who exposed to type errors strategy, the second experimental group (section B) had (33) students who exposed to token errors strategy, and the control group (section C) had (34) students who randomly chosen and exposed to traditional teaching strategy. The three groups took grammar teaching tests before and after the research experiment.

In this research study, the three groups equalized according to certain variables. These variables involve the ages of participants, the father and mother educational backgrounds of the participants, and the participants' performance on the pretest scores. The equalization of the three groups at the academic level of the student's father and mother calculated by Chi-Square values at 0.05 significance level. By applying ANOVA for age variable, it found out that the computed F-ratio is (0.422) which is lower than the tabulated F-ratio which is (3.021) at 2 and 116 degrees of freedom and 0.05 significance level. As a result, there is no statistically significant differences among the three groups at the age variable as shown in Table One.

Table 1. *ANOVA Statistics of Equalization among the three Groups in the Age Variable*

Variance	Sum of Squares	d.f	Mean Square	F-ratio	Level of Significance
Between Groups	274.351	2	157.121	0.422	0.05
Within Groups	457.222	116	245.654		
Total	731.573	118			

Applying ANOVA to pretest variable, it found out that the computed F-ratio (1.72) was lower than the tabulated F-ratio (2.02), at 2,116 degrees of freedom and 0.05 significance level. Thus, there is no statistically significant difference among the three groups in the pretest variable. It concludes that the three groups are equivalent depending on this variable as shown in Table Two.

Table 2. *The ANOVA Statistics of the Equalization among the three Groups in the pretest Variable*

Groups	Sum of Squares	d.f	Mean square	F –ratio Computed	Tabulated	Level of significance
Between groups	20.543	2		1.72	2.02	0.05
Within groups	541.324	116				
Total	561.867					

After finishing the pretest in the two experimental groups and the control group, and then applying the ANOVA formula for the three groups, the statistical treatment of mean scores for the three groups calculated. The mean score of the 1st experimental group is (40.14) whereas it is (42.41) for the 2nd experimental group, and (24) for the control one. The three groups' respective standard deviations are (3.312), (4.25), and (5.93).

Table 3. *Mean scores and standard deviations of three groups on grammar performance posttest*

Groups	Sample size	Mean	Standard Deviation
1st Exp.	33	40.14	3.312
2ndExp.	33	42.41	4.25
Control	34	24.12	5.93
Total			

0.05 significance level and with 2,116 degrees of freedom, the computed F- ratio which is (90.66) is higher than the tabulated F-ratio (3.055). It shows that there is a statistically significant difference among the three groups on grammar teaching posttest as shown in Table Four.

Table 4. *ANOVA results of the three groups of grammar performance posttest*

Source of variance	Sum of squares	d.f	Mean square	F-ratio Computed/tabulated	Level of significance
Between groups	454.22	2	2511.22	90.66 3.055	0.05
Within groups	263.31	116	21.23		
Total	717.53	118			

Research Instrument

The instrument for this research study was a pre-posttest on grammar teaching. It has provided to the three groups after fifteen weeks of intensive English grammar instructions. The researcher gave grammar teaching instructions in type error and token error strategies to the three groups then took a three-question test. Certain rubric standards for testing grammar were included such as grammar structures accuracy, range of grammar usage, error frequency and severity, sentence formation, suitability of grammar to context, consistency, and clarity and comprehensibility. The test was a multiple- choice exam that included specific grammatical structures and several comprehension passages. These syntactic areas outlined in the teaching curriculum for study during the first session. Grammar areas selected carefully to relate to certain grammar aspects and the subjects of the units in the prescribed textbook. The students had enough time to write their responses to the test questions. Sooner or later, to avoid the expected results of the other irrelevant factors in this study, the investigator herself determined to comprise a control combination of this research. It is essential to mention that the three combinations were, to a similar extent, influenced by the English settings. The achievement of these three combinations was much the same as the beginning stage previous to the experiment. The three question test involves as follows: Question one includes a reading passage, followed by choosing the most suitable answer from 10 multiple-choice questions (MCQs). The second question deals with correcting the verbs between brackets for 10 items. The third question involves in remedying grammatical sentences.

Posttest

The posttest administered following an intervening time of five months. Though the interval was long, notable exam impact was not seen. Participants could not logically influence by the pretest. Exam outcomes were numerically interpreted and calculated.

Questionnaire

Questionnaire used by the researcher to determine and assess the ideas and viewpoints by students with applying grammar teaching strategies. It consisted of items with responses ranging from always to never. To decide if there is statistically significant variance among the three combinations of grammar teaching strategies, Chi-square method has applied. The calculated (***x 2***) value is (8.10) is less than the tabulated (***x 2***) value that is (12.15) at eight grades of freedom and 0.05 level of significance. Thus, there are no numerically significant variations among the three combinations in grammar teaching strategies. This means that the three groups are equal in applying these strategies as shown in Table Five.

Table 5. *Chi-square for statistical Difference among the three groups in grammar teaching strategies*

Group	Always	Often	Sometimes,	Rarely,	Never	Total	Calculated	**X2**	Tabulated	d.f	Level of significance
1st observed	38	131	380	820	886	2055	8.10		12.15	8	0.05
Expected	55.5	140.5	305.7	810.9	890.70						
2nd observed	51	312	360	705	811	2239					
Expected	46	140.05	370.04	701.8	815.55						
Control group Observed	60	155	301	703	707						
Expected	59	100	390.70	725.33	680.60						
Total	149	598	1041	2228	2204	6220					

Weighted mean and weighted percentiles of the three groups in grammar teaching strategies have indicated that the above strategies have not achieved by the three groups. The three groups are equivalent, so there is no statistically significant differences among them as shown in Table Six.

Table 6. *Weighted mean, weighted percentiles for the three study groups in grammar strategies*

Groups	Weighted mean	Weighted percentiles
1st exp.	1.90	38.32
2nd exp.	1.80	38.24
CG	1.95	38.66

Validity

Richards and Schmidt (2002:197) elucidate that face validity signifies the extent to which an instrument measures what it intends to measure, depending on the personal opinions of the evaluator. A jury of specialists, containing eight university professors and ten supervisors of ELT, validates the face validity of the research by testing the pre-post examination. The jury members confirmed the face validity of the instrument and provided significant input, getting correct views about the usefulness and adequacy of the items. The posttest has confirmed the validity and

appropriateness of the third- year university participants in the English Language Department. The questions attain the exam aim specification and gain 100% agreement of the total jury members.

Pilot Study

The pilot study of the research tool aims to determine that the test constituents and instructions are obvious. It estimates time required to administer the tool, and calculate too. 40 Third-year students from Baghdad University, College of Education for Women, randomly participated in this research study on September 12, 2022. The outcomes of the pilot study indicated that there is no fundamental suspicion concerning the given grammatical items and the instructions, and that the time required to implement the grammar test ranges from twenty to twenty-five minutes. Adopting the analytical scoring scheme of the test papers, the outcomes are arranged from highest to lowest. The researcher then divides the marks into two groups: An upper group comprising 27% of the total, which gained the highest marks of the test, as well as a lower group containing an equal number of test papers from those who achieved the lowest marks. Every group consists of (18) points. It completed to confirm the degree of difficulty and the discriminating power of each test item. The discriminating power of grammar test items found out to range between (0.55& 0.90), as well as the difficulty range is between (0.25 & 0.75). Therefore, all grammar teaching test items have agreeable and appropriate discriminating power and difficulty levels, regarding to Mehrens and Lehmann (1991, p. 180).

Reliability of the Research Instrument

ASQ (2024:90) signifies that reliability suggests stability of test scores of the same testees under similar conditions. The dependable procedure relies on the test results of the subjects being fixed and consistent; but if not, the test is unreliable. To calculate the reliability of the current study, it used by Alpha Cronbach method. The outcomes showed that all the points of the questionnaire are reliable and appropriate. The consistency coefficient reveals to be (0.90). This result considers agreeable according to Mehrens and Lehmann (1991, p. 255). To calculate the reliability of the posttest, first, the Alpha Cronbach Method is used by the researcher, yielding a score of (0.88). Second, is by applying Pearson Formula. It discovered that Pearson Correlational Coefficient rates at (0.80). This value is reasonable by Mehrens and Lehmann (1991, p. 255). The test question items took from student grammar book titled "An Introductory to English Grammar" by Stageberg (2000).

Teaching Material for Experimental Study

After supplying the grammatical teaching material, the teacher selected a lesson in grammar teaching and suitable learning resources. The teacher brainstormed a discussion, asked students grammatical questions, and focused on various grammatical structures as well as on grammatical concepts and problems. Then, she responded to students' feedback after producing the selected lesson until the intended instructional objectives were achieved. Later, the teacher completed grammar tasks designed to assess and facilitate the students' ability to perceive grammatical structures that cause vagueness. In contrast with traditional structural methods of teaching grammar, the study inferred that specific grammatical structures have improved students' ability to comprehend grammar both functionally and syntactically. In addition, sentence cards, mind maps, comparison cards, group activities, and quizzes, along with graphic innovations, provided different grammatical learning materials. These grammar- supportive media stimulated

students to attain a clear understanding of structural elements. These learning aids helped them think critically and study deeply and accurately. These learning tools provided students with opportunities to think, structure, and produce content correctly. Teachers supported the students in structuring and sequencing the content and provided them with teaching resources that could assist in organizing information in class. During grammar lessons, students can track the structure and grade that the teacher has designed. Attentively, they become aware of and can notice every structure the teacher presented earlier while explaining the grammatical points in class. Students organized and structured class content in every session. Then, the teacher accumulated such structures and asked students to sequence them into grammatical patterns previously taught. This approaches helped them generate flexible and varied grammatical patterns by dividing the structures into their constituents, then combining or putting them together to form complete units by applying grammatical strategies. The teacher chose to document and store the results of this instructional method to teach other problematic grammatical areas in the future.

Final Administration of the Posttest

After confirming the posttest's validity and reliability, it was finally administered for the current research. On February 1, 2023, both the two experimental groups and the control group underwent a similar administration of the posttest on grammar teaching.

Analytical Scoring Scheme of the Posttest

The posttest analytical scoring scheme for the three questions of grammar teaching strategies is that one point is allotted to the correct item while (0) is allotted to the incorrect one, as shown below: In question 1, the teacher instructs the participants to read the comprehension passage carefully, and answer the items of MCQs that have some grammatical difficulty. Then, they select the most suitable choice. The total number for items of the first question is (10). Each item has (1) score for correct answer and (0) score for incorrect answer. The second question deals with (10) items of correcting the verbs between brackets. Each item has one score. The third question is about (10) items of correcting grammatical sentences, i.e. circle the part of the sentence that is ungrammatical. Each item has (1) score as shown in Table Seven

Table 7. *Specification of the achievement posttest*

Q No	Type	No. of Items	Category	Total score
1	MCQ	10	Objective	10
2	Correct verb between brackets	10	Objective	10
3	Correcting grammatical sentences	10	Objective	10
Total				30

- Question one (10 scores):10 MCQs
- Question two (10 scores):10 items of correcting the verbs between brackets
- Question three (10 scores):10 items of correcting grammatical sentences

ISSN: 2229-9327

Data Analysis

This section elucidates the data analysis attained from the participants' answers on the posttest achievement and how the outcomes concern the aim and hypotheses of this research study. The current study utilizes the Two Independent Samples ANOVA test to investigate the two hypotheses that there is no numerically notable distinction between grammar average variation of the participants who are employing type error procedure as well as those of the participants who are teaching relying on the conventional approach. The second null hypothesis states that there is no numerically notable distinction between the grammar average variation of the participants who are employing token error procedure and those of the participants who are teaching relying on the classical approach. The current study may enrich the data to fulfill its stated aim, which is to study the influence of new grammatical strategies regarding type and token errors in different grammatical settings and to examine how their use affects EFL students. The data collected from the ANOVA test for independent samples was obtained after participants' completed responses on the posttest achievement and the questionnaire. The data was then analyzed to determine the statistical differences among the two experimental groups and the control group.

Results

Results Belonged to Grammar Teaching Posttest

This section elucidates the data analysis, how do the outcomes associate with the aim as well as the hypothesis that began that there are no numerically notable distinctions among the students' groups who are employing type error strategy, token error strategy as well as the traditional approach in their mean scores in grammar teaching? Mean scores and standard deviations are counted for the three groups as indicated in Table Eight.

Table 8. *Mean scores and Standard Deviations of the three groups in the Grammar teaching post-test*

Grs.	N.	M.	S.D.
CG	33	26.31	2.60
EXP1	33	44.60	2.17
EXP2	34	40.50	2.25
Total	100	10.22	2.018

To reveal the value of the diversity among the mean marks of the three combinations in grammar, analysis of variance applies to point the place results are indicated in Table Nine for teaching grammar.

Table 9. *Analysis of variance for revealing the value of diversity among the Mean scores in Grammar*

source of variance	Sum of squares	Mean square	Calculated F-ratio	Tabulated F-ratio	d.f	Level of value
between combinations	90.22	42.333	5.22	3.05	2.97	0.05
within combinations	811.15	7.17				
Total	901.37	49.503				

Table 9 indicates that the calculated F-ratio that is 5.22 is greater than the tabulated F-ratio that is 3.05 at a level of value (0.05) as well as (2) & (97) degrees of freedom. That shows a numerically important diversity among the three combinations on the grammar posttest.

To distinguish the source of variance among the three combinations in the grammar posttest. Scheffe′ test uses for post contrasts between type error and token error strategies. Results related to grammar are indicated in Table Ten.

Table10. *Scheffe' Value for the contrasting among the three groups in the grammar posttest*

Group		Mean variation	Crucial Scheffe
1st exp.	-2nd exp.	2.1	3.5
	-CG	16.04	4.055
2nd exp.	-1st exp.	2.1	3.5
	-CG	13.22	4.055
CG	-1st exp.	16.04	4.055
	-2nd exp.	13.22	4.055

Table 10 clarifies the following:

-The calculated Scheffe merit for the distinction between the 1st experimental group and the 2nd experimental group is 2.1.This value is less than the crucial Scheffe value of 3.5 at the 0.05 significance level. That shows no analytically important diversity between the 1st and 2nd empirical groups.

-The calculated Scheffe value for the distinction between the 1st empirical and control combinations is 16.04, which is higher than the crucial Scheffe value of 4.055 at the 0.05 significance level. That shows a significant numerical differences, favoring the 1st empirical combination, which taught by using a type error strategy.

-The calculated Scheffe static for the variance between the mean score of the 2nd empirical combination and the control combination is 13.22, which is higher than the crucial Scheffe value of 4.055 at the 0.05 significance level. That shows there is analytically significant variation in favor of the 2nd empirical combination, which taught by token error strategy.

-Table 9 shows the comparing results between the first and second empirical combinations. The calculated Scheffe value for the distinction between the average marks of the 1st empirical combination with the control combination is 16.04, which is the same for the second empirical combination with the conventional combination 13.22. At the same time Scheffe crucial merit of 3.5 for first and second empirical groups found to be significant at the 0.05 significance level. Therefore, the calculated Scheffe merit for the distinction between the average marks of the first empirical combination is greater than those of the second empirical group. The current study concludes that there is a numerically substantial variation for the benefit of the first empirical combination, which studied by the type error strategy. This justifies the impact of type error strategy on EFL university students' performance in the grammar teaching posttest as show in Table Eleven.

Table 11. *Scheffe' Merit for the comparison between 1st & 2nd empirical combinations*

Groups	Average distinction	Scheffe crucial merit
1st exp.	16.04	4.055
2nd exp.	13.22	3.5

When comparing the 1st and the2nd empirical combinations with the third, the highest average mark is for the 1st empirical combination 44.60.In contrast, the second empirical combination has average of 40.05, and the control combination has 26.31. The average difference between first and second empirical combinations and the control combination is 14.22.The crucial Scheffe merit is 4.055, while the mean difference is14.22, which is greater than the essential Scheffe merit. The current study findings indicate that a numerically substantial variation in favor of the 1st empirical combination that taught by the type error strategy. Based on the ANOVA test, the first experimental group's performance (M=44.60, S .D=2.17) was significantly higher than the other two groups. It concluded that the first strategy is successful in mastery and accuracy in grammar teaching in class as shown in Figure one.

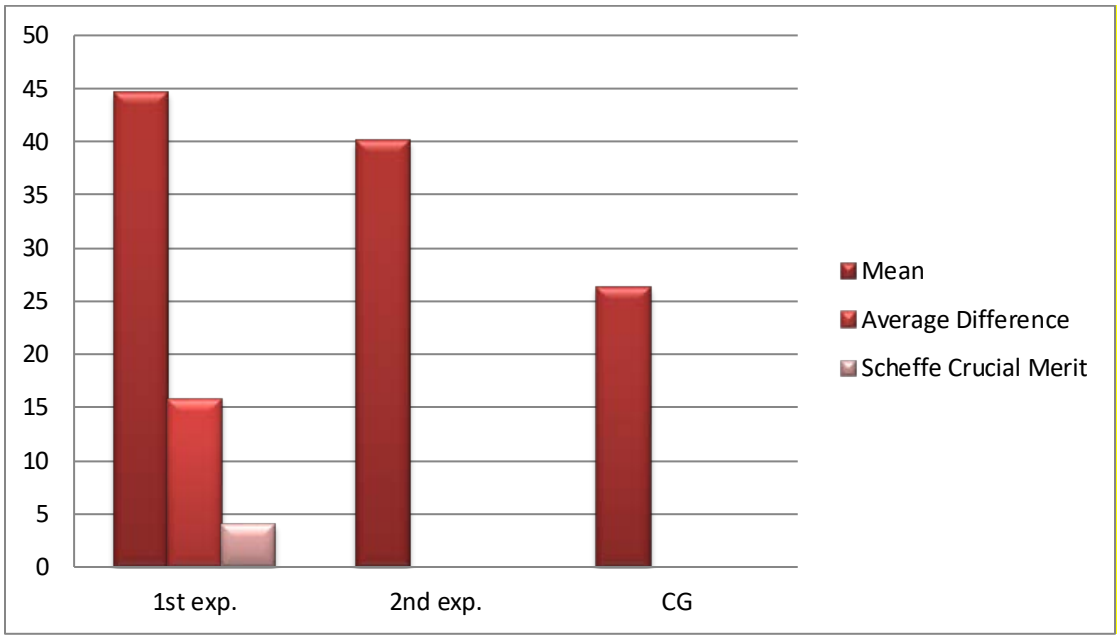


Figure 1. Bar Chart of compound comparisons among combinations by Scheffe Test

Discussion

Literary studies highlight meta analyses and innovative classroom studies. These studies concentrate on grammar accuracy and observe grammatical patterns to attain meaningful improvement. It has been illustrated that type-error strategies in EFL classes engage students to plan, organize, and create structural patterns that stimulate communicative outputs and enhance learner autonomy creation (AlTameemi and Al-Tamimi, 2025, p. 439). Through this research study, the researcher intended to highlight some significant points regarding the strategy used. Findings show that type error strategies focused on greater grammar accuracy and transfer than the second token strategies. Students achieved higher comprehension and utility for typo questionnaire practices. Students with high motivation and independence attain better results under the type strategy, whereas those with low motivation availed less outright. Therefore, motivated students exploit profound conversion. They tend to use AI grammatical tools as valid applications, they do not constantly clarify structures or compare L1/L2 grammatical structures, which reduces transfer. The study indicated that feedback with no metalinguistic clarification proposes native remedy, not rule popularization. Token practices rarely face bad evidence in texts; thus, students learn by heart the correct appearance with no control over previous reflexive concord. Typo strategies make these chains clear (Philipose, 2022, p. 22). It is implicated that token tasks reprocess similar lexical structures, encouraging learning material. The research outcomes reveal how type and token error strategies influence participants' grammar teaching. Examining the posttest outcomes of the students' grammar performance among the three groups, it discovered that the first experimental group's achievement scores in applying the type error strategy were statistically significantly higher than those of the second experimental group and the control group, with a significance level of 0.05. In agreement with the studies of Borg (2024), Alhasan (2024), and Muhammad et al. (2023), the type error strategy has an immense and vital role in improving EFL college students' level in correcting grammatical errors. Also this strategy offers modern and

significant methods of creative grammatical correction. Gradually, it starts from a simple to a complex level of correction. Therefore, the examiner concentrates primarily on structural errors to enrich examinees' academic performance. This innovative strategy fosters the participants' ability to correct structural errors fussily. As a result, students are encouraged by their teacher to avoid committing grammatical errors and have the opportunity to build good structural phrases, and sentences ideally. They feel no pressure and relax while constructing their structural items successfully.

Conclusion

According to the findings of the present study, innovative grammar teaching strategies—particularly the type and token error strategies—proved useful in addressing structural problems among Iraqi EFL college students. A central conclusion is that the type error strategy was statistically more effective than both the token strategy and traditional methods, as shown by posttest results at the 0.05 significance level. This indicates that type error strategies lead to greater grammatical accuracy, comprehension, and error correction.

The study also revealed that students' motivation and autonomy play an important role in the effectiveness of the strategies. Highly motivated and independent students benefited most from type error practices, while less motivated learners showed limited improvement. Furthermore, feedback without excessive metalinguistic explanation encouraged students to correct errors naturally, reducing negative transfer from L1 structures.

Token strategies, in contrast, underperformed due to L1 interference, low variability, and limited feedback, leading to memorization without deeper control over grammatical patterns. Nevertheless, token practices still offered opportunities to recycle lexical structures and reinforce basic forms.

Overall, the type error strategy stands out as a pivotal approach for ELT grammar lessons, since it fosters structural accuracy, supports communicative use of grammar, and encourages students to build correct phrases and sentences with confidence. By gradually moving from simple to complex correction, this strategy enables students to internalize grammatical structures effectively. Therefore, the integration of type error strategies into college curricula is highly recommended, as they not only improve grammar achievement but also enhance learner autonomy, motivation, and communicative competence.

Recommendations and Suggestions

Based on the outcomes and conclusions, the following suggestions are significant: When students are not positively taking part in the grammar correction process, some obstacles that would face instructors in grammar teaching process. Teachers have to clear the way for students' selection new methods focusing on how to correct structural errors by teaching new methods of grammar teaching strategies. The proposed type error strategy is studied as a learner-centered method inside the class to stimulate intellectual grammatical structures. Similar to an established approach, students have to choose their method of learning grammar to increase their engagement. Instructors still play a vital role in helping students to attain their learning objectives by acting as effective facilitator of content structures. Each pace is stimulated, pictured, and planned by the teacher. Teachers have a vital role to orient the students if they have not the stimulation to learn or cannot organize the goals of the suggested learning scheme. To increase the effectiveness of students' grammatical learning strategies, teachers must conduct a complete investigation of

curriculum content, the educational environment, and the learners' readiness. Upon submitting recommendations for recent studies, it is important that researchers examine the proposed grammar teaching strategies to learners in various environments, such as colleges and different academic sections. To assist students in enriching their grammar correction accuracy in different aspects, grammar teaching strategies must focus on these correction techniques. Exploring how grammar strategies influence students' proficiency in improving and inventing other grammar strategies, such as self-correction of structural errors, is the main concern of this study. The current study is concerned with how grammar strategies affect EFL students' stimulation and standard of participation during teaching grammar.

Funding: This research paper has not funded
Acknowledgement: It is not applicable
Conflicts of interest: The author declares no conflict of interest
Authenticity: This manuscript is an original work
Artificial Intelligence Statement: AI and AI-assisted technologies were not used**.**

About the Author:
Raghad Khalaf Ulaywi is an Assistant Professor with a doctoral degree in Methods of ELT. She has more than 15 years of experience teaching at Imam Jafar Al-Sadiq and Al-Mustansiriyah University /English Language departments. Her interests include many topics, such as recognition and production skills, strategies and techniques in teaching and learning, English language teaching approaches, psycholinguistics, and creative and critical thinking skills.
ORCID ID: https://orcid.org/0009-0009-9774-6493

References

Alahmadi, N. (2019). A study of grammatical errors of subject-verb agreement in writing. *International Journal* of *English Language and Linguistics Research*, *7(*6), 48-55.

Alhasan, A. (2024). Obstacles of applying alternative strategies from the point of view of the first cycle teachers in the schools of Hama city and ways of solving them. *Alustath Journal for Human and Social Sciences 63*(2), 68. https://doi.org/10.36473/x3md1066.

Al-Jarrah, M., Rayyan, M., Al-Shawashreh, E., & Zuraikat, M. J. (2021). Subject-verb agreement in Arabic, Spanish, and English: A typological view. *Dirasat Journal for Human and Social Sciences*, *47*(1), 144.

Alqahtani, A. (2022).Review and analysis of theories underlying grammar teaching methodologies. *Arab World English Journal (AWEJ), 13(*4*),* 82& 86. https://dx.doi.org/10.24093/awej/vol13no 4.6

AlTameemi, M., & Al-Tamimi, E. (2025).Visual learning environment and analytical thinking skills. *Arab World English Journal (AWEJ), 16(*1*),* 439. https://dx.doi.org/10.24093/awej/vol16no1.27.

Ascione, L. (2024). *Current trends in education. E school media*, Washingtonian Boulevard: Suite.

Aslan, A. (2021). *Relevance of inquiry-based learning in basic reading skills exercises for improving student learning*. San Francisco: Pearson Education Longman.

ASQ. (2024). *Learning about quality*. ASQ Statistical Quality Press.

Atkins, J. (2023). *Skills development methodology*. Addis Ababa University. Addis Ababa Printing Press.

Barlow. (2019). *Detecting grammatical change in a previously undescribed endangered language*. University of Helsinki: Helsinki.

Bayou, Y. (2015). *Grammar learning strategy use of grade 11 students at Medhanealem preparatory school: Gender in focus*. Ethiopia: Addis Ababa University.

Bolliger, D. (2018). *Engagement matters: Student perceptions on the importance of engagement Strategies in the online learning environment*. Carolina: Charlotte University of Wyoming Dress.

Borg, S. (2024). Teacher cognition in grammar teaching: A literature review. *Journal of Language Awareness Strategies, 12*(6).

Borkovska, I. (2024). Integration of AI into the distance learning environment: Enhancing soft skills. *Arab World English Journal, 16* (Special Issue on Chat GPT), 58. http://dx.doi.org/10.24093/awej/ChatGPT.2.

Celce-Murcia, M. (1988). *Techniques and resources in teaching grammar*. New York: Oxford University Press.

Chang, C. (2021). *Six trends for the new future of English grammar*. British Council Worldwide.

Dania, V. (2024). Trends in education technology that will have a major impact in the future. *The Digital Textbook Platform, Education &Technology*, 2(4), 56.

Doff, A. (2000). *Teach English: A training course for teachers*. Cambridge: Cambridge University Press.

Eckert, M. (2023). How do token-based strategies elicit reward error and increase student engagement? *Educational Psychology Journal*, 14(2). http://doi.org/10.3389/fpsyg.2023.1077406

Elzmarabyan, L. (2020). *Eight amazing new trends in English language learning*. S.A: Haber.

Fokert,G. (2019). *Grammar writing and language contact*. University of Helsinki: Helsinki.

Fotheringham, J. (2024).*The best strategies for language learning*. Lenox Education Limited.

Harb, M. *(*2014*)*. A closer look at the English article system: Internal and external dimensions of difficulty revisited. *International Journal of Applied Linguistics, 6(*4), 87.

Hashemi, A., & Daneshfar, S. (2018). The impact of different teaching strategies on teaching grammar to college students. *Theory and Practice in Language Studies, 8*(3).

Janadhvla, D. (2018). *Innovative methods for teaching English at the secondary level*. Lincoln: Blackwell. Retrieved from http://galleryteachers.com.

Kellaghan, T., & Greaney, V. (2001). *Using assessment to improve the quality of education.* UNESCO. https://unesdoc.unesdoc.unesco.org/ark:/48223/pf0000126231.

Kurtz, C. (2023). *The most effective language learning strategies*. Cleve don: Pergamum Press.

Litharge, Albert. , & Shankar, K. (2020). Data weighted training strategies for grammatical error correction. *Journal* of *Applied Linguistics, 21*(4), 101.

Mathur, R. (2022). Innovations in English language teaching: Strategies, challenges and future directions. *International Journal of Social Impact, 7(*4). DOI: 10:25215/2455/0702015.www.ijsi.in

Mehrens, W., & Lehmann, I.J. (1991). *Measurement and evaluation in education and psychology*. New York: Holt, Rinehart, and Winston.

Mesham, Nagata. , Wanaka, T., & Masui, F. (2023). Detecting article errors based on the mass count distinction. *Journal of Computational Linguistics*, 1(3), *26*.

Muhammad, I., Darmayanti, R., & Arif, V. (2023). Discovery learning research in mathematics learning: A bibliometric review. *Alustath Journal for Human & Social Sciences*, *63*(3), *57*.

Muthmainuah, Y. (2022).Engaging learners' book. *World Journal of English Language*, 1(4), *13*.

Noor, Abdul Sahib. (2024). Assessing Iraqi ESL fourth year college of physical education and sport sciences students' writing anxiety. *Alustath Journal for Human and Social Sciences*, *3(1*), 95.DOI:http://doi.org/10.36473/7sv84632.

Nunan, D. (2002). Learner strategy training in the classroom: An action research study. *TESOL Journal, 2(*1*),* 94.

Ozkaya, P, & Dogan,B. Grammar Teaching through Digital Stories: A Study on Developing a Module. *International Technology and Education Journal, 7(*2), 01-15. http://itejournal.com/.

Padgett, J. (2008). Glides, vowels, and features. *Lingua*, *8*(4), 80.

Pak, M. (2016). How allomorphic is English article allomorph? Glossy: *Journal* of *General Linguistics*, *1*(3), 62. DOI: https://doi.org/10.5334/gjgl.62.

Pawlak, M. (2023).Grammar learning strategies: Towards a pedagogical intervention. *Language teaching research quarterly*, *3(*2*),* 17. https://doi.org/10.1177/1362168819876156.

Pawlak, M. (2024). *Exploring the interface between individual difference variables and the knowledge of second language grammar*. Springer.

Philipose, C. (2022).Type-errors and token-errors in grammar evaluation. *Journal* for *Research Scholar and Professionals of English Language Teaching, 1 (*3*),* 22.

Qingfang, X. (2023). An exploration of grammar teaching strategies based on new curriculum standards. *Frontiers in Science*, 2(1), 54 .DOI:10.54691/v3i7.5321

Raghad, U. (2012). The effect of using English culture awareness strategies on Iraqi EFL university students' reading and writing performance .*M. A. Thesis*. College of education Ibn- Rushd. University of Baghdad.

Raghad, U. (2020). Investigating EFL college students' level in recognition and production of English collocations in reading comprehension. *Journal* of Linguistics &*Translation Studies*, *33*(3), January, 25.

Raghad, U. (2021). Cognitive activation strategy as an innovative teaching trend to EFL students' in writing skill. *AL- Adab Journal*, *137*(1), *June*. DOI:10.31973/aj.v2i137.1645.

Raghad, U. (2021). Spur students to brainstorm ideas and deliver a presentation as classroom management strategies in essay writing. *Aliraqia University Journal*, *50*-2.

Raghad, U. (2023).The effect of vocabulary teaching strategies on linguistically diverse classroom reading comprehension. *Journal* of *Sustainable Studies, 5*(4), 725. www.joss-iq.org.

Richards, J., & Schmidt, R. (2002). *Longman dictionary of language teaching and applied linguistics*. New York: Pearson Education Ltd.

Savignon, S. (1991). Communicative language teaching. *TESOL Quarterly, 1(*2*),* 25.

Scott. (2000). Explicit and implicit grammar teaching strategies. New empirical data. *The French Review, 3(*3*),* 63.

Simpson, A. (2018). *How to use songs in English Language classroom.* Cambridge: CUP.

Syafryadin, S. (2022). Actualizing learning English grammar strategy using new inventory for university students. *Parole Journal* of *Linguistics and Education,* 1(4), *25*. DOI:10.14710/parole.v12i1.25

Tetreault, J., & Chodorow, M. (2008). *Ups and downs of preposition error detection in ESL writing*. Manchester: United Kingdom.

Ur, P. (1996). *A course in language teaching*. Cambridge: Cambridge University Press.

Ur, P. (2003). *Grammar practice activities: A practice guide for teachers*. Cambridge: Cambridge University Press.

Wang, H., Li, D., & Chen, X. (2022).Token-based complication error categorization and its applications. *Journal of software*: *Evolution and Process*, *35(*2*).*

Yousif, A. (2023). The perception of faculty members and EFL learners to proficient instructors. *World Journal of English* Language, *13(*5).

Zarrinabadi, N. (2023). The links between grammar learning strategies and language mindsets among L2 and L3 learners. *International Journal of Multilingualism*, *2(*2*)*, 347.*http*://doi.org/10.1080/14790718.2023.1871356.

Zhang, Q. (2010). A study of Chinese learning of English tag questions. *Journal* of *Language Teaching and Research*, *1(*5), 99.

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.113-128
DOI: https://dx.doi.org/10.24093/awej/vol16no3.6

First Language Transfer Among Chinese High School Students in the English Writing Process

YANQIN MA
Language Academy, Faculty of Social Sciences and Humanities, Universiti Technologi Malaysia, 81310 Skudai, Johor Bahru, Malaysia
Email: mayanqin@graduate.utm.my

MAHANI STAPA
Language Academy, Faculty of Social Sciences and Humanities, Universiti Technologi Malaysia, 81310 Skudai, Johor Bahru, Malaysia.
Email: m-mahani@utm.my
Corresponding Author: mayanqin@graduate.utm.my

Received: 06/04/2025 Accepted: 08/20/2025 Published:09/20/2025

Aabstract
English learning surges around the world, while foreign language instruction in writing falls behind. The teaching of English emphasizes language input, yet writing skills lag. Such condition accentuates the need to develop English writing skills among English as a Foreign Language students, considering the intrinsic link between second language writing and their first language. Therefore, this study investigates the transfer of the first language during second language writing. It tries to answer: 1) When do learners transfer L1 the most in their English writings? 2) What are the reasons for using L1 at different writing stages? 3)How do learners perceive the influence of L1 use in L2 writing? To address these questions, this study adopted a sequential explanatory mixed-methods design, which involved quantitative questionnaires followed by qualitative interviews. Sixty students participated in answering the questionnaire. Subsequently, four respondents participated in qualitative interviews. The results revealed that Chinese students tended to rely on their first language during the pre-writing stages the most. They mainly used their first language for ideation and lexical searching. In addition, students consider L1 as a cognitive scaffold as it helps them generate ideas, search for lexical items, and reduce cognitive load in second language writing. However, they are aware of the interference caused by literal translation. This study deepens our knowledge of how learners of English as a Foreign Language utilize their native language during the process of writing in a second language. It offers important insights for improving English writing skills. The results provide educators with useful guidance to develop writing instruction that better addresses the specific needs of second language learners.
Keywords: Cognitive load, first language transfer, mixed-methods research, second language writing, writing instruction, writing proficiency, Chinese learners

How to cite: MA, Y., & STAPA, M. (2025). First Language Transfer Among Chinese High School Students in the English Writing Process. *Arab World English Journal, 16*(3): 113-128. https://dx.doi.org/10.24093/awej/vol16no3.6

Introduction

With China's economic rise, the demand for English learning has never been stronger. Foreign language teaching in China has its peculiarities. English is a foreign language in China. Chinese students only learn English for 45 minutes in the classroom, and even teachers only focus on grammatical knowledge (Chen & Li, 2016; Sun & Lan, 2023), so there is little communication in English in the classroom, let alone outside school. In addition, there is a vast difference between English and Chinese, which means that Chinese students face substantial difficulties learning English. Besides, Chinese English learners emphasize input over output, scoring higher in listening and reading but lower in writing and speaking (Moghadam et al., 2012; Zhu et al., 2022). Therefore, studies should pay more effort to English writing among Chinese learners.

Research on second language (L2) writing has two major strands: one focusing on the written product and the other on the writing process. Since the 1980s, there has been a notable shift from a product-oriented to a process-oriented perspective, emphasizing the cognitive and strategic dimensions of writing (Flower & Hayes, 1981). Initial research in this area focused on the writing processes in learners' first language (L1) (Abas & Abd Aziz, 2016; Abdi Tabari, 2022; Manchón et al., 2000; Sasaki, 2002), establishing a foundation for later investigations into second language (L2) writing. Subsequent studies identified both commonalities and distinctions between L1 and L2 writing processes.

An increasing number of studies indicate that writing in a second language (L2) presents greater difficulties than writing in a first language (L1) because it requires the simultaneous handling of two language systems (Cheong et al., 2021; Michel et al., 2020; Zhu, 2000; Zhu et al., 2022). In early theoretical frameworks, especially those influenced by behaviorism, the use of L1 during L2 writing was seen as interference and thus discouraged in instructional settings. Starting in the 1980s, the emergence of cognitive writing theories led researchers to reevaluate the function of the first language (L1), suggesting that it can positively impact second language (L2) writing by aiding in idea development, text organization, and planning strategies.

Despite this growing recognition, much of the existing research still tends to focus on either general cognitive strategies or broad comparisons between L1 and L2 writing. Less attention has been given to the specific influence of L1 on the writing process in L2, particularly how students employ L1 during different stages such as planning, drafting, and revising. Besides, research among high school students in English as a Foreign Language settings is scarce. Therefore, this study aims to fill this gap by investigating how learners draw on their L1 during the L2 writing process, and what role it plays in facilitating or hindering their writing performance. This research examines how Chinese high school students use their L1 at various stages of the English writing process and investigates the effects of this usage on their cognitive engagement. It focuses on identifying stage-specific patterns of L1 use and understanding their influence on L2 writing performance.

The study adapted Wang and Wen's (2002) model of L1 use in L2 composing, which outlines how Chinese learners engage with their L1 during planning, formulation, and revision. Their model further refines the L2 writing process into five cognitive stages: task-examining, idea-generating, idea-organizing, text-generating, and process-controlling. Drawing on this framework, the present study adapts these stages to examine when and why Chinese high school students use their L1 while writing in English, and how this use shapes their composing behavior and performance.

The findings are expected to provide valuable insights into learners' cognitive processes and offer implications for more effective writing instruction. By aligning pedagogical strategies with students' natural composing behaviors, particularly their use of L1, this research aims to contribute to the development of more supportive and cognitively informed EFL writing pedagogy.

The study addresses the following research questions:

1. When do learners transfer L1 the most in their English writings?
2. What are the reasons for using L1 at different writing stages?
3. How do learners perceive the influence of L1 use in L2 writing?

Literature Review

Cognitive Model of L2 writing

The exploration of L2 writing, dating back to the 1950s, has expanded since the 1990s, evolving into a robust, multifaceted field (Silva, 1993; Sun & Lan, 2023). L2 writing refers to the examination and understanding of writing in a language learned after one's native tongue. It covers a wide range of elements, including writing techniques, strategies, and cognitive processes used by individuals who are learning and using a second language for written expression (Nystrand, 2023). Researchers have scrutinized trends in L2 writing research to gain insights into the intricacies of the field. Literature reviews have mapped the historical evolution of L2 writing (Silva, 1993) and delved into pedagogical approaches and implications (Grabe & Kaplan, 2014). It shifted from focusing on the writing product to the writing process. Hence, a theory of the L2 writing process becomes essential.

This study adapts Wang and Wen's (2002) L2 composing model. Wang and Wen's (2002) L2 composing model extends Flower and Hayes' (1981) cognitive composing process model to incorporate language elements. Wang and Wen's (2002) L2 writing process model comprises three core components: the task environment, the composing processor, and the writer's long-term memory. The task environment encompasses the writing prompt and the existing written text. The composing processor delineates the cognitive processes undertaken by L2 writers during composition. The model outlines five stages in the L2 writing process: task examination, idea generation, idea organization, text generation, and process control. The authors also acknowledge the significant influences of both the L1 and the L2 on these processes. This model provides the theoretical foundation for analyzing how learners utilize their L1 at various stages of L2 writing, a phenomenon rarely examined in recent process-oriented writing studies.

L1 transfer in L2 writing

During the 1960s and 1970s, behaviorism influenced a strict avoidance of L1 in writing instruction (Choi & Lee, 2006). In Chinese English education, the practice of forbidding L1 transfer during writing lessons continues, with the goal of reducing errors and improving writing abilities. However, the rise of process-oriented teaching methods in the 1980s (Raimes, 1985) redirected attention to the writing process itself, prompting renewed interest in the role of L1 in second language writing. This study investigates how Chinese learners use their first language when writing in English.

Since the 1980s, the process-oriented approach has encouraged research into second language (L2) writing. Early studies compared the writing processes in L2 and first language (L1), identifying similar stages such as pre-writing, drafting, and revising (Uzawa, 1996; Arndt, 1987). Silva (1993) emphasized that L2 writing involves greater complexity. Conversely, Woodall (2002)

argued that using L1 can improve fluency and support idea generation, providing cognitive advantages. These differing views indicate that the role of L1 in L2 writing remains a topic of debate. Therefore, this study examines how Chinese learners incorporate their L1 during the writing process.

Employing L1 for L2 writing aids idea generation, topic search, and information organization. Students use L1's ease for coping with cognitive load (Al-Buainain, 2023; Bruen & Kelly, 2017) and enhancing analytical thinking (Tsuji, 2021). L1 aids idea generation, organization, and content improvement (Zhao& Hassan, 2024; Murphy & de Larios, 2010; Woodall, 2002; Wang & Wen, 2002). Recent studies support these conclusions, showing that purposeful use of L1 enhances writing fluency and coherence, particularly for learners with lower proficiency levels (Liu et al., 2024). For Chinese learners with limited language skills, L1 proves especially beneficial. Its role in L2 writing varies, including functions such as generating ideas, easing cognitive demands, assisting with vocabulary retrieval, and facilitating text revision.

Despite increasing recognition of the potential cognitive benefits of L1 use in L2 writing, existing research has primarily focused on university students or general writing behaviors, with limited attention to stage-specific L1 transfer in secondary education contexts. Furthermore, most research has focused on urban or well-funded educational environments, often neglecting learners in rural areas who may depend more on their first language due to limited English exposure. In China's EFL settings, especially in rural high schools, banning L1 use during writing instruction remains prevalent, reflecting outdated behaviorist views. However, recent studies indicate that strategic L1 use can aid various stages of the writing process, especially for students with lower language proficiency. This study seeks to fill this gap by examining how rural Chinese high school students use and are influenced by L1 throughout different stages of English writing.

Research Methodology

This section outlines the research methodology employed in the study, including a description of the participants, data collection techniques, and analytical procedures. The research collected both quantitative and qualitative data through questionnaires and interviews to explore how Chinese high school students in an EFL context use and are influenced by their first language during the second language writing process.

Research Design

To address the research question, Table One delineates the research inquiries, the employed instruments, and the data analysis methods. This study embraced a mixed-methods approach, commencing with a quantitative phase followed by a qualitative one. The quantitative segment utilised a questionnaire as the primary tool for data collection, while the qualitative component employed interviews. The rationale behind opting for a mixed-methods design stems from the recognition that solely relying on one method may not yield a comprehensive exploration of the phenomenon under investigation. A mixed-methods approach amalgamates the merits of both quantitative and qualitative methodologies. Consequently, it enables not only an extensive examination of students on a large scale through quantitative means but also a nuanced exploration of reasons and perceptions through qualitative methodologies. Thus, this investigation adopted a mixed-methods approach, employing a questionnaire followed by interviews, to acquire a thorough and insightful comprehension of the utilization of the first language (L1) in the English writings of Chinese students.

Table 1. *Research method of the study*

Research questions	Instruments	Data analysis
RQ1:When do learners transfer L1 the most in their English writings?	Questionnaire Interview	Descriptive analysis
RQ2:What are the reasons for using L1 in different writing stages?	Questionnaire Interview	Descriptive analysis Thematic analysis
RQ3:How do learners perceive the influence of L1 use in L2 writing?	Interview	Thematic analysis

In Table One shows that to address the initial research question, a questionnaire and an interview were employed. Following the prescribed sequence of the research methodology, which adhered to a sequential explanatory design, this study initially utilised a questionnaire as the research instrument. The questionnaire encompassed background information and the application of the first language (L1) in second language (L2) writing. Adapted from Stapa (Stapa, 2014), the questionnaire incorporated insights from pertinent literature to respond to the primary research question. A semi-structured interview was also adopted to explore the reasons why they transfer L1 at that time.

In addressing the second research question, the study combined quantitative and qualitative methods. Specifically, the researcher used a questionnaire to examine how students applied their first language. At the same time, the researcher conducted interviews to explore the reasons behind L1 use and the participants' perceptions of its impact on L2 writing. To answer the third research question, the researcher asked targeted interview questions to investigate perceptions of L1 transfer and its influence on L2 writing.

Participants

In this research, the researcher randomly selected 60 high school students from a rural high school in China. From this group, the researcher chose 6 participants with high English proficiency as the qualitative respondents. Details appear in Table Two.

Table 2. *Participants and sampling methods*

Instruments	Sample size	Participants	Sampling strategy
Questionnaire	60	High school students	Random sampling
Interview	4	Questionnaire respondents	Voluntary sampling

Table Two provides an overview of the participants, sample size, and the sampling strategy employed in this study. To gain insights into the transfer of the L1 among students, a cohort of 60 secondary school students in a rural area with varying English proficiency was selected. English proficiency was categorized based on students' final exam scores, as the exam follows a standardized format and administrators deliver it uniformly across schools, ensuring fairness and consistency. In this study, English proficiency was categorized based on the final exam, which has a total score of 150. Students scoring 120 and above were classified as the high-proficiency group, those scoring between 90 and 120 as the intermediate group, and those below 90 as the low-proficiency group. According to approximate equivalence, scores above 120 correspond to CEFR level B2, scores between 90 and 120 align with B1 to low B2 levels, and scores below 90 are close to A2 or lower.

The researcher based this choice on the notable disparities in English education between rural and urban settings in China. Rural areas often lack English communication opportunities and focus more on grammatical aspects, resulting in a deficiency in L2 writing skills and more reliance

on the first language for students in these regions. The decision to target second-year secondary school students is grounded in the observation that, despite hailing from different rural areas, they tend to adopt similar learning modes after a year of high school education. Consequently, they exhibit comparable approaches to English writing. The researcher determined the sample size of 60 students by considering the population of approximately 6,000 second-year secondary school students in rural areas and selecting 10% to ensure a representative sample. The researcher chose simple random sampling because it allows drawing participants from a large pool.

For the interview sessions, the researcher selected a subset of 4 students with varying levels of English proficiency through volunteer sampling. This approach acknowledges the substantial academic pressure faced by students, who are tasked with learning eight subjects, often leaving them with limited free time. Hence, fewer can attend for the qualitative phase. In addition, although the number of students was restricted, the data achieved saturation, indicating that the participants provided similar answers with no new information. Therefore, volunteer sampling was deemed appropriate to accommodate the availability and willingness of students to engage in the interview process.

Data collection

The questionnaires were distributed through the internet and were analyzed using SPSS. Descriptive analysis was adopted. Besides, after collecting quantitative data, those who were willing to attend the qualitative part were chosen as the qualitative respondents. The researcher called each participant individually to the office and conducted one-on-one interviews. The researcher audio recorded the entire process. Later, the researcher manually transcribed the audio recordings and conducted a thematic analysis.

Research instruments

The study employed two main instruments: a questionnaire and semi-structured interviews. The questionnaire was designed based on Wang and Wen's (2002) model of the L2 composing process and previous research on L1 use in writing (Stapa, 2014; Murphy & de Larios, 2010; van Weijen et al., 2009). It consisted of items that assessed participants' use of L1 at different stages of the writing process (pre-writing, while-writing, and post-writing). The questionnaire used a five-point Likert scale ranging from "strongly disagree" to "strongly agree". The reliability of the questionnaire was confirmed with a high Cronbach's alpha of 0.947, indicating strong internal consistency. Construct validity was verified through exploratory factor analysis (EFA) using SPSS, with the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy at 0.944, surpassing the recommended threshold of 0.75.

Semi-structured interviews were conducted to gain in-depth insights into participants' experiences and perceptions regarding the use of L1 in their English writing process. The interview questions focused on learners' perceived influence of L1 transfer on their writing performance.

Data Analysis

The researcher chose the data analysis methods to align with the specific research questions and the types of data collected.

For Research Questions 1 and 2 ("When do learners transfer L1 the most?" and "What are the reasons for using L1 in different writing stages?"), quantitative data from the questionnaires were analyzed using descriptive statistics (means, frequencies) to identify patterns in L1 use.

Qualitative data from interviews were analyzed using thematic analysis to uncover deeper explanations and themes regarding learners' motivations and experiences.

For Research Question 3 ("How do learners perceive the influence of L1 use in L2 writing?"), the analysis focused on qualitative interview data through thematic analysis (Buaun& Clarke, 2006), enabling exploration of learners' attitudes and reflections on the impact of L1 use in their writing.

Findings

Stages of using L1 in L2 writing

Firstly, to answer the first research question "When do learners transfer L1 the most in their English writings?", the questionnaire result is analyzed as indicated in Table Three.

Table 3. *Descriptive analysis of L1 transfer in L2 writing*

	N	Mean	Std. Deviation
Pre-writing	60	3.8722	.93093
While-writing	60	3.3306	.48201
Post-writing	60	2.7444	.94354

The results indicate that Chinese high school students most frequently transfer their first language during the pre-writing stage, where the mean L1 usage score is 3.87. This finding suggests that learners rely heavily on L1 to support activities such as idea generation, planning, and understanding the writing task at this early phase. L1 use decreases during the while-writing stage, with a mean of 3.33, indicating that although learners still use L1 to some extent (likely for lexical retrieval or translating thoughts), their dependence on L1 lessens as they engage in composing their texts in English. The lowest mean score for L1 transfer appears in the post-writing stage (mean = 2.74), showing that learners seldom use L1 during revision and editing. This reduction might reflect time constraints or a greater focus on producing English output during finalization. Overall, the pattern suggests an apparent decline in L1 usage as the writing process progresses, with the highest reliance during pre-writing and the least during post-writing.

To further examine whether L1 use differs significantly among learners with varying L2 proficiency levels across these stages, a one-way ANOVA was conducted. The results of this analysis are presented in Table Four.

Table 4. *ANOVA of L1 transfer in L2 writing among different L2 proficiency*

		Sum of Squares	df	Mean Square	F	Sig.
Pre-writing	Between Groups	9.493	2	4.747	6.937	.002
	Within Groups	30.109	44	.684		
	Total	39.603	46			
While-writing	Between Groups	1.419	2	.709	2.867	.068
	Within Groups	10.888	44	.247		
	Total	12.307	46			
Post-writing	Between Groups	3.417	2	1.709	1.866	.167
	Within Groups	40.295	44	.916		
	Total	43.712	46			

ISSN: 2229-9327

Table Four presents the results of a one-way ANOVA conducted to examine whether L1 usage across the pre-writing, while-writing, and post-writing stages differs significantly among learners with varying levels of L2 proficiency.

The ANOVA results show a statistically significant difference in L1 transfer during the pre-writing stage among learners with different L2 proficiency levels ($F = 6.937$, $p = .002$). This finding indicates that L2 proficiency is related to how much learners rely on L1 before they start writing.

Conversely, at the while-writing stage ($F = 2.867$, $p = .068$) and post-writing stage ($F = 1.866$, $p = .167$), no significant differences were found between proficiency groups in terms of L1 usage.

Further breakdown reveals that learners with lower L2 proficiency tend to use L1 more extensively during the pre-writing phase than their higher proficiency peers. This trend suggests an inverse relationship between proficiency and pre-writing L1 reliance.

Supporting these quantitative results, interview data provide illustrative examples. For instance, a high-proficiency student reported using L1 sparingly throughout the writing stages (20% pre-writing, 5% while-writing, and 10% post-writing). In contrast, a low-proficiency student described a persistent dependence on L1 during all stages, citing challenges with English language confidence and formulation.

Together, the questionnaire and interview findings underscore that L1 use is most prominent during the pre-writing phase and decreases during subsequent stages, with proficiency level influencing the degree of L1 transfer particularly at the outset of the writing process.

Reasons for using L1 at different writing stages

After analyzing the stage of L1 use, the study next explores the second research question: "What are the reasons for using L1 at different writing stages?" Learners transfer their L1 for various purposes depending on the writing stage. A closer examination of the questionnaire results, presented in Table Five, reveals these patterns.

Table 5. *L1 use at the pre-writing stage*

	N	Mean
Use L1 to think of the content	60	4.13
Use L1 to plan the structure	60	4.03
Use L1 to understand the writing task	60	3.45
Valid N (listwise)	60	

Table Five summarizes how Chinese high school students use their first language (L1) during the pre-writing stage. The data show that students most frequently rely on L1 for thinking about content ($M = 4.13$) and planning the structure of their essays ($M = 4.03$). Understanding the writing task itself also involves L1 use, though slightly less often ($M = 3.45$).

This pattern suggests that before starting to write in English, students lean heavily on Chinese to help generate ideas and organize their thoughts. This reliance on L1 for conceptualizing the essay is echoed in interviews. One student explained, "I first translate the writing prompt into Chinese to fully understand it. Then I plan my ideas and the structure in Chinese before I start writing." Similarly, Student 3 described her process as first thinking and composing sentences in Chinese, which she later translates into English. After finishing, she even translates back into

Chinese to check if the essay makes sense logically. Student 2 noted that brainstorming in Chinese helps her clear her mind and focus better on writing.

Together, the questionnaire and interview data confirm that using L1 at the pre-writing stage is a common and vital strategy for Chinese learners, supporting comprehension, structural planning, and content development. Subsequently, it delves into an examination of the reasons for using L1 by learners during the while-writing stage, as elucidated in Table Six.

Table 6. *L1 use at the while-writing stage*

	N	Mean
I come up with L1 words and translate them into L2.	60	4.13
I use L1 to search for specific words.	60	3.88
I compare L1 and L2 to find a suitable word.	60	3.47
I come up with L1 sentences and translate them into L2.	60	3.45
I compare L1 and L2 to find suitable sentence.	60	2.98
I compare L1 and L2 to choose the a suitable writing style.	60	2.72
Valid N (listwise)	60	

Table Six outlines L1 use during the while-writing phase. Here, the most common function of L1 is thinking in Chinese and translating thoughts into English (M = 4.13), followed by using Chinese to search for specific words (M = 3.88). Students also reported constructing sentences in Chinese first and then translating them (M = 3.45). Less frequent were activities involving comparing L1 and L2 to select suitable sentences or writing styles, both scoring below 3.0.

Interview insights support these findings. One student remarked, "When I'm actually writing, I try not to think in Chinese because this is an English task, but sometimes I can't help it. I often translate my Chinese thoughts into English." Student 4 shared a strategy of writing down difficult English words in Chinese to bypass vocabulary gaps and keep her writing flow.

Overall, L1 appears to function as a cognitive aid during writing—helping with translation and lexical retrieval, reducing the mental burden caused by limited English vocabulary. This use aligns with prior research emphasizing L1's role in easing working memory load during second language production. The subsequent delves into the examination of how learners employ L1 during the post-writing stage, as elucidated in Table Seven.

Table 7. *L1 use at the post-writing stage*

	N	Mean
I use L1 for backtracking.	60	3.02
I use L1 to control the writing time.	60	2.70
I use L1 to check the word number.	60	2.52
Valid N (listwise)	60	

Table Seven focuses on L1 use after writing. The results show limited reliance on L1 during this stage, with backtracking or reviewing the text being the most common use (M = 3.02). Other activities, such as controlling writing time or checking the word count, show even lower mean scores.

Interviews reveal that when students do have time to review, some use L1 by translating their English drafts back into Chinese to verify coherence and ensure they have addressed the prompt correctly.

The findings paint a clear picture of how Chinese high school learners strategically use L1 throughout the writing process. During pre-writing, L1 is heavily relied on for understanding the task, planning structure, and generating ideas. In the while-writing phase, L1 primarily supports lexical searching and translation to help manage cognitive load. Finally, in post-writing, L1 use declines but remains a tool for reviewing and revising content.

These stage-specific patterns reflect the unique challenges Chinese students face when writing in English and underscore the functional role of L1 in supporting their writing process within the Chinese educational context. In the next section, how students perceive the use of L1 is presented.

Learners' perception of L1 use in L2 writing

In response to the third research question, "How do learners perceive the influence of L1 use in L2 writing?" an interview was conducted to gather and analyse responses. The interview analysis revealed two prominent themes related to learners' perceptions of L1 use in L2 writing: L1 as a Cognitive Scaffold and Awareness of L1 Interference.

L1 as a cognitive scaffold

Most participants expressed a generally positive attitude toward the use of their first language during the writing process. They viewed L1 as a crucial cognitive tool that supports multiple aspects of their L2 writing. Many described how switching to their native language helped them better understand complex writing tasks, organize their thoughts, and retrieve appropriate vocabulary. One student explained, "Using Chinese helps me clear my mind and plan better before writing in English," highlighting how L1 acts as a mental workspace where ideas can be developed without the immediate pressure of producing grammatically correct English.

Several learners emphasized that L1 use reduced cognitive load during the early stages of writing. By thinking in their native language, they could focus more on content and logical flow without getting overwhelmed by the demands of writing in a less familiar language. For example, another student stated, "When I first read the prompt, I translate it into Chinese to make sure I really understand what is being asked. Then I plan my ideas and outline in Chinese, which makes writing in English easier." This process suggests that L1 functions as a scaffolding mechanism, providing a mental framework that supports effective L2 composition.

Awareness of L1 Interference

Despite acknowledging the benefits, learners also demonstrated a clear awareness of the potential pitfalls of over-reliance on their first language. Many expressed concern that literal translation from Chinese to English sometimes resulted in unnatural, awkward, or incorrect sentences. This finding is particularly problematic in areas such as syntax, idiomatic expressions, and word order, where the structural differences between Chinese and English are significant. One participant candidly remarked, "Sometimes, translating word-for-word from Chinese leads to strange sentences. I know the meaning is clear to me, but it might confuse the reader or sound wrong."

This awareness indicates that learners are not uncritically dependent on L1; instead, they are actively negotiating when and how to use their first language strategically. They recognize that while L1 can aid comprehension and planning, excessive or careless transfer can introduce errors

ISSN: 2229-9327

and reduce the overall quality of their writing. Some students described a conscious effort to "check" or revise their writing to avoid the pitfalls of direct translation, showing metalinguistic sensitivity and growing bilingual competence.

Discussion

Discussion on Chinese high school students' use L1 at different stages of L2 writing

The study reveals a clear pattern in L1 use across the three main stages of writing, with the pre-writing stage being the most intensive in terms of L1 reliance. During pre-writing, learners predominantly use their first language as a cognitive tool to generate ideas, plan essay structures, and comprehend the task requirements. This heavy reliance is unsurprising given that idea generation and strategic planning require deep cognitive engagement, which learners find more accessible when using their dominant language.

This finding reinforces Wang and Wen's (2002) theoretical framework, which positions L1 as essential in the early phases of L2 writing, particularly for learners with intermediate or lower proficiency who have yet to develop fully automatic L2 processing skills. The findings also support Bruen and Kelly's (2017) argument that L1 functions as a "cognitive scaffold," allowing learners to externalize their thoughts before translating them into the L2. This stage-specific reliance underscores the importance of providing learners with sufficient time and space to harness their L1 strategically during planning, a practice that has been undervalued in some L2 writing pedagogies that emphasize immediate English output.

During the while-writing phase, the study identifies a shift in L1 usage—from broad cognitive planning to specific linguistic support, notably lexical retrieval and sentence translation. Learners appear to use their L1 as a resource to bridge gaps in vocabulary and syntactic knowledge, facilitating smoother composition despite their limited English proficiency. This finding aligns with previous research (Murphy & de Larios, 2010; Tsuji, 2021), showing that L1 serves as a translational tool that reduces the cognitive load associated with retrieving and organizing unfamiliar L2 words and structures. However, the data also reveal that while learners report frequent use of L1 in interviews, quantitative measures suggest a more moderate level of use during writing. This discrepancy indicates that learners' perceptions of "L1 use" may extend beyond overt language switching to include covert mental translation processes, highlighting the complexity of accurately measuring L1 reliance.

In contrast, L1 use in the post-writing stage is minimal, primarily limited to backtracking activities such as reviewing and minor revising. This limited use is consistent with the constraints of Chinese high school education, where students often face strict time limits that preclude extended editing. The finding aligns with Wang and Wen's (2002) observation that cognitive resources are primarily allocated to initial stages of writing and less so to post-writing revision, especially under timed conditions. It also points to a potential area for pedagogical improvement: encouraging learners to engage more fully in post-writing revision could enhance writing quality and reduce errors related to L1 interference.

Another finding drawn from the questionnaire and interview results is that the use of L1 decreases with the improvement of L2 proficiency. The observed decrease in L1 use as learners' L2 proficiency increases is consistent with established research that positions L1 as a compensatory resource during early language development (Zhao& Hassan, 2024; Murphy & de Larios, 2010). This pattern indicates that learners with lower English proficiency tend to rely more heavily on their first language to manage the complex cognitive demands of writing, such as idea

generation, task analysis, and lexical retrieval. As their proficiency improves, learners can process these demands more directly in the target language, resulting in diminished dependence on L1.

This finding aligns with cognitive load theory, which suggests that learners with limited L2 skills experience higher mental effort when composing in English. Utilizing L1 helps reduce this load by facilitating pre-writing planning and lexical access. With growing L2 competence, learners develop greater automaticity and fluency in English, enabling them to conceptualize and organize ideas without reverting to their native language.

Discussion on learners' perceptions of L1 use in L2 writing

Learners in this study generally hold a positive view of L1 transfer, seeing it as an indispensable aid in making sense of the writing task and organizing their thoughts. Their perception reflects an understanding that L1 facilitates comprehension and lowers cognitive barriers, which is crucial for navigating complex writing assignments. This positive attitude corresponds with findings from Qi (1998) and Tsuji (2021), who similarly noted that L1 use can enhance cognitive efficiency and strategic planning in L2 writing.

Nonetheless, learners demonstrate a sophisticated awareness of the potential downsides of L1 use. They acknowledge that direct, literal translation from Chinese to English often leads to unnatural phrasing and grammatical errors, echoing concerns raised in cross-linguistic transfer literature about interference (Odlin, 1989). This awareness indicates that learners are not simply passive users of their L1 but critically monitor its influence on their English output. Such metalinguistic insight suggests that learners are capable of developing strategies to mitigate errors, provided they receive appropriate instruction and feedback.

This view contrasts with earlier studies that sometimes portray L1 use as either wholly detrimental or beneficial. Instead, it highlights the nuanced, dynamic nature of bilingual writing, where L1 can simultaneously serve as a scaffold and a source of error, depending on how and when it is deployed. This complexity aligns with Bruen and Kelly's (2017) conceptualization of L1 as a flexible, strategic resource, and it calls for pedagogical approaches that encourage reflective, controlled L1 use rather than strict prohibition.

Discussion on howL1 functions vary across writing stages

The distinct functions of L1 at different writing stages emphasize its multifaceted role in L2 composition. At the pre-writing stage, L1 acts primarily as a "Cognitive Scaffold," enabling learners to clarify ideas, map out essay structures, and internalize task demands in a familiar linguistic medium. This function is critical for learners with lower English proficiency, as it offloads cognitive effort and helps organize complex thoughts before engaging with the more demanding process of writing in L2. Such use of L1 mirrors Wang and Wen's (2002) emphasis on L1 in planning and idea generation.

During while-writing stage, L1's role shifts to that of a "Translational Tool," assisting learners in retrieving vocabulary and translating sentences into English. This stage-specific function addresses learners' lexical and syntactic gaps, enabling them to produce coherent L2 text despite linguistic limitations. Prior research (Murphy & de Larios, 2010; Tsuji, 2021) supports this finding, illustrating how L1 reduces working memory load and eases lexical access during composition.

In the post-writing stage, L1 use is confined to a "Review Aid," employed sparingly to backtrack and refine drafts. This limited engagement contrasts with pre- and while-writing stages,

reflecting practical constraints such as time pressure and exam-oriented writing tasks typical of Chinese secondary education. The minimal post-writing use highlights a potential weakness in current instructional practices, where revision is underemphasized, possibly limiting learners' opportunities to correct L1-based errors and improve overall text quality.

The findings of this study provide strong empirical support for the complex, stage-dependent role of L1 in L2 writing, reinforcing the theoretical model proposed by Wang and Wen (2002). The study advances this framework by providing nuanced insights into how L1 functions evolve dynamically throughout the writing process and how learners strategically regulate their L1 use according to task demands and proficiency levels.

From a pedagogical perspective, these findings challenge rigid prohibitions against L1 use in L2 writing classrooms. Instead, they advocate for instructional approaches that recognize L1 as a legitimate cognitive and linguistic resource, particularly valuable in the pre-writing and while-writing stages. Moreover, scaffolding lexical and syntactic skills through targeted vocabulary and grammar instruction could help reduce excessive reliance on L1 during the composition phase. This approach could free cognitive resources for more sophisticated English output and gradually foster greater L2 independence. Additionally, increasing opportunities for post-writing revision, with guided attention to L1 interference, could enhance learners' ability to self-monitor and improve their writing quality.

Conclusion

This study closely examined how Chinese high school students use their first language (L1) during English writing. The findings reveal that students frequently rely on their L1 in the pre-writing stage to generate ideas, comprehend the writing task, and plan their essays. Additionally, they strategically use L1 during later stages for lexical retrieval, translating thoughts, and ensuring coherence. However, participants also noted that excessive reliance on literal translation during drafting can hinder fluency and naturalness in English writing. These insights have important implications for English teaching in China. Educators and curriculum developers should recognize the constructive role of L1, particularly in brainstorming and planning, while guiding students to minimize direct translation during drafting. Instructional designs that separate idea generation in L1 from composition in L2—such as encouraging students to jot down ideas in Chinese before drafting in English with vocabulary support—may enhance writing fluency and accuracy.

Nevertheless, this study has limitations. The participant sample was primarily drawn from rural schools, which may limit the generalizability of the findings to urban or more diverse contexts. Moreover, reliance on self-reports and interviews may have led to under- or overestimation of actual L1 use. Future research could address these limitations by including more varied samples and employing methods like think-aloud protocols or direct observation to capture real-time L1 use. It would also be valuable to explore instructional interventions that help students transition from L1-based planning to English drafting, and to investigate how L1 use varies across different writing tasks and proficiency levels.

Overall, this study sheds light on the complex yet crucial role of L1 in L2 writing for Chinese learners, offering valuable directions for pedagogical practices aimed at fostering more confident and skilled English writers.

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: AI and AI-assisted technologies, specifically ChatGPT-4.0, were used in this study for grammar and language refinement. The authors take full responsibility for the accuracy and integrity of the content.
About the Authors:

Yanqin Ma is a PhD candidate in Applied Linguistics at the Language Academy, Universiti Teknologi Malaysia. Her research focuses on first language transfer in English as a Foreign Language (EFL) writing. She is particularly interested in second language writing processes and cross-linguistic influence. She has presented her work at Scopus-indexed conferences, Web of Science, and is preparing publications in peer-reviewed journals.
ORCID: http://orcid.org/0009-0000-8637-7180

Mahani Stapa is an Associate Professor at the Language Academy, Universiti Teknologi Malaysia. She holds a PhD in Applied Linguistics and has extensive experience in supervising postgraduate research. Her research interests include second language writing, bilingual education, and language assessment. She has published widely in Web of Science and Scopus-indexed journals and serves as a reviewer for several academic publications.
ORCID: http://orcid.org/0000-0002-8220-9531

References:

Abas, I. H., & Abd Aziz, N. H. (2016). Classification of L2 Writing Process and Writing Strategies. *Proceedings of The ICECRS, 1(1)*. https://doi.org/10.21070/picecrs.v1i1.505

Abdi Tabari, M. (2022). Investigating the interactions between L2 writing processes and products under different task planning time conditions. *Journal of Second Language Writing, 55, 100871*. https://doi.org/10.1016/j.jslw.2022.100871

Al-Buainain, H. (2023). The Effect of L1 Use on L2 Academic Writing Performance among EFL Learners. *Arab World English Journal*, *14(2)*, 35-50. https://dx.doi.org/10.24093/awej/vol14no2.3

Arndt, V. (1987). Six writers in search of texts: a protocol-based study of L1 and L2 writing *ELT Journal 41(4)*, 257-268.

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

Bruen, J., & Kelly, N. (2017). Using a shared L1 to reduce cognitive overload and anxiety levels in the L2 classroom. *The Language Learning Journal*, *45(3)*, 368-381.

Chen, Y., & Li, X. (2016). Effects of L1 Transfer on English Writing of Chinese EFL Students. *Cross-Cultural Communication*, *12*（4）, 50-53. https://doi.org/10.3968/8281

Cheong, C.-M., Zhu, X., & Xu, W. (2021). Source-Based Argumentation as a Form of Sustainable Academic Skill: An Exploratory Study Comparing Secondary School Students' L1 and L2 Writing. *Sustainability*, *13(22)*, 12869. https://doi.org/10.3390/su132212869

Choi, Y. H., & Lee, J. (2006). L1 use in the L2 writing process of Korean EFL students. *English Teaching, 61(1),* 205-226.
Flower, L., & Hayes, J. R. (1981). A Cognitive Process Theory of Writing. *Collece composition and communication, 32,* 365-387.
Grabe, W., & Kaplan, R. B. (2014). *Theory and practice of writing: An applied linguistic perspective*. Routledge.
Liu, Y., Cheong, C. M., Ng, R. H. W., & Tse, S. K. (2024). The role of L1 self-efficacy in L2 reading comprehension: an exploration of L1–L2 cross-linguistic transfer. *International Journal of Bilingual Education and Bilingualism, 27(7),* 883-897.
Manchón, R. M. A., de Larios, J. R., & Murphy, L. (2000). An approximation to the study of backtracking in L2 writing. *Learning and Instruction, 10(1),* 13-35.
Michel, M., Révész, A., Lu, X., Kourtali, N.-E., Lee, M., & Borges, L. (2020). Investigating L2 writing processes across independent and integrated tasks: A mixed-methods study. *Second Language Research, 36(3),* 307-334. https://doi.org/10.1177/0267658320915501
Moghadam, S. H., Samad, A. A., & Shahraki, E. R. (2012). Code switching as a medium of instruction in an EFL classroom. *Theory and Practice in Language Studies, 2(11),* 2219-2225.
Murphy, L., & de Larios, J. R. (2010). Searching for words: One strategic use of the mother tongue by advanced Spanish EFL writers. *Journal of Second Language Writing, 19(2),* 61-81.
Nystrand, M. (2023). *What writers know: The language, process, and structure of written discourse*. BRILL.
Odlin, T. (1989). *Language transfer* (Vol. 27). Cambridge University Press. Cambridge.
Qi, D. (1998). An Inquiry into Language-switching in Second Language Composing Processes. *The Canadian Modern Language Review, 54(3),* 413-435.
Raimes, A. (1985). What unskilled ESL students do as they write: A classroom study of composing. *TESOL Quarterly, 19(2),* 229-258.
Sasaki, M. (2002). Building an empirically based model of EFL learners' writing processes. *Journal of Second Language Writing, 11(3),* 259–291. https://doi.org/10.1016/S1060-3743(02)00053-1
Silva, T. (1993). Toward an understanding of the distinct nature of L2 writing. *TESOL Quarterly, 27,* 657-677.
Stapa, M. B. (2014). *Language Switching Among Malay Undergraduates in an English as a Second Language Writing*. Universiti Teknologi Malaysia.
Sun, Y., & Lan, G. (2023). A bibliometric analysis on L2 writing in the first 20 years of the 21st century: Research impacts and research trends. *Journal of Second Language Writing, 59. 100963.* https://doi.org/10.1016/j.jslw.2023.100963
Tsuji, K. (2021). Effects of L1 Use on L2 Text Quality: Rethinking Cognitive Process of Formulating L1 Texts during L2 Writing. *English Language Teaching, 14(7),* 44-57. https://doi.org/10.5539/elt.v14n7p44
Uzawa, K. (1996). Second language learners' processes of L1 writing, and translation from L1 into L2. *Journal of Second Language Writing, 5(3),* 271-294.
van Weijen, D., van den Bergh, H., Rijlaarsdam, G., & Sanders, T. (2009). L1 use during L2 writing: An empirical study of a complex phenomenon. *Journal of Second Language Writing, 18(4),* 235-250.

Wang, W., & Wen, Q. (2002). L1 use in the L2 composing process: An exploratory study of 16 Chinese EFL writers. *Journal of Second Language Writing*, *11*, 225-246.

Woodall, B. R. (2002). Language-switching: Using the first language while writing in a second language. *Journal of Second Language Writing*, *11(1)*, 7-28.

Zhao, L., & Hassan, M. (2024). Rethinking Cognitive Models in EFL Writing: An Empirical Study in Secondary Schools. *Arab World English Journal*, *15(1)*, 22–40. https://dx.doi.org/10.24093/awej/vol15no1.2

Zhu, J. (2000). The impact of language transfer on English writing of second language learners. *Foreign Language Education*, *21(1)*, 5.

Zhu, X., Guan, Y., & Yao, Y. (2022). Transfer of ideal L1 and L2 writing selves and their impacts on L2 writing enjoyment and integrated writing performance. *Assessing Writing*, *54*. 100674. https://doi.org/10.1016/j.asw.2022.100674

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.129-143
DOI: https://dx.doi.org/10.24093/awej/vol16no3.7

Effects of English Songs on EFL Learners' Speaking Performance: Lexical Complexity, Accuracy, and Fluency

Tantip Kitjaroonchai
Faculty of Liberal Arts, North Bangkok University, Thailand
Corresponding Author: thanthip@apiu.edu

Krittat Sukman
Faculty of Liberal Arts, North Bangkok University, Thailand
Email: krittat.su@northbkk.ac.th

Received: 06/30/30/3025 Accepted: 08/26/2025 Published: 09/20/2025

Abstract
Although songs are widely valued in language instruction, there is limited research on their measurable impact on English as a foreign language (EFL) learners' speaking performance. This study investigates the effects of integrating English songs into EFL instruction on the overall speaking performance of EFL undergraduate students and examines the impact of songs on lexical complexity (measured by lexical density), accuracy, and fluency. Twenty-eight first-year EFL students at a private university in Thailand participated in a fifteen-week intervention, with two sessions per week, using contemporary English songs integrated into lesson plans. The study employed a one-group pretest-posttest design. Data were collected via speaking tests (pretest, posttest, and delayed posttest) to evaluate participants' speaking performance. Quantitative data were analyzed using paired-sample t-tests, frequency, mean, and standard deviation. Results revealed a statistically significant improvement ($p < .05$) in overall speaking performance between the pretest and posttest. Lexical density showed delayed improvement in the post-delayed test, while accuracy increased significantly. Fluency also improved, evidenced by higher word counts and longer pruned speech runs ($p < .001$), suggesting enhanced grammatical control and speech fluidity. These findings underscore the pedagogical potential of English songs in developing EFL learners' speaking skills. Limitations include the small sample size and the absence of a control group. The study offers implications for educators seeking engaging, practical methods for teaching speaking in EFL contexts.

Keywords: Complexity, English as a foreign language learners, English song-based learning, fluency, lexical density, speaking performance

How to cite: Kitjaroonchai, T. & Sukman, K. (2025). Effects of English Songs on EFL Learners' Speaking Performance: Lexical Complexity, Accuracy, and Fluency. *Arab World English Journal*, *16*(3):129-143. https://dx.doi.org/10.24093/awej/vol16no3.7

Introduction

English serves as a global lingua franca, essential for international communication in contexts such as the Association of Southeast Asian Nations (ASEAN), and offers access to employment and higher education (Franco & Roach, 2018; Kirkpatrick, 2012; Suoc et al., 2025). However, achieving oral fluency in English remains a common challenge for many EFL learners, as shown in studies highlighting the need for immersive, communicative environments to improve speaking performance (Al Zoubi, 2018; Keomany et al., 2020; Santiwatthanasiri, 2018).

In Thailand, English is a mandatory subject from primary school to university; however, instruction predominantly uses Thai, limiting authentic English exposure. Consequently, many Thai university students struggle with proficient English communication, often feeling shy, lacking confidence, and fearing mistakes (Sha'ar & Boonsuk, 2021; Yuh & Kaewurai, 2021). Teachers' excessive use of Thai further hinders speaking abilities. This limited exposure presents a key challenge for English instructors: designing learning experiences that increase students' English exposure and improve their speaking skills (Phisutthangkoon, 2024; Santiwatthanasiri, 2018; Triwittayayon & Sarobol, 2019).

One such strategy is the integration of English songs into instruction. The Affective Filter Hypothesis (Krashen, 1985) supports the idea that enjoyable and low-anxiety activities—such as using songs—can reduce emotional barriers like anxiety and increase motivation, thereby enhancing language acquisition. Empirical studies have shown that songs improve vocabulary, pronunciation, listening comprehension, and learners' confidence in speaking (Astuti & Parmawati, 2022; Khan, 2024; Palupi et al., 2022; Saldiraner & Cinkara, 2021; Zhang et al., 2023).

However, although some studies have explored the use of English songs in language learning (Astuti & Parmawati, 2022; Nipattamanon, 2018; Phoowong, 2016; Songkhro, 2023), systematic research on their specific impact on speaking performance—particularly in terms of lexical complexity, accuracy, and fluency—remains limited. To address this gap, the present study examines the effects of integrating English songs into instruction on EFL undergraduates' speaking performance by comparing results from pretest, posttest, and delayed posttest assessments. The aim is to provide pedagogical insights into the potential of musical input to enhance EFL learners' spoken English skills.

The present study seeks to answer the following research questions:

1. What are the effects of English songs on the overall speaking performance of EFL undergraduate students?
2. How do English songs influence EFL undergraduate students' speaking production in terms of lexical complexity, accuracy, and fluency?

Literature Review

This literature review is grounded in the Complexity, Accuracy, and Fluency (CAF) triad (Housen et al., 2012), which offers a structured lens for analyzing learners' spoken performance and evaluating pedagogical interventions, such as the use of songs in EFL classrooms. It also incorporates Krashen's (1985) Affective Filter Hypothesis to explain how music can reduce anxiety and facilitate more effective language input and speaking development in EFL contexts.

Speaking Challenges Among EFL Learners

According to Fedorenko et al. (2024), language serves as a fundamental tool for communication. Speaking, a productive skill, is often considered particularly challenging for English as a Second Language (ESL) or Foreign Language (EFL) learners due to the complexities of mastering English structures for accurate and fluent oral production. Brown (2001) stated, "The benchmark of successful language acquisition is almost always the demonstration of an ability to accomplish pragmatic goals

through interactive discourse with other speakers of the language" (p.267). Similarly, Gebhard (2017) describes conversing in a second language as "knowing how to maintain interaction and focus on meaning; use conversational grammar; introduce, develop, and change topics; take turns; apply conversational routines; and adapt style to match the setting/context" (p.175). Thus, dedicated practice is essential for EFL learners to achieve an acceptable level of speaking proficiency (Ratnasari, 2020; Winnie et al., 2023).

However, EFL learners face common challenges, including limited vocabulary, grammatical difficulties, pronunciation issues, and psychological barriers such as anxiety and fear of making mistakes (Alrasheedi, 2020; Hannaoui, 2017; Purwati et al., 2023). A lack of exposure to English and overuse of the mother tongue further reduce opportunities for authentic communication (Amoah & Yeboah, 2021; Chand, 2021; Normawati et al., 2023).

Vocabulary limitations, even with theoretical knowledge, often prevent learners from expressing ideas confidently, while insufficient practice applying grammar leads to hesitation and reduced fluency (Keomany et al., 2020; Nety, 2022; Scrivener, 2011). Pronunciation errors and a lack of confidence negatively affect speaking engagement (Chand, 2021; Winnie et al., 2023). In contrast, regular exposure to English enhances English proficiency (Al Zoubi, 2018; Santiwatthanasiri, 2018).

In the Thai context, where English is a foreign language, these hindrances are particularly pronounced. Despite years of English education, Thai students often exhibit low proficiency and struggle with oral communication due to limited exposure (Waluyo, 2019). Survey findings corroborate that Thai undergraduate students attribute their speaking difficulties to a lack of vocabulary, grammar issues, and insufficient English exposure. Fear of making mistakes and anxiety are also commonly cited as significant deterrents (Cherdchoopong, 2020; Santiwatthanasiri, 2018; Sha'ar & Boonsuk, 2021; Yaorm, 2019; Yuh & Kaewurai, 2021).

The Pedagogical Value of Songs in Speaking Instruction

Scrivener (2011) observed that songs are often used in EFL classrooms as filler activities to shift the mood or tempo of a lesson. However, songs can also be purposefully integrated into the course design to support language learning. Their value lies in offering a multisensory and motivational learning experience, which helps develop speaking skills by engaging learners through rhythm, melody, and meaningful context. Songs support oral language development by integrating key linguistic components such as vocabulary, grammar, pronunciation, and listening comprehension, all of which contribute to fluent and intelligible speech. Numerous studies highlight the role of songs in promoting vocabulary acquisition and retention. Through repetition and context, songs enable learners to internalize and recall language more effectively (Riensumettharadol, 2019; Tilwani et al., 2022; Zaharani, 2023; Zamin et al., 2020). In addition, grammar and syntax—critical to constructing effective speech—are reinforced through repetitive and contextualized patterns found in song lyrics (Busse et al., 2021). Pronunciation skills also benefit, as songs expose learners to natural phonological features such as stress, rhythm, and intonation. These features help learners imitate native-speaker models (Astuti & Parmawati, 2022; Khan, 2024; Palupi et al., 2022; Saldiraner & Cinkara, 2021; Zhang et al., 2023).

Songs as a Bridge Between Listening and Speaking Skills

In addition to their positive effects on productive skills, songs also support the development of listening comprehension—a foundational component of effective spoken communication. Listening to songs exposes learners to authentic input, including connected speech and colloquial expressions, helping them interpret meaning more accurately and respond appropriately in conversation (Mobbs & Cuyul, 2018; Rahayu & Rifansyah, 2022). This exposure builds auditory discrimination and improves learners' sensitivity to nuances in pronunciation, tone, and pacing. The affective benefits of music further support

language acquisition. Songs help create a low-anxiety environment, aligning with Krashen's (1985) Affective Filter Hypothesis, and enabling learners to speak with greater confidence and less fear of making mistakes. Learners themselves have reported positive attitudes toward learning English through songs, describing them as enjoyable and anxiety-reducing (Al-Efeshat & Baniabdelrahman, 2020; Fuekfon & Thumawongsa, 2023; Kim et al., 2024; Lume et al., 2023). Studies also note that learners feel more motivated and emotionally engaged when songs are used in the classroom, and they often perceive singing or listening activities as fun and socially inclusive (Güzel & Altay, 2023; Kim et al., 2024), which can further reduce communicative apprehension.

Practical Considerations for Effective Song Integration

While songs offer clear pedagogical benefits, effective integration into EFL instruction requires careful planning. Teachers may face challenges such as selecting songs suitable for learners' proficiency levels, working within time constraints, and lacking training in song-based methods (Bokiev & Ismail, 2021; Xi, 2021). Learners, too, may struggle with fast-paced lyrics or idiomatic language (Muhamad & Rahmat, 2020). Therefore, songs should be chosen based on learners' age, proficiency, vocabulary demands, and cultural relevance (Ludke, 2009; Mobbs & Cuyul, 2018; Simpson, 2015; Tasnim, 2022).

As previously discussed, songs support the development of speaking-related skills that contribute to improvements in complexity, accuracy, and fluency. Lexical complexity, often assessed through lexical density, improves through repeated exposure to varied vocabulary in context. Research on complexity, accuracy, and fluency provides a valuable framework for understanding these gains (Housen et al., 2012). Songs also reinforce grammatical accuracy through consistent, patterned input (Busse et al., 2021; Gürsoy & Karakas, 2018; Lan et al., 2020) and enhance fluency by promoting automaticity and familiar language chunks (Jain et al., 2024; Lume et al., 2023).

Together, these benefits highlight the multifaceted role of songs in enhancing both learner engagement and key aspects of speaking. While previous research has explored the effects of song integration on specific elements such as pronunciation, vocabulary acquisition, or grammar improvement—each of which contributes to speaking development—few studies have examined the comprehensive impact of songs on overall speaking performance. In particular, limited research has assessed how songs influence learners' speaking using the CAF triad, especially within the Thai EFL context. To address this gap, the present study investigates the effects of song-based instruction on the speaking development of EFL learners, employing the CAF triad to provide a more holistic evaluation.

Methodology

The study employed a one-group pretest-posttest design. Data were collected via speaking tests (pretest, posttest, and delayed posttest) to assess the impact of English songs on EFL learners' speaking performance. Quantitative methods were used to analyze three key dimensions of speaking ability: lexical complexity, accuracy, and fluency. Instructional interventions were integrated into regular speaking lessons over ten weeks. Data were collected through planned speaking activities and analyzed using standard measurement tools and statistical methods.

Participants

This study was conducted during the first semester of the 2024–2025 academic year at the Faculty of Arts and Humanities, Asia-Pacific International University (AIU), Thailand. Participants were 28 students enrolled in the ENGL2210: Dialogue and Conversation Skills in English course, which focuses

on speaking and aims to improve communication through everyday conversational practice. Using purposive sampling, the study targeted students with the Common European Framework of Reference for Languages (CEFR) A1–A2 proficiency levels based on placement test scores. The group included 14 students from the Professional Communication program, four from Accounting, and 10 ESL students preparing for international degree entry. Participants came from Thailand, China, Malaysia, and India, providing a diverse sample for examining the impact of English songs on EFL speaking development.

Instruments

This study used two main instruments for data collection: (1) a series of speaking tests (pre-test, post-test, and delayed post-test) and (2) speaking lessons incorporating English songs as an instructional intervention.

Speaking Tests

The primary instrument was a sequence-picture-based speaking test administered individually at three points: before instruction (pre-test, Week two), after ten weeks of instruction with English songs (post-test), and three weeks later (delayed post-test) to assess retention. Each test used four colored sequence pictures from *Sequencing Pictures: Picture Sequences to Talk or Write About* (Easylearn, 2015), depicting a familiar scenario of a woman and a boy traveling by bus. Participants had three minutes to observe and take notes, then two minutes to narrate while being audio recorded. Test content was validated by three language teaching experts to ensure validity.

Speaking Lessons with English Songs

Ten lesson topics (e.g., Daily Routines, Our Family, Nature Around Us) were chosen for relevance to students' daily lives. Each topic was paired with a thematically related English song, including *This Is The Way*, *What Are You Doing?*, *We Are a Family*, *Beautiful Sunday*, *What a Wonderful World*, *The Gift I'll Give, At My Worst, and Don't Cry Joni*. These songs were used as a starting point for communicative speaking exercises that targeted improvements in word choice, sentence accuracy, and overall fluency.

Procedures

The study took place during the first semester of 2024–2025 at Asia-Pacific International University. After selecting participants by placement test scores, the research proceeded as follows:

1. Participants completed a pre-speaking test using sequence pictures in Week Two.
2. They attended ten weeks of speaking lessons incorporating English songs.
3. A post-speaking test identical to the pre-test was administered after instruction.
4. A delayed post-test was given three weeks later to assess skill retention.
5. All speaking tests were recorded, transcribed, and analyzed.

Data Analysis

Lexical Density Analysis

To measure lexical complexity, this study examined lexical density, defined as the proportion of content words (nouns, verbs, adjectives, and adverbs) in a text. Participants' speech from the pretest, posttest, and delayed posttest was transcribed using *VEED.IO* (n.d.) and analyzed using *Analyze My Writing* (n.d.), a free online text content and readability analyzer that calculates lexical density by dividing the number of lexical words by the total word count. This method was chosen for its accessibility and alignment with the study's objectives. To evaluate changes over time, paired samples t-tests were

conducted to compare lexical density scores across the three tests, determining whether statistically significant improvements occurred and whether learners retained lexical richness after instruction.

Formula:

$$Lexical\ Density = \frac{\text{Total Number of Words } \times 100}{\text{Number of Lexical Words}}$$

Speaking Accuracy Analysis

Accuracy was assessed using methods adapted from Khan (2010), focusing on two indicators: the number of errors per 100 words and the percentage of error-free clauses (see Table 1). This approach is particularly suitable for EFL beginners, as it allows shorter, simpler clauses to be fairly assessed. Transcriptions from all speaking tests were carefully coded for grammatical errors, and both metrics were calculated for each participant across the pretest, posttest, and delayed posttest. Paired samples t-tests were used to compare mean accuracy scores across the three test stages to determine whether the use of English songs contributed to statistically significant improvements in speaking accuracy and whether such gains were sustained over time.

Table 1. *Calculation of the accuracy rate*

Terms	Definition
Total clauses (T-units) ➡	An independent clause and any clauses dependent on it.
Percentage of error-free clauses ➡	The percentage of clauses that do not contain any errors of the total number of clauses.
Errors per 100 words ➡	Number of errors divided by the total number of words produced, divided by 100

Speaking Fluency Analysis

Fluency was evaluated based on the number of meaningful words spoken per minute, using the pruned speech rate (PSR) adapted from Bui and Huang (2018). PSR excludes elements such as long pauses, fillers, repetitions, and incomplete words, thus offering a more accurate measure of actual speech production. The researcher calculated each participant's PSR from the transcribed data collected during the two-minute speaking task in the pretest, posttest, and delayed posttest. To assess improvements and retention of fluency, paired samples t-tests were conducted to compare fluency scores across the three tests. This analysis provided insight into whether learners' speech became smoother and more fluent after engaging in the song-integrated speaking lessons and whether the fluency gains persisted over time. The calculation of PSR is the formula outlined below.

$$\text{PSR} = \frac{\text{total words} - \text{vocal fillers} - \text{incomplete words} - \text{repeated words}}{\text{Total time of speech (in seconds)}} \times 60$$

Additional variables analyzed under speaking fluency included:

- The number of long pauses (LP) lasting more than three seconds
- The duration of pauses (DP)
- The frequency of filled pauses (FP), such as 'um,' 'like,' 'you know,' or similar vocalizations.

Results

To address Research Question one, which explored the effects of English songs on the overall speaking performance of EFL undergraduate students, a paired-sample t-test was employed to examine

changes in speaking scores across three assessment points: pre-speaking test, post-speaking test, and delayed post-speaking test. The descriptive and inferential statistics are presented in Table 2.

Table 2. *Results of overall speaking performance in pretest, posttest, and delayed posttest*

Pair	Comparison	N	M	SD	MD	*t*	df	*p*
1	Pre-Speaking Test	28	3.67	0.99	-1.14	-9.63	27	0.000
	Post-Speaking Test	28	4.81	0.66				
2	Pre-Speaking Test	28	3.67	0.99	-1.27	-10.41	27	0.000
	Delayed Post-Speaking Test	28	4.95	0.59				
3	Post-Speaking Test	28	4.81	0.66	-0.13	-1.89	27	0.069
	Delayed Post-Speaking Test	28	4.95	0.59				

Note. N = Number of participants; M = Mean; SD = Standard Deviation; MD = Mean Difference; *t* = t-value; *df* = degrees of freedom; *p* = p-value. $p < .001$ (2-tailed).

The results revealed a statistically significant improvement in speaking performance from the pre-speaking test (M = 3.67, SD = 0.99) to the post-speaking test (M = 4.81, SD = 0.66), with a mean difference of -1.14, $t(27) = -9.63$, $p < .001$. This indicates that the intervention using English songs had a considerable immediate positive impact on students' speaking abilities.

Furthermore, a significant improvement was also found between the pre-speaking test and the delayed post-speaking test (M = 4.95, SD = 0.59), with a mean difference of -1.27, $t(27) = -10.41$, $p < .001$. This finding suggests that the improvement in speaking performance was not only immediate but also sustained over time. In contrast, the comparison between the post-speaking test (M = 4.81, SD = 0.66) and the delayed post-speaking test (M = 4.95, SD = 0.59) showed no statistically significant difference (M difference = -0.13, SD = 0.36), $t(27) = -1.89$, $p = .069$. Although there was a slight increase in the mean score, it was not statistically significant, indicating that students maintained their enhanced speaking skills during the delayed posttest period.

Taken together, these findings provide strong empirical support for the effectiveness of English songs in enhancing and retaining EFL learners' speaking skills.

To address Research Question two, which examines the effects of English songs on EFL undergraduate students' speaking production in terms of lexical density, accuracy, and fluency, a series of paired-sample t-tests was conducted. These tests compared students' scores across three time points—pretest, posttest, and delayed posttest—for each of the three dimensions. The results are presented in the following tables.

Table 3. *Mean scores and standard deviations of lexical density in the pre-, post-, and delayed post-speaking tests*

	Variables	N	M	SD	MD	*t*	df	Sig. (2-tailed)
Pair 1	Pre VCD	28	49.47	4.64	-0.04	-0.03	27	.971
	Post VCD	28	49.52	4.08				
Pair 2	Post VCD	28	49.52	4.08	-1.71	-2.27	27	.031
	Delayed Post VCD	28	51.23	3.90				
Pair 3	Pre VCD	28	49.47	4.64	-1.75	-1.42	27	.165
	Delayed Post VCD	28	51.23	3.90				

Note. N = Number of participants; M = Mean; SD = Standard Deviation; MD = Mean Difference; *t* = t-value; *df* = degrees of freedom; Pre VCD = Pre-vocabulary density; Post VCD = Post-vocabulary density; Delayed Post VCD = Delayed post-vocabulary density.

To investigate the effects of English songs on EFL undergraduate students' lexical density, (see table 3) the results revealed no statistically significant difference in lexical density between the pre-speaking test (M = 49.47, SD = 4.64) and the post-speaking test (M = 49.52, SD = 4.08), $t(27) = -0.03$, $p = .971$. The negligible mean difference of -0.04 indicates that the intervention did not lead to an immediate change in lexical density. However, a statistically significant increase in lexical density was observed between the post-speaking test and the delayed post-speaking test (M = 51.23, SD = 3.90), $t(27) = -2.27$, $p < .05$, with a mean difference of -1.71. This finding suggests that improvements in lexical density may have emerged over time, potentially as a delayed effect of the intervention.

Although the comparison between the pre-speaking test and the delayed post-speaking test showed an increase in lexical density (MD = -1.75), this difference did not reach statistical significance, $t(27) = -1.42$, $p = .165$. Overall, these findings suggest that while the use of English songs did not produce an immediate impact on lexical density, a delayed improvement was observed, indicating that gains in lexical complexity may require time to consolidate following the intervention.

To examine the impact of English songs on the speaking accuracy of EFL undergraduate students, two indicators were analyzed: errors per 100 words (ERW) and error-free clauses (EFC). Paired-sample t-tests were conducted to compare students' performance before and after the intervention. The results are presented in Table 4.

Table 4. *Mean scores and standard deviations for errors per 100 words and error-free clauses*

	Variables	N	M	SD	MD	*t*	df	Sig. (2-tailed)
Pair 1	Pre_ERW	28	11.52	2.76	1.78	2.58	27	.016
	Post_ERW	28	9.74	2.21				
Pair 2	Pre_EFC	28	8.67	13.94	-15.58	-3.14	27	.004
	Post_EFC	28	24.25	17.74				

Note. N = Number of participants; M = Mean; SD = Standard Deviation; MD = Mean Difference; *t* = t-value; *df* = degrees of freedom; Pre ERW = Pre-error per 100 words; Pre EFC = Pre-error-free clause; Post ERW = Post-error per 100 words; Post EFC = Post-error-free clause.

The analysis revealed a statistically significant reduction in the number of errors per 100 words from the pre-speaking test (M = 11.52, SD = 2.76) to the post-speaking test (M = 9.74, SD = 2.21), $t(27) = 2.58$, $p < .05$. The mean difference of 1.78 indicates that students produced fewer errors relative to word count following the intervention, suggesting improved grammatical accuracy. Similarly, the number of error-free clauses increased significantly from the pre-speaking test (M = 8.67, SD = 13.94) to the post-speaking test (M = 24.25, SD = 17.74), $t(27) = -3.14$, $p < .05$. The mean difference of -15.58 suggests that students became more proficient at forming grammatically correct clauses following the use of English songs.

Overall, these findings indicate that incorporating English songs had a statistically significant and positive effect on students' speaking accuracy, as evidenced by both a reduction in errors and an increase in error-free clause production.

To evaluate the impact of using English songs on students' speech fluency, a series of paired-sample t-tests were conducted to compare pre- and post-speaking test scores on five fluency-related measures: total word count (WC), pruned speech rate (PSR), number of long pauses (LP), duration of pauses (DP), and frequency of filled pauses (FP). Descriptive statistics are summarized in Table 5.

Table 5. *Mean scores and standard deviations of fluency in the pre- and post-speaking tests*

	Variables	N	M	SD	MD	*t*	df	Sig. (2-tailed)
Pair 1	Pre_WC	28	62.50	34.80	-38.14	-5.77	27	.000
	Post_WC	28	100.64	39.33				
Pair 2	Pre_PSR	28	27.52	14.40	-16.87	-5.66	27	.000
	Post_PSR	28	44.39	17.45				
Pair 3	Pre_LP	28	2.43	2.20	0.53	1.16	27	.253
	Post_LP	28	1.89	1.79				
Pair 4	Pre_DP	28	13.57	13.89	6.41	2.63	27	.014
	Post_DP	28	7.16	7.61				
Pair 5	Pre_FP	28	5.36	4.60	-2.35	-2.36	27	.026
	Post_FP	28	7.71	6.79				

Note. N = Number of participants; M = Mean; SD = Standard Deviation; MD = Mean Difference; *t* = t-value; *df* = degrees of freedom; Pre WC = Pre-word count; Post LP = Post-long pause; Post WC = Post-word count; Pre DP = Pre-duration of pause; Pre PSR = Pre-pruned speech run; Post DP = Post-duration of pause; Post PSR = Post-pruned speech run; Pre FP = Pre-filled pause; Pre LP = Pre-long pause; Post FP = Post-filled pause.

The analysis revealed a statistically significant increase in word count, with participants producing more words per minute in the post-speaking test (M = 100.64, SD = 39.33) compared to the pre-speaking test (M = 62.50, SD = 34.80), $t(27) = -5.77$, $p < .001$. The mean difference of -38.14 indicates a substantial improvement in overall verbal output.

Similarly, a significant increase was observed in pruned speech rate, rising from a pre-test mean of 27.52 (SD = 14.40) to a post-test mean of 44.39 (SD = 17.45), $t(27) = -5.66$, $p < .001$. The mean difference of -16.87 suggests enhanced fluency after the intervention, with participants producing longer, uninterrupted stretches of speech. In terms of pausing behavior, the duration of pauses significantly decreased from a pre-test mean of 13.57 seconds (SD = 13.89) to 7.16 seconds (SD = 7.61) in the post-test, $t(27) = 2.63$, $p < .05$, indicating a reduction in hesitation time. However, the number of long pauses did not show a statistically significant change ($t(27) = 1.16$, $p = .253$), suggesting that while participants hesitated less in duration, the frequency of longer pauses remained relatively stable.

Regarding filled pauses (e.g., "um," "uh"), a significant increase was noted from pre-test (M = 5.36, SD = 4.60) to post-test (M = 7.71, SD = 6.79), $t(27) = -2.36$, $p < .05$. The mean difference of -2.35 may reflect participants' greater verbal effort during the post-test, albeit with a trade-off in disfluency markers.

Collectively, these findings suggest that the use of English songs positively influenced several dimensions of speech fluency, particularly in increasing speech output, enhancing the flow of speech, and reducing pause duration. However, the increase in filled pauses may indicate a transitional phase in fluency development, wherein learners produce more speech but continue to rely on verbal fillers.

Discussion

This study investigated the effects of integrating English songs on the overall speaking of EFL undergraduate students and examined the impact of songs on lexical complexity (measured by lexical density), accuracy, and fluency. The findings demonstrated meaningful progress in many areas, highlighting the instructional benefits of using music as a tool in English language teaching.

Overall Speaking Performance

The substantial improvements recorded between the pretest and both the posttest and delayed posttest suggest that integrating songs into instruction had a positive impact on students' overall speaking performance. This outcome aligns with previous research highlighting how music contributes to low-anxiety learning environments that support oral communication and reduce learners' fear of speaking (Kim et al., 2024; Lume et al., 2023). The maintained progress evident in the delayed posttest reinforces the idea that music supports the notion that songs aid in long-term retention of language skills (Güzel & Altay, 2023; Lume et al., 2023).

Drawing on Krashen's (1985) Affective Filter Hypothesis, music lowers emotional barriers such as fear of errors and lack of confidence, which often inhibit EFL learners' oral performance (Al-Efeshat & Baniabdelrahman, 2020; Fuekfon & Thumawongsa, 2023). In emotionally supportive settings, students are more willing to engage in speaking activities, leading to improved oral performance. Moreover, the multisensory and emotionally engaging nature of music boosts motivation and supports retention of language skills (Güzel & Altay, 2023; Lume et al., 2023). The social aspects of classroom singing further reduce communicative anxiety and encourage participation, particularly in culturally diverse settings (Mobbs & Cuyul, 2018; Saldiraner & Cinkara, 2021). These affective, cognitive, and social dimensions collectively explain how song-based instruction leads to meaningful and lasting improvements in EFL learners' speaking performance.

Lexical Complexity

In terms of lexical complexity, measured by lexical density, no immediate improvement was observed from pretest to posttest; however, a statistically significant gain was recorded in the delayed posttest. This delayed improvement aligns with the observation of Housen et al. (2012), who noted that EFL learners often begin with limited vocabulary and tend to rely on a small set of familiar words to express broader ideas. Similarly, Lume et al. (2023) found that singing combined with interactive activities can enhance vocabulary retention. The delayed increase in lexical density in the present study suggests that the benefits of learning through songs may require extended exposure and time to manifest, with gains in lexical complexity emerging gradually. Moreover, the delayed increase in lexical density aligns with the view that lexical development requires not only exposure but also active use, which may occur over time through classroom interaction and outside practice. Thus, while the short-term effect was negligible, the longer-term gain affirms the value of English songs as a scaffold for expanding learners' lexical repertoires (Riensumettharadol, 2019; Tilwani et al., 2022).

Accuracy

The improvements in speech accuracy were both immediate and statistically significant, evidenced by the reduction in errors per 100 words and the increase in error-free clauses. These results corroborate earlier studies emphasizing the role of music in reinforcing grammatical structures (Busse et al., 2021; Gürsoy & Karakas, 2018). In their research, Gürsoy and Karakaş (2018) found that while English songs did not improve learners' fluency, they did help participants use the provided vocabulary more accurately, supporting gains in grammatical precision. Songs provide repetitive and well-structured language input, which allows learners to internalize correct grammatical forms through repeated exposure. This evidence supports Khan's (2010) view that musical input enhances grammar acquisition and matches evidence from Lan et al. (2020), who found improvements in learners' grammatical competence through song-based instruction.

Fluency

Perhaps the most pronounced impact of the intervention was observed in the domain of fluency. Statistically significant improvements in word count and pruned speech rate demonstrate that students became more verbally productive and able to sustain speech with fewer hesitations. These findings are in line with the literature, noting that songs support the development of fluency by providing formulaic sequences, rhythm, and intonation patterns that help learners internalize everyday speech chunks and speak more smoothly (Jain et al., 2024; Lume et al., 2023; Zhang et al., 2023). One possible explanation for the students' improvement in fluency is that, as students became more familiar with language elements through songs, their fluency improved when using them in speaking activities. This finding is consistent with Bui and Huang (2018), who found that content familiarity enhances fluency. However, it is essential to note the unexpected increase in filled pauses (e.g., "um," "uh"), which may point to an intermediate developmental phase in fluency acquisition. As students attempted to speak more and at greater length, they may have resorted to fillers to maintain fluency while organizing their thoughts. The lack of significant change in the frequency of long pauses further supports the interpretation that while students produced more and longer utterances, their pause behavior remained relatively consistent. Taken together, these results suggest that song-based instruction positively influences fluency by fostering both increased output and smoother delivery, even as learners work through normal developmental disfluencies.

Conclusion, Implications, and Future Research

This study offers strong evidence that incorporating English songs into EFL teaching significantly improves students' speaking abilities, especially in grammatical accuracy and fluency, with delayed but notable enhancements in lexical density. The statistically significant gains across pretest, posttest, and delayed posttest demonstrate not only the immediate benefits of song-based instruction but also its potential to foster long-term language growth. These findings reinforce the theoretical underpinnings of Krashen's (1985) Affective Filter Hypothesis and are consistent with a vast body of research highlighting the linguistic, cognitive, and emotional benefits of music in language learning (Keomany et al., 2020; Kim et al., 2024; Zamin et al., 2020).

From a teaching perspective, the findings highlight the usefulness of English songs as a multifaceted instructional tool. Songs promote vocabulary retention, facilitate grammar acquisition, and foster fluency by providing repeated exposure to rhythmically patterned and contextually meaningful language (Busse et al., 2021; Gürsoy & Karakas, 2018). Additionally, they contribute to creating inclusive, low-anxiety learning environments that encourage greater oral participation and reduce communication apprehension (Fuekfon & Thumawongsa, 2023; Kim et al., 2024; Lume et al., 2023; Saldiraner & Cinkara, 2021). Teachers are encouraged to adopt songs strategically, choosing materials that matches learners' proficiency, reflects cultural relevance, and aligns with curriculum goals.

Despite the promising outcomes, limitations must be acknowledged. The absence of a control group and the relatively small, purposively selected sample limit the generalizability of the findings. Future studies should employ larger, randomized samples and comparative experimental designs to confirm causality. Moreover, longitudinal studies could explore how sustained exposure to song-based instruction influences other areas of language proficiency, such as pronunciation, discourse competence, and pragmatic use. Finally, qualitative inquiries into learner engagement, identity, and motivation would complement quantitative findings and offer richer pedagogical insights into how music facilitates second language development.

Funding: This research is funded.

Acknowledgments: We would like to express our heartfelt gratitude to the Research, Ranking, and Development Committee for their ethical approval of this project (RRDC 2024-173).

Conflicts of Interest: The authors declared no conflict of interest.

Authenticity: This manuscript is an original work.

Artificial Intelligence Statement: During the preparation of this work, the authors used ChatGPT to check grammar, paraphrase text, and enhance language and readability. After using the tool, the authors reviewed and edited the content as needed. All content and interpretations remain the responsibility of the authors.

About the Authors:

Tantip Kitjaroonchai is a Ph.D. candidate at North Bangkok University. Her research interests include second language acquisition and language teaching methodologies.
ORCID: https://orcid.org/0009-0002-1076-7742

Krittat Sukman, Ph.D., is a lecturer in the Graduate School of English at North Bangkok University, Thailand. His research interests include language teaching methodology and corpus linguistics. ORCID: https://orcid.org/0009-0001-0540-048X

References

Al-Efeshat, H., & Baniabdelrahman, A. (2020). The EFL teachers' and students' attitudes towards the Use of Songs in Learning English. International Online Journal of Education and Teaching, 7(3), 844–858.

Alrasheedi, S. (2020). Investigation of factors influencing speaking performance of Saudi EFL learners. *Arab World English Journal*, *11*(4). 66-77. https://dx.doi.org/10.24093/awej/vol11no4.5

Al Zoubi, S. M. (2018). The impact of exposure to English language on language acquisition. *Journal of Applied Linguistics and Language Research, 5*(4), 151-162.

Amoah, S., & Yeboah, J. (2021). The speaking difficulties of Chinese EFL learners and their motivation towards speaking the English language. *Journal of Language and Linguistic Studies*, *17*(1), 56-69. https://dx.doi.org/10.52462/jlls.4

Analyze My Writing. (n.d.) *Analyze My Writing*: A free online text analysis tool. https://www.analyzemywriting.com

Astuti, K., & Parmawati, A. (2022). The influence of English songs to improve student's pronunciation ability. *Professional Journal of English Education*, *5*(1), 75-80.

Bokiev, D., & Ismail, L. (2021). Malaysian ESL teachers' beliefs and practices regarding the use of music and songs in second language teaching. *The Qualitative Report*, *26*(5), 1497-1521. https://doi.org/10.46743/2160-3715/2021.4350

Brown, H. D. (2001). *Teaching by principles: An interactive approach to language pedagogy* (2nd ed.). New York: Longman.

Bui, G., & Huang, Z. (2018). L2 fluency as influenced by content familiarity and planning: Performance, measurement, and pedagogy. *Language Teaching Research*, *22*(1), 94-114. https://doi.org/10.1177/1362168816656650

Busse, V., Hennies, C., Kreutz, G., & Roden, I. (2021). Learning grammar through singing? An intervention with EFL primary school learners. *Learning and instruction*, *71*, 101372. https://doi.org/10.1016/j.learninstruc.2020.101372

Chand, G. B. (2021). Challenges faced by bachelor level students while speaking English. *Indonesian Journal of English Language Teaching and AppliedLinguistics*, *6*(1), 45-60.

Cherdchoopong, J. (2020). *A study of English communication problems in listening and speaking of third-year students majoring in English at a public university in Bangkok* [Unpublished independent study]. Thammasat University.

Easylearn. (2015). *Sequencing pictures: Picture sequences to talk or write about*. https://www.crumpsalllaneprimary.org/uploads/1/2/2/8/122801150/sequencingpictures2.pdf

Fedorenko, E., Piantadosi, S. T., & Gibson, E. A. (2024). Language is primarily a tool for communication rather than thought. *Nature, 630*(8017), 575-586. https://doi.org/10.1038/s41586-024-07522-w

Franco, A., & Roach, S. S. (2018). An assessment of the English proficiency of the Thai workforce and its implication for the ASEAN economic community: An empirical inquiry. *Open Journal of Business and Management*, *6*(3), 658-677. https://doi.org/10.4236/ojbm.2018.63050

Fuekfon, S. & Thumawongsa, N. (2023). Using English songs to enhance linking sound's awareness. *Rajapark Journal*, *17*(53), 18-32.

Gebhard, J. G. (2017*). Teaching English as a foreign or second language: A teacher self-development and methodology guide.* University of Michigan Press.

Gürsoy, E., & Karakas, O. (2018). The effects of singing in English on the speech fluency of Turkish teenagers learning English as a foreign language. *International Online Journal of Education and Teaching*, *5*(1), 190-204.

Güzel, E. & Altay, I. F. (2023). EFL learners' perceptions of song-based classes integrated into English language preparatory program. *International Online Journal of Education and Teaching*, *10*(4). 2466-2479.

Hannaoui, A. E. (2017). Mending students' speaking deficiencies in Moroccan EFL classes. *Arab World English Journal*, *8*(1), 386-401. https://dx.doi.org/10.24093/awej/vol8no1.27

Housen, A., Kuiken, F., & Vedder, I. (Eds.). (2012). *Dimensions of L2 performance and proficiency, Complexity, accuracy, and fluency in SLA*. John Benjamins Publishing Company.

Jain, J.; Luaran, J. & Wenzhen, Y. (2024) Harmonizing fluency: The impact of English songs on speaking skills among rural junior high school students in China. *Journal of Tianjin University Science and Technology*, *57*(2), 219-229. https://doi.org/10.5281/zenodo.10690828

Keomany, B., Kanokkamalade, V., & Klinchan, N. (2020). The problems of English speaking of Grade 12 students at Paktha Secondary School in Laos. *Journal of MCU Humanities Review*, *6*(2), 277-287.

Khan, S. (2010). *Strategies and spoken production on three oral communication tasks: A study of high and low proficiency EFL learners* [Doctoral dissertation, Universitat Autònoma de Barcelona].

Khan, T. A. (2024). The impacts of English songs on the pronunciation of ESL learners. *Research Journal of Language and Literacy Studies*, *3*(1), 59-80.

Kim, H. J., Chong, H. J., & Lee, M. (2024). Music listening in foreign language learning: perceptions, attitudes, and its impact on language anxiety. *Frontiers in Education*, *9*, 1-13. https://doi.org/10.3389/feduc.2024.1372786

Kirkpatrick, A. (2012). English in ASEAN: Implications for regional multilingualism. *Journal of Multilingual and Multicultural Development*. *33*(4), 331-344. https://doi.org/10.1080/01434632.2012.661433

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.

Lan, D. T. H., Van, D. T., & Huong, N. T. (2020). Use of pop songs in enhancing English speaking Skills of first-year students at Thai Nguyen University of Economics and Business Administration-TNU. *IOSR Journal of Research & Method in Education*, *10*(2), pp. 29-36.

Ludke, K. M. (2009). *Teaching foreign languages through songs*. University of Edinburgh.

Lume, L., Anggraini, S. W. P., & Laili, U. (2023). The effect of singing method in teaching vocabulary to improve students' speaking skill at VIII class of MTs Plus Nurul Madinah NW Repok Asmari in academic year 2022-2023. *Linguistics and English Language Teaching Studies*, *4*(1), 53-63.

Mobbs, A., & Cuyul, M. (2018). Listen to the music: Using songs in listening and speaking classes. *English Teaching Forum*, *56*(1), 22-29.

Muhamad, N., & Rahmat, N. H. (2020). Investigating challenges for learning English through songs. *European Journal of English Language Teaching*, *6*(1), 1-15. http://dx.doi.org/10.46827/ejel.v6i1.3270

Nety, N. (2022). A study of students' reluctance to speak English at English Department of Unidayan. *English Education Journal*, *8*(1), 45-51. https://doi.org/10.55340/e2j.v8i1.859

Nipattamanon, J. (2018). Utilizing English pop songs with karaoke application among young adults to enhance connected speech production. *The New English Teacher*, *12*(2), 15-31.

Normawati, A., Nugrahaeni, D. A., Manggolo, N. S. K. H., & Susanto, A. I. F. (2023). EFL learners' difficulties in speaking English. *English Language and Education Spectrum*, *3*(1), 1-9. https://doi.org/10.53416/electrum.v1i1.116

Palupi, D., Prasetyaningsih, E. T., & Bilandari, S. (2022). Improving young learners' pronunciation skill through songs. *ETERNAL*, *13*(1), 21-31.

Phisutthangkoon, K. (2024). Thai EFL university students' beliefs in English-speaking fluency. *rEFLections*, *31*(2), 478-500.

Phoowong, S. (2016). Development of English pronunciation skills through English songs. *Academic Journal of Humanities and Social Sciences*, *8*(Special Volume), 189-202.

Purwati, D., Ubaidillah, M. F., & Restall, G. C. (2023). "Sorry, I can't speak": English teachers' challenges of teaching EFL speaking in an Indonesian vocational high school sector. *MEXTESOL Journal*, *47*(1), 1-9.

Rahayu, Y. S., & Rifansyah, T. (2022). The effectiveness of English songs to improve students in listening comprehension. *ELang An English Language Education Journal*, *7*(1), 8-23.

Ratnasari, A. G. (2020). EFL Students' challenges in learning speaking skills: A case study in mechanical engineering. *Journal of Foreign Language teaching and learning*, *5*(1), 20-38 https://doi.org/10.18196/ftl.5145

Riensumettharadol, M. (2019). Effectiveness of out-of-class instructional media integrated songs to enhance vocabulary knowledge of EFL university students. *International Journal of Arts & Sciences*, *12*(1), 147-163.

Saldiraner, G., & Cinkara, E. (2021). Using songs in teaching pronunciation to young EFL learners. *PASAA*, *62*(1), 119-141. https://doi.org/10.58837/CHULA.PASAA.62.1.5

Santiwatthanasiri, T. (2018). *Factors affecting the English speaking ability of Thai university students* [Master's thesis, Thammasat University]

Scrivener, J. (2011). *Learning teaching: The essential guide to English language teaching*. Oxford: Macmillan Education.

Sha'ar, M. Y. M. A., & Boonsuk, Y. (2021). What hinders English speaking in Thai EFL learners? Investigating factors that affect the development of their English speaking skills. *MEXTESOL Journal*, *45*(3), 1-16.

Simpson, A. J. (2015, March 4). How to use songs in the English language classroom. *British Council Voices Magazine*. https://www.britishcouncil.org/voices-magazine/how-use-songs-english-language-classroom

Songkhro, J. (2023). Achievement of using English songs for enhancing English vocabulary knowledge and listening strategies of English as a foreign language second-year students majoring in English for communication at a public university in the lower southern region of Thailand. *Pathumthani University Academic Journal*, *15*(2), 39-53.

Suoc, M. F., Lastino, J., Lukingan, J., Baraoil, M. R., & Pelila, J. R. (2025). How is English as a lingua franca (ELF) transforming education in ASEAN universities? *Journal of Multidisciplinary Research and Development*, *2*(1), 102-115. https://doi.org/10.56916/jmrd.v2i1.1024

Tasnim, Z. (2022). Songs for EFL/ESL class: how to teach listening skill through music. *MEXTESOL Journal*, *46*(3), 1-12. https://doi.org/10.61871/mj.v46n3-16

Tilwani, S. A., Amini MosaAbadi, F., Shafiee, S., & Azizi, Z. (2022). Effects of songs on implicit vocabulary learning: Spoken-form recognition, form-meaning connection, and collocation recognition of Iranian English as a foreign language learners. *Frontiers in Education, 7*, Article 797344. https://doi.org/10.3389/feduc.2022.797344

Triwittayayon, T., & Sarobol, N. (2019). Factors enhancing English speaking ability: Perspectives from Thai high school students and their teachers. *ThaiTESOL Journal*, *31*(1), 49-64.

VEED.IO. (n.d.). *Online video editing and transcription platform.* https://www.veed.io

Waluyo, B. (2019). Examining Thai first-year university students' English proficiency on CEFR levels. *The New English Teacher*, *13*(2), 51-71.

Winnie, W., Leong, H. J., Badiozaman, I. F., & Yap, A. (2023). Negotiating the challenges in speaking English for Indonesian undergraduate students in an ESL university. *Studies in English Language and Education*, *10*(2), 822-840. https://doi.org/10.24815/siele.v10i2.26563

Xi, X. (2021). English teachers' perspectives on using music in English language teaching in Thai secondary schools. *Theory and practice in language studies*, *11*(2), 145-150. https://doi.org/10.17507/tpls.1102.05

Yaorm, P. (2019). *English speaking problems of undergraduate students in a public university* [Independent study paper, Thammasat University].

Yuh, A. H. & Kaewurai, W. (2021). An investigation of Thai students' English-speaking problems and needs and the implementation collaborative and communicative approaches to enhance students' English-speaking skills. *The Golden Teak: Humanity and Social Science Journal*, *27*(2), 91-107.

Zaharani, H. (2023). The effect of using songs to elevate students' vocabulary skill at SMP Yabes School. *Journal of Classroom Action Research*, *2*(2), 52-65. https://doi.org/10.52622/jcar.v2i2.164

Zamin, A. A. M., Adzmi, N. A. H., & Mohamad, M. (2020). Learning vocabulary through songs: A study on the role of music in teaching verbs. *Humanities & Social Sciences Reviews*, *8*(1), 550-557. https://doi.org/10.18510/hssr.2020.8167

Zhang, Y., Bailles, F., & Prieto, P. (2023). Singing songs facilitates L2 pronunciation and vocabulary learning: A study with Chinese adolescent ESL Learners. *Languages*, *8*(3), Article 219. https://doi.org/10.3390/languages8030219

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.144-158
DOI: https://dx.doi.org/10.24093/awej/vol16no3.8

Harnessing AI in CALL: Theories, Pedagogies, and Applications for Adult Language Learning

Lakshmi Kakumanu
Faculty of English Department, Al-Baha University, Saudi Arabia
Email: lakshmi@bu.edu.sa

Received: 06/01/2025 Accepted:09/04/2025 Published: 09/20/2025

Abstract

The growing integration of Artificial Intelligence in Computer-Assisted Language Learning is reshaping the landscape of second language acquisition, particularly, for adult learners seeking flexible, personalized, and effective language instruction. This paper explores the pedagogical impact of AI-enhanced CALL systems through the lens of three foundational theoretical frameworks: Sociocultural Theory, Cognitive Load Theory, and Interactionist perspectives. These perspectives offer complementary insights into how AI technologies, such as intelligent tutoring systems, natural language processing, and adaptive learning platforms, work. Which support learner autonomy, engagement, and communicative competence through personalized feedback, adaptive learning paths, and interactive practice. Drawing on recent empirical studies, the paper highlights both the transformative potential and practical challenges of implementing AI in adult language education. Including concerns over data privacy, teacher readiness gaps, and the risks of algorithmic bias that necessitate mitigation to ensure equitable and effective adoption. Building on these insights, the paper proposes a framework for ethically and pedagogically grounded AI integration to guide researchers, educators, and developers in designing responsible and impactful CALL practices. Ultimately, this study contributes to the evolving discourse on technology-enhanced language learning by providing theoretical depth and practical guidance for leveraging AI in adult second-language contexts.

Keywords: artificial intelligence, computer-assisted language learning, cognitive load theory, second language acquisition, sociocultural theory

How to cite: Kakumanu, L. (2025). Harnessing AI in CALL: Theories, Pedagogies, and Applications for Adult Language Learning. *Arab World English Journal, 16*(3): 144-158. https://dx.doi.org/10.24093/awej/vol16no3.8

Introduction

In today's interconnected and multilingual world, the demand for effective methods of Second Language Acquisition (SLA) continues to rise, particularly among adults engaged in higher education, professional growth, and lifelong learning. Unlike younger learners, adults often bring diverse linguistic backgrounds, shifting motivations, and limited time for consistent study. These realities make language learning more complex, while also increasing the value of Computer-Assisted Language Learning (CALL). AI-driven CALL systems, in particular, respond to these challenges by offering personalized pacing, targeted support, and on-demand practice—options that traditional, uniform approaches struggle to provide (Al-Said, 2022).

Recent advances in Artificial Intelligence have further reshaped CALL environments, introducing levels of personalization and responsiveness that were previously unattainable. Natural Language Processing (NLP), machine learning, and Intelligent Tutoring Systems (ITS) can adapt materials to individual needs, track progress dynamically, and deliver feedback that mimics human responsiveness. These tools are now being used in both formal classrooms and informal learning spaces, and their potential has attracted growing attention within Arab higher education contexts, where issues of teacher readiness and learner autonomy remain particularly salient (Al-Khalifa & Alharbi, 2023). Despite these advances, it is necessary to evaluate AI-enhanced CALL not only for its pedagogical promise but also through critical theoretical and ethical lenses.

This study draws on Sociocultural Theory (Vygotsky, 1978), Cognitive Load Theory (Sweller, 1988), and Interactionist perspectives (Long, 1981; Gass & Mackey, 2007) to examine how AI technologies shape language learning. These frameworks complement one another: Sociocultural Theory explains how AI functions as a mediational tool within the learner's zone of proximal development; Cognitive Load Theory clarifies how adaptive systems help balance mental effort; and Interactionist perspectives illuminate the role of AI-driven feedback and conversational agents in promoting input, negotiation, and output. Together, these perspectives provide a foundation for understanding the interactional, cognitive, and social dimensions of AI-mediated language learning.

This paper emphasizes adult learners, a group still underrepresented in CALL research. By synthesizing theoretical insights and empirical findings, it seeks to:

- Evaluate the alignment of AI-driven CALL with core SLA principles,
- Identify the benefits and limitations of its use in adult learning, and
- Propose recommendations for ethical and practical integration.

Accordingly, the study addresses three research questions:

1. How do AI-enhanced CALL systems respond to the particular needs of adult learners in areas such as personalization, scaffolding, and assessment?
2. In what ways do Sociocultural Theory, Cognitive Load Theory, and Interactionist perspectives explain the processes underlying AI-mediated learning?

3. What benefits, limitations, and ethical issues arise when AI is incorporated into adult SLA?

The sections that follow present a theoretical review, examine empirical evidence, discuss challenges, and outline future directions for AI in CALL, contributing to broader debates at the intersection of artificial intelligence, adult education, and second language pedagogy.

Literature Review: Theoretical Frameworks in CALL

The effectiveness of Computer-Assisted Language Learning (CALL) is increasingly understood through well-established theories of language acquisition and instructional design. Among these, Sociocultural Theory, Cognitive Load Theory, and Interactionist Perspectives provide robust foundations for designing and evaluating AI-enhanced CALL systems, particularly for those who require adaptable, efficient, and socially meaningful learning environments (Golonka et al., 2014; Reinders & White, 2016).

Sociocultural Theory and CALL

Rooted in the work of Vygotsky (1978), Sociocultural Theory emphasizes the centrality of social interaction, mediation, and cultural context in learning. A cornerstone of this theory is the Zone of Proximal Development (ZPD)—the gap between what a learner can do independently and what they can achieve with appropriate guidance. In CALL contexts, this guidance often takes the form of technological scaffolding, such as adaptive feedback, guided practice, and interaction with AI-based virtual agents or peer collaboration platforms (Blake, 2016; Kukulska-Hulme & Viberg, 2018).

Research supports the relevance of Sociocultural Theory in CALL. For example, Lai and Li (2011) found that adult ESL learners using social platforms such as blogs and wikis developed language skills more effectively when peer interaction and collaborative tasks were embedded. Warschauer and Kern (2020) further argue that digital environments replicate key social elements of the ZPD, as learners engage with both human and machine-based facilitators who help shape their understanding.

Recent findings confirm that AI integration can significantly increase learner motivation and satisfaction through personalized interaction and feedback (Chen et al., 2025). AI uniquely operationalizes SCT principles by providing dynamic, just-in-time scaffolding based on learner input, directly mediating within the ZPD. Intelligent tutoring systems, conversational agents, and context-aware feedback tools extend learners' capabilities while tailoring support to their immediate needs. For adult learners, who often balance learning with professional and personal commitments, such self-paced ZPD support ensures that scaffolding is available whenever and wherever they engage with the system, making socially mediated learning accessible beyond the classroom.

Cognitive Load Theory and Multimedia Learning in CALL

Cognitive Load Theory (CLT), proposed by Sweller (1988), posits that learning is most effective when instructional materials are designed to minimize unnecessary mental effort and optimize working memory usage. For CALL environments—especially those enhanced by AI and multimedia tools—this theory underscores the importance of instructional design that integrates modality, coherence, and segmenting principles (Mayer, 2009).

AI-driven CALL platforms often incorporate text, audio, animation, and interactive assessments. While such tools have the potential to deepen comprehension, poorly designed interfaces can overwhelm learners, especially adults returning to formal education after a long break. Studies by Plass and Jones (2021) and Makransky and Mayer (2022) show that learners benefit from platforms that balance information across modalities, such as pairing spoken explanations with synchronized visuals rather than overloading one sensory channel.

More recent research demonstrates that AI systems capable of adaptive sequencing and personalized content reduction can effectively manage intrinsic and extraneous cognitive load, enhancing focus and retention in online learning contexts (Al-Saadi et al., 2024; Zhang & Li, 2025). AI can tailor complexity, eliminate redundancy, and adapt content pacing to learner readiness. These features are particularly beneficial for adult learners—often facing cognitive overload when re-engaging academically—because they manage cognitive load to foster sustainable learning without overwhelming the learner.

Interactionist Perspectives on Language Acquisition in CALL

Interactionist theories—advanced by Long (1981) and Gass and Mackey (2007)—emphasize that interactional feedback, negotiation of meaning, and output opportunities are essential for second language development. Learners must engage in meaningful conversation, process comprehensible input, and be allowed to modify their output based on corrective feedback.

CALL environments, especially those powered by AI, align well with these principles. AI chatbots, virtual speaking partners, and automated evaluators simulate real-time dialogue, analyze learner input, and generate individualized feedback that promotes language restructuring (Godwin-Jones, 2023; Li, 2022). Such interaction-rich environments foster linguistic noticing and metalinguistic reflection, which are essential mechanisms in SLA (Schmidt, 1990).

Recent studies highlight how AI conversational agents provide scalable, interactive opportunities for language practice, feedback, and negotiation of meaning that were previously resource-prohibitive (Nguyen, 2023; Wang & Chen, 2024). For adult learners who may experience anxiety or self-consciousness in live classrooms, AI-driven interaction creates a safe practice space where they can rehearse workplace, academic, or social communication scenarios with confidence.

Synthesis and Implications

Together, these three theoretical perspectives offer a comprehensive foundation for understanding how AI-driven CALL systems support second language acquisition. Sociocultural

Theory informs the importance of interactive, scaffolded learning; Cognitive Load Theory ensures multimedia instruction remains cognitively efficient; and Interactionist approaches validate the need for practice-rich, feedback-driven environments.

Notably, the combination of these frameworks provides a more holistic evaluative lens than any single one alone. SCT explains the mediational and social aspects of AI scaffolding, CLT addresses the cognitive efficiency of adaptive instruction, and Interactionist theory captures the communicative value of AI-driven feedback and negotiation. For adult learners, whose needs span self-paced support, cognitive efficiency, and safe communicative practice, this integrated framework ensures that AI-enhanced CALL is analyzed in a way that reflects the full spectrum of their learning realities.

AI Technologies in Language Learning

The integration of Artificial Intelligence (AI) in Computer-Assisted Language Learning (CALL) has introduced innovative, adaptive, and learner-centred solutions to address the diverse needs of adult language learners. Among the most impactful AI technologies in CALL are Natural Language Processing (NLP), Machine Learning (ML), and Intelligent Tutoring Systems (ITS). These technologies enhance learner autonomy, provide timely feedback, and personalize instructional delivery—features especially valuable for adults balancing education with professional and personal responsibilities.

Natural Language Processing (NLP)

NLP enables machines to interpret, analyze, and generate human language. In language education, it powers tools for grammar checking, text summarization, error detection, and writing enhancement. Applications such as Grammarly, Write & Improve, and Turnitin Revision Assistant provide immediate, form-focused feedback that helps learners independently identify and correct syntactic and lexical errors (Chen et al., 2024; Katinskaia, 2025).

Recent innovations include chatbots and conversational agents that simulate realistic dialogue through NLP algorithms. These tools deliver non-judgmental, interactive speaking practice, which is crucial for adults who may experience anxiety in classroom interactions. Studies demonstrate that chatbot-based practice improves fluency, pragmatic competence, and motivation (Karataş et al., 2024; Godwin-Jones, 2023). However, NLP technologies still struggle with pragmatic subtleties and context-dependent errors, such as tone, irony, or cultural nuances, which means that adult learners may require supplementary guidance when dealing with more complex communicative contexts.

Machine Learning and Adaptive Learning Systems

Machine Learning (ML), a core branch of AI, empowers CALL systems to analyze learner data and adapt instruction in real-time. ML algorithms can detect patterns in grammatical error types, vocabulary acquisition speed, and task completion behaviors, and then adjust instructional

trajectories accordingly (Chen et al., 2025). This creates personalized pathways that allow learners to progress at their own pace and revisit challenging material when needed.

Beyond personalization, ML-driven platforms utilize predictive analytics to identify early signs of disengagement or underperformance, enabling timely intervention and differentiated instruction. Research confirms that such systems enhance satisfaction and proficiency outcomes, especially in heterogeneous classrooms (Mohsen et al., 2023; AlAqlobi et al., 2024). However, ML is highly dependent on the quality and quantity of learner data, and flawed or biased datasets can result in ineffective or even misleading adaptation. For adults with highly diverse linguistic and professional backgrounds, ensuring representative data is particularly important.

Intelligent Tutoring Systems (ITS)

Intelligent Tutoring Systems (ITS) provide individualized instruction and feedback that mirrors one-on-one tutoring. Grounded in cognitive and pedagogical models, ITS tracks learner performance and applies scaffolding strategies, adjusting task complexity as proficiency increases (Makransky & Mayer, 2022; Plass & Jones, 2021). Examples include platforms offering dynamic listening comprehension tasks, grammar explanations, and interactive reading support.

For adult learners with varied educational and linguistic backgrounds, ITS provides tailored pathways that address both form and meaning while encouraging reflective learning and metacognitive awareness (Godwin-Jones, 2023; Tafazoli, 2023). Recent ITS developments integrate multimodal input (visual, auditory, and textual cues) and immersive technologies such as Augmented Reality (AR) and speech recognition, enabling authentic communicative practice (Caetano et al., 2024). Still, creating truly "intelligent" pedagogical models that capture the complexity of human tutoring remains a challenge, as ITS often relies on scripted scaffolding rather than nuanced instructional judgment.

Synthesis

AI technologies in CALL collectively reshape language learning—particularly for adult learners. NLP enhances written and spoken communication but still struggles with deeper pragmatics. ML facilitates personalized, data-driven pathways, yet it depends heavily on robust datasets. ITS provides real-time scaffolding, while facing limitations in replicating human-level pedagogical intelligence. Importantly, these technologies often work in tandem: NLP powers the interface for written and spoken interaction, ML drives the adaptation engine using data parsed by NLP, and ITS integrates both to deliver holistic tutoring experiences. When thoughtfully combined, they promote autonomy, engagement, and long-term language development. At the same time, effective implementation requires careful pedagogical design, digital literacy support, and ongoing research into their long-term impact.

Empirical Studies on AI-Enhanced CALL

The effectiveness of AI-enhanced Computer-Assisted Language Learning (CALL) has been increasingly validated through a range of recent empirical studies. These studies provide valuable insights into the pedagogical benefits and implementation outcomes of AI technologies across diverse language learning settings, particularly among learners in higher education and continuing education contexts.

Efficacy of AI in Language Learning

Numerous studies have shown that AI-supported CALL environments enhance language proficiency in grammar, writing, pronunciation, and vocabulary acquisition. For example, a randomized controlled trial by Chen et al. (2024) investigated a multi-strategy AI-powered EFL writing system that utilizes deep learning to provide personalized feedback on grammar and content. Learners in the experimental group demonstrated statistically significant improvements in writing fluency, coherence, and grammatical accuracy compared to a control group receiving traditional feedback.

Similarly, a meta-analysis by Chen et al. (2025), synthesizing findings from 32 studies on AI-enabled assessment tools, concluded that intelligent evaluation platforms yielded consistent, statistically significant gains in both formative and summative performance indicators in L2 learning. The authors emphasized that the immediacy, consistency, and adaptiveness of AI feedback were central to these outcomes, though effect sizes varied depending on task type and learner proficiency.

Learner Outcomes and Engagement

AI-enhanced CALL systems also influence learner engagement, autonomy, and motivation—key drivers of adult SLA. A mixed-methods study by Karataş et al. (2024) examined ChatGPT-based instruction for adult EFL learners and reported moderate effect sizes on motivation and interactional competence. Participants noted a reduction in anxiety and an increased willingness to experiment with language, attributing these improvements primarily to the tool's safe and responsive feedback environment.

Godwin-Jones (2023) echoed these findings, observing that AI-mediated conversational agents and adaptive writing support promoted metacognitive awareness and learner autonomy. These affordances are especially important for adults, who value tools that accommodate diverse prior knowledge and support reflective, self-paced practice. In addition, learner log data from AI platforms often show increased time-on-task and higher completion rates compared with traditional digital tools, indicating observable improvements in sustained engagement (Mohsen et al., 2023).

Case Studies of Successful Implementation

Case-based research further validates the role of AI in enhancing language learning success. A longitudinal study by AlAqlobi et al. (2024) evaluated an AI-integrated platform combining pronunciation feedback, grammar correction, and vocabulary scaffolding among working professionals in a business English program. Over the course of 12 weeks, participants demonstrated measurable improvements in oral fluency and grammatical range, accompanied by consistently high satisfaction ratings.

Similarly, a case study by Katinskaia (2025) examined AI-supported CALL in an adult ESL program at a Korean university. The integration of chatbots and automated writing tutors resulted in improved self-correction strategies and increased linguistic complexity in writing. The study emphasized that clear instructional design and well-defined task goals are crucial for effective AI use.

Other studies highlight the role of AI in differentiated instruction and inclusivity. For example, adaptive systems that adjust to learners' pace, proficiency, and preferred input modes (e.g., text, audio, visuals) have been shown to narrow gaps in mixed-ability classrooms (Caetano et al., 2024; Tafazoli, 2023).

Synthesis

Collectively, these empirical studies confirm that AI-enhanced CALL systems contribute to improved linguistic outcomes, increased learner engagement, and greater instructional efficiency. However, several limitations are common: many studies focus on short-term interventions, target specific skills (e.g., writing or fluency) rather than holistic proficiency, or rely on self-reported engagement metrics in conjunction with performance data. While findings are promising, more longitudinal, large-scale research is needed to understand the sustained and comprehensive effects of AI integration in adult language learning.

Benefits of Integrating AI into CALL

The integration of Artificial Intelligence (AI) into Computer-Assisted Language Learning (CALL) environments has generated significant pedagogical benefits in diverse educational contexts. These benefits, however, are potential rather than guaranteed, and their realization depends heavily on the quality of implementation, alignment with pedagogical principles, and teacher facilitation. Among the most notable advantages are improvements in learner autonomy and motivation, the provision of immediate and personalized feedback, and the scalability of high-quality language instruction across varied learning settings.

Enhanced Learner Autonomy and Motivation

AI-based CALL platforms empower learners to take control of their own learning process. Through adaptive content delivery, self-paced modules, and goal-setting features, these tools support self-directed learning—an application that aligns closely with Sociocultural Theory, where

AI acts as a mediational tool within the Zone of Proximal Development, and with Machine Learning, which adapts to individual learner profiles. For adult learners managing multiple responsibilities, such self-paced scaffolding is especially beneficial (Godwin-Jones, 2023). The ability to track progress, revisit challenging tasks, and receive ongoing support fosters intrinsic motivation and sustained engagement; however, studies indicate that these gains are often moderate in effect size and depend on the thoughtful integration of AI into broader learning design (Karataş et al., 2024).

Immediate and Personalized Feedback

One of the core strengths of AI in CALL is its ability to provide real-time, individualized feedback. Tools equipped with Natural Language Processing (NLP) detect grammatical errors, suggest lexical alternatives, and assess pronunciation instantly, thereby operationalizing principles of Interactionist Theory, which emphasizes feedback and negotiation of meaning. This immediacy reduces learner frustration and accelerates revision cycles, supporting deeper learning outcomes (Chen et al., 2024, 2025). However, NLP systems still struggle with pragmatic subtleties and cultural appropriateness, occasionally mis flagging correct usage or reinforcing dominant language norms (e.g., privileging U.S. English over regional varieties). Such algorithmic bias highlights the need for teacher oversight and diversified training datasets.

Scalability and Accessibility of Resources

AI technologies also enable CALL to be scaled across regions, populations, and learning formats. Cloud-based platforms can deliver instruction to thousands of learners simultaneously, while Intelligent Tutoring Systems (ITS) simulate individualized tutoring at scale by combining NLP-driven interfaces with ML-powered adaptation. This scalability reflects Cognitive Load Theory principles, as adaptive sequencing ensures learners are not overwhelmed when accessing content remotely. Moreover, AI allows personalization based on proficiency level, cultural background, or learning preferences, enhancing inclusivity (Mohsen et al., 2023; Caetano et al., 2024). Nevertheless, scalability brings challenges: poor infrastructure, insufficient teacher readiness, and unchecked algorithmic models can undermine accessibility. Mitigation requires investment in digital literacy for educators and learners, as well as transparent oversight of AI-driven systems.

Synthesis

Overall, the benefits of AI integration in CALL are compelling but closely tied to the frameworks and technologies discussed earlier. SCT explains how AI enhances autonomy, CLT ensures scalable systems manage cognitive load, and Interactionist Theory clarifies the role of real-time feedback. However, these benefits are interdependent: personalization relies on high-quality data (machine learning), feedback depends on NLP accuracy, and scalability requires robust integration of ITS. Challenges such as algorithmic bias, teacher readiness, and infrastructural

limitations are interconnected and require coordinated solutions to address them effectively. Mitigation involves teacher reskilling—not just general training, but specific competencies in data literacy, critical evaluation of AI outputs, and the ability to integrate AI feedback with human instruction. Addressing these challenges is essential to ensure that AI-enhanced CALL fulfils its promise for adult learners.

Challenges and Considerations in Integrating AI into CALL

While the integration of Artificial Intelligence (AI) in Computer-Assisted Language Learning (CALL) offers numerous benefits, it also presents several pedagogical, technical, and ethical challenges that must be addressed to ensure effective and equitable implementation—especially in adult education contexts. These challenges relate to data privacy, algorithmic bias, teacher readiness, accessibility issues, and the evolving balance of roles between teachers, learners, and AI systems.

Data Privacy and Ethical Concerns

AI-driven CALL systems rely heavily on collecting and analyzing large volumes of learner data, including language performance, behavioral patterns, and personal identifiers. This raises concerns about privacy, consent, and security, particularly in online environments. Studies have emphasized the need for transparency in how learner data is stored, processed, and used for algorithmic decision-making (AlAqlobi et al., 2024; Tafazoli, 2023). Without robust governance policies, adult learners—especially those less familiar with digital platforms—may be hesitant to fully engage with AI systems. Mitigation requires more transparent consent processes, encrypted storage, and institution-wide policies on the ethical use of data.

Algorithmic Bias and Content Fairness

AI tools trained in limited or non-diverse language corpora can reproduce biases in grammar correction, vocabulary suggestions, or accent evaluations, potentially disadvantaging learners from specific linguistic or cultural backgrounds. Karataş et al. (2024) caution that unchecked biases may reinforce stereotypes and harm learner confidence. For example, pronunciation engines may penalize regional accents, or essay scorers may undervalue non-Western rhetorical styles. Addressing these issues requires diverse training datasets, regular fairness audits, and the inclusion of culturally varied benchmarks.

Teacher Training and Digital Readiness

Despite the growing availability of AI technologies, many educators lack the technical skills and pedagogical readiness to integrate these tools effectively. As Godwin-Jones (2023) notes, resistance often stems from insufficient exposure during professional development. Teachers need not only general training but also specific competencies, such as data literacy, the ability to critically interpret AI-generated feedback, and strategies for blending automated feedback with

human judgment. Mitigation involves structured professional development, peer-learning communities, and the integration of AI modules into teacher education curricula.

Infrastructure and Accessibility Barriers

Access to reliable internet, modern devices, and platform-specific software remains a challenge, particularly in under-resourced regions. While AI offers scalability, it risks widening the digital divide without parallel investment in infrastructure. Caetano et al. (2024) emphasize the importance of accessibility features—such as multilingual interfaces, text-to-speech capabilities, and offline modes—for promoting inclusivity. Mitigation requires policy-level support, institutional investment in devices and bandwidth, and a design that prioritizes universal access.

Evolving Roles and Collaboration

A frequently overlooked challenge lies in defining the optimal collaboration among teachers, learners, and AI systems. AI can mediate feedback and adapt pacing, but it cannot replace the cultural, ethical, and motivational roles that teachers play. Adult learners particularly benefit from teachers who act as facilitators, guiding them in interpreting AI feedback, setting realistic goals, and maintaining critical awareness of system limitations. Moving forward, compelling CALL will require a partnership model, where AI handles scalability and personalization, teachers provide human mediation and ethical oversight, and learners actively co-direct their learning paths.

Synthesis and Future Research Directions

Addressing these challenges is crucial to ensuring that AI in CALL functions not only as a technological innovation but also as a pedagogically sound and ethically responsible tool. Importantly, these issues are interconnected: teacher readiness affects how bias is managed, infrastructure shapes accessibility, and privacy policies influence learner trust. Future research should move beyond short-term outcomes to explore:

- How can teacher–AI–learner collaboration be structured to maximize learning without overburdening educators?
- What long-term effects do AI-driven CALL systems have on holistic proficiency, beyond targeted skills?
- How can algorithmic bias be detected, measured, and systematically mitigated in multilingual and multicultural contexts?
- What models of professional development best prepare teachers for AI-integrated classrooms in adult education?

By addressing these questions, researchers and practitioners can ensure AI-enhanced CALL evolves responsibly and inclusively.

Future Directions in AI-Enhanced CALL

As Artificial Intelligence (AI) continues to evolve, its potential to reshape language education through Computer-Assisted Language Learning (CALL) is both promising and

expansive. Future developments are expected to focus on greater personalization, immersive technologies, and ethical design frameworks that ensure equitable and learner-centered experiences.

Greater Personalization Through Multimodal AI

Future AI-driven CALL platforms will increasingly incorporate multimodal learning environments, combining speech, text, video, and gesture to create more context-rich, interactive learning experiences (Chen et al., 2025). AI algorithms will better interpret individual learner profiles, offering highly customized feedback and dynamic learning paths that adjust in real-time (Plass & Jones, 2021).

Integration of Immersive Technologies

Emerging technologies such as Augmented Reality (AR), Virtual Reality (VR), and haptic feedback systems are likely to redefine engagement in language learning. For instance, AR applications allow learners to practice vocabulary and communication skills in location-based or task-specific contexts, enhancing retention and cultural immersion (Caetano et al., 2024; Makransky & Mayer, 2022).

Ethical and Inclusive AI Design

As AI adoption accelerates, attention must turn to ensuring transparency, fairness, and inclusiveness in system design. Future CALL systems will need built-in mechanisms to audit for algorithmic bias, protect learner privacy, and promote universal accessibility—particularly for marginalized adult learners (Godwin-Jones, 2023; Tafazoli, 2023).

Conclusion

This study set out to examine how Artificial Intelligence can support Computer-Assisted Language Learning for adult learners. By drawing on Sociocultural Theory, Cognitive Load Theory, and Interactionist perspectives, it has been shown that tools such as Natural Language Processing, Machine Learning, and Intelligent Tutoring Systems can help manage cognitive load, provide scaffolding, and create more opportunities for meaningful practice.

The evidence points to measurable gains in accuracy, fluency, and learner engagement. At the same time, many studies remain short-term and skill-focused, which leaves important questions about long-term impact unanswered. Concerns over bias, data privacy, and uneven teacher readiness also remind us that the success of AI in CALL depends less on the technology itself than on how it is designed and implemented.

To move forward, this paper has proposed a framework that emphasizes inclusivity, pedagogical alignment, teacher competence, and learner agency. Within this model, AI adds scale and efficiency, teachers provide cultural and ethical judgment, and learners play an active role in shaping their progress.

In short, AI holds real promise for adult language education, but its value lies in its thoughtful and ethical use rather than novelty. When educators and learners work without technology, CALL can evolve into a more accessible, responsive, and human-centered practice.

Funding: This research did not receive any specific grant or funding support.
Acknowledgements: Not Applicable.
Conflicts of Interest: The author declares that there are no conflicts of interest regarding the publication of this manuscript.
Authenticity: This manuscript is an original work.
Artificial Intelligence Statement: During the preparation of this work, the author used Grammarly AI to check the grammar. After using this tool, the author reviewed and edited the content as needed, taking full responsibility for the content of the publication.

About the author

Dr. Lakshmi Kakumanu is a faculty member in the English department at Al-Baha University, Saudi Arabia. She holds a Ph.D. in Applied Linguistics and has over 20 years of experience in teaching English at the tertiary level. Her research interests focus on second language acquisition, technology-enhanced language learning, and collaborative learning practices in higher education.
ORCID ID -- https://orcid.org/0009-0001-5122-0543

References

Al-Said, R. (2022). Artificial intelligence and CALL: Opportunities and challenges in Arab higher education. *Arab World English Journal (AWEJ), 13*(4), 254–269. https://doi.org/10.24093/awej/vol13no4.17

Al-Khalifa, H., & Alharbi, S. (2023). Teacher readiness and learner autonomy in AI-assisted CALL: Evidence from Saudi universities. *Arab World English Journal (AWEJ), 14*(1), 98–115. https://doi.org/10.24093/awej/vol14no1.7

AlAqlobi, O., Alduais, A., Alasmari, M., & Qasem, F. (2024). Artificial intelligence in language acquisition: A balancing act of potential and challenges. *Forum for Linguistic Studies, 6*(6), 101–117. https://doi.org/10.54254/2753-7048/6/20240818

Blake, R. J. (2016). Technology and the four skills. In F. Farr & L. Murray (Eds.), *The Routledge handbook of language learning and technology* (pp. 102–115). Routledge.

Caetano, A., Lawson, A., Liu, Y., & Sra, M. (2024). ARLang: An outdoor augmented reality application for Portuguese vocabulary learning. *arXiv preprint arXiv:2411.05211.* https://doi.org/10.48550/arXiv.2411.05211

Chen, B., Li, J., Zhang, Y., & Lim, C. P. (2024). A multi-strategy computer-assisted EFL writing learning system with deep learning incorporated. *Journal of Educational Computing Research, 62*(1), 45–69. https://doi.org/10.1177/07356331231152607

Chen, B. et al. Zhang, Y., Jia, J., Liang, M., Cha, Y., & Lim, C. P. (2025). A systematic review and meta-analysis of AI-enabled assessment in language learning. *Journal of Computer Assisted Learning, 41*(2), 221–245. https://doi.org/10.1111/jcal.12786

Gass, S. M., & Mackey, A. (2007). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 175–200). Routledge.

Godwin-Jones, R. (2023). Emerging Technologies: AI Language Tutors and Their Evolving Role in Language Learning *Language Learning & Technology, 27*(1), 1–14. https://doi.org/10.10125/73467

Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer Assisted Language Learning, 27*(1), 70–105. https://doi.org/10.1080/09588221.2012.700315

Karataş, F., Abedi, F. Y., & Gunyel, F. O. (2024). Incorporating AI in Foreign Language Education: The Effect of ChatGPT on Learners. *Education and Information Technologies, 29*(4), 19343–19366. https://doi.org/10.1007/s10639-024-12987-1

Katinskaia, A. (2025). An overview of artificial intelligence in computer-assisted language learning. *arXiv preprint arXiv:2505.02032*. https://doi.org/10.48550/arXiv.2505.02032

Kukulska-Hulme, A., & Viberg, O. (2018). Mobile collaborative language learning: State of the art. *British Journal of Educational Technology, 49*(2), 207–218. https://doi.org/10.1111/bjet.12580

Lai, C. (2017). *Autonomous language learning with technology: Beyond the classroom.* Bloomsbury Academic.

Lai, C., & Li, G. (2011). Technology and task-based language teaching: A critical review. *CALICO Journal, 28*(2), 498–521. https://doi.org/10.11139/cj.28.2.498-521

Lee, L. (2020). A systematic review of virtual exchange for second language learning. *Computer Assisted Language Learning, 33*(1–2), 1–28. https://doi.org/10.1080/09588221.2018.1501069

Li, M. (2022). Chatbots for language learning: Perspectives, practices, and challenges. *Language Teaching Research, 26*(5), 869–889. https://doi.org/10.1177/13621688211048969

Long, M. H. (1981). Input, Interaction, and Second-Language Acquisition. *Annals of the New York Academy of Sciences, 379*(1), 259–278. https://doi.org/10.1111/j.1749-6632.1981.tb42014.x

Makransky, G., & Mayer, R. E. (2022). Benefits of Integrating Immersive Virtual Reality into Multimedia Learning *Educational Psychology Review, 34*(1), 235–256. https://doi.org/10.1007/s10648-021-09618-2

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.

Mohsen, M. A., Althebi, S., Alsagour, R., Alzahrani, A., Alghamdi, F., & Altalhi, A. (2023). Forty-two years of computer-assisted language learning research: A scientometric study. *ReCALL, 35*(3), 259–279. https://doi.org/10.1017/S0958344023000126

Plass, J. L., & Jones, L. C. (2021). *Multimedia learning and second language acquisition.* Cambridge University Press.

Reinders, H., & Benson, P. (2017). Research agenda: Language learning beyond the classroom. *Language Teaching, 50*(4), 561–578. https://doi.org/10.1017/S0261444817000182

Reinders, H., & White, C. (2016). 20 years of autonomy and technology: How far have we come and where to next? *Language Learning & Technology*, *20*(2), 143–154. https://doi.org/10.10125/44477

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

Stockwell, G. (2012). Using mobile phones for vocabulary activities: Examining the effect of the platform. *Language Learning & Technology*, *14*(2), 95–110. https://doi.org/10.10125/44284

Tafazoli, D. (2023). Critical appraisal of artificial intelligence-mediated communication. *arXiv preprint arXiv:2305.11897.* https://doi.org/10.48550/arXiv.2305.11897

van Lier, L. (2004). *The ecology and semiotics of language learning: A sociocultural perspective.* Springer.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Warschauer, M., & Kern, R. (2020). *Network-based language teaching: Concepts and practice* (2nd ed.). Cambridge University Press.

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.159-174
DOI: https://dx.doi.org/10.24093/awej/vol16no3.9

Which Engages Students More in English Writing: Teacher+Peer Feedback or Teacher-Only Feedback? -A Study in Chinese Tertiary Education

Min Deng
Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Malaysia
School of Foreign Languages, China West Normal University (CWNU), Nanchong, China
Email: P121758@siswa.ukm.edu.my

Nur Ehsan Mohd Said
Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Malaysia
Corresponding author: nurehsan@ukm.edu.my

Nur Ainil Sulaiman
Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Malaysia
Email: nurainil@ukm.edu.my

Yunyun Zhao
Jiujiang New Town Middle School, Longdu Rd.Shuangliu District, Chengdu, Sichuan, China
Email: 2469397928@qq.com

Received:06/04/2025 Accepted:08/28/2025 Published: 09/20/2025

Abstract

This study investigates the impact of "teacher+peer" mixed feedback on undergraduate English writing engagement, aligning with the emphasis on comprehensive feedback in English Language Teaching and student-centered learning principles outlined by the *College English Curriculum Guide* (2020) in China to answer two main research questions: how does the "teacher+peer" mixed feedback pattern influence engagement in English writing among Chinese tertiary students behaviorally, cognitively, and emotionally and does mixed feedback yield significantly higher engagement than teacher-only feedback? Conducted a 12-week semester with 85 first-year undergraduates, the study compared an experimental group receiving "teacher+peer" feedback with a control group receiving"teacher-only" feedback, utilizing data from pre- and post-intervention questionnaires and post-intervention focus group discussion. Findings reveal that the mixed feedback approach significantly enhances learning engagement across all dimensions—emotional, cognitive, and behavioral—outperforming "teacher-only" feedback. A substantial post-intervention mean increase is observed in the experimental group, in contrast to minimal changes in the control group. It expanded the existing literature on writing feedback, engagement theory, and peer involvement in the ELT domain through empirical data from the Chinese context. Additionally, the pedagogical implications include rotating mixed-ability groups, overcoming cultural hesitancy among reserved Chinese students, and accommodating variations in English proficiency, facilitating the conduction of the research-informed "teacher+peer" writing feedback in similar contexts outside China.

Keywords: EFL writing, Chinese Tertiary Education, "teacher+peer" mixed feedback, writing engagement

How to cite: Deng, M., Said, N. E.M., Sulaiman, N. A., & Zhao, Y. (2025). Which Engages Students More in English Writing: Teacher+Peer Feedback or Teacher-Only Feedback? -A Study in Chinese Tertiary Education. *Arab World English Journal, 16*(3):159-174. https://dx.doi.org/10.24093/awej/vol16no3.9

Introduction

The 2020 College English Guide of China (*Guide*s for short) states that college English writing instruction aims to enable students to "write to express personal opinions on general topics in English; to be able to write English abstracts and papers related to their specialties." (MOE, 2020, p. 9). Concerning the approaches to achieve these goals, the document highlights the importance of fostering "students' initiative, autonomy and strategies in learning," (MOE, 2020, p.33), suggesting that curricula should cater to students' personal interests, academic goals, and future career demands, thus arousing their engagement in learning. However, the gaps still retain, for example, teachers are mostly the only subject of evaluation even until recently although the growing instructional workload and expanding class size in higher education made them already exhausted in China (Shi, 2019), and students scarcely were involved as members of the evaluation group, being unable to conduct self-evaluation and peer feedback thus to boost their engagement in the writing process.

To investigate whether the "teacher+peer" mixed feedback enhances learners' engagement than "teacher-only" feedback in English writing, this study is conducted to achieve these objectives: to identify whether learner writing engagement is impacted by "teacher+learner" feedback behaviorally, cognitively, and emotionally, and to compare the engagement yielded by mixed feedback and teacher-only feedback. The current study seeks answers to the following research questions:

i. How does the "teacher+peer" mixed feedback pattern influence engagement in English writing among Chinese tertiary students behaviorally, cognitively, and emotionally?
ii. Does mixed feedback yield significantly higher engagement than teacher-only feedback?

The "teacher+peer" mixed feedback approach, which integrates teacher guidance with peer remarks, may meet the demands of the updated *Guides*, stating the importance of combining learner-centered teaching and teacher guidance (MoE, 2020) and thus promote learners' engagement in writing. Learning engagement, defined as the degree to which learners actively participate in educational activities and experiences (Reeve, 2012), is evaluated for understanding the learning process and fostering future progress (Ren, 2023). It is considered one of the most pivotal feedback types in tertiary education (Noroozi et al., 2023). Active student engagement impacts academic outcomes (Gao et al., 2019). Thus, strong writing engagement fostered in the feedback process is essential for the effectiveness of writing outcomes.

Review of Literature

In this section, the existing literature about"teacher+peer" mixed feedback, writing engagement, and the combination of the two dimensions is reviewed to lay a foundation for the current study and identify the research gaps.

"Teacher+Peer" Mixed Feedback

Back in the 1960s, teacher feedback dominated the feedback domain in the teaching of writing in English as a second or foreign language (ESL/EFL), and its authority was unanimously recognized by learners. There are various definitions for teacher feedback, which encompass the corrective information provided by teachers to students (Flower, 1979) and a feedback mechanism through which teachers provide comments and suggestions to improve learners' writing (Hyland

& Hyland, 2019). This study defines teacher feedback as a teacher's review of students' writing and the provision of revision information via written feedback (Zhao, 2024). Ferris and Roberts claim that learners prefer teachers' written feedback over other forms of feedback (2001), echoed by Mustafa et al. 's research findings that teacher feedback is the most helpful measure from learners' viewpoints (2022). On the other hand, peer feedback usually refers to students providing oral or written feedback on their peers' writing in pairs or small groups (Yu & Lee, 2016). With other names as peer assessment or peer review, it features several aspects: effectiveness, which is proven by Cai (2011), feasibility (Liu & Ji, 2018), and learners' perceptions and attitudes towards it (Bai & Li, 2020).

Studies largely support combining teacher and peer feedback, namely "teacher+peer" mixed feedback, finding it boosts motivation, reduces anxiety, and improves writing quality, autonomy, and engagement (Yu, 2020). Xing and Wang (2023) rank it as the most effective among feedback sources, and students prefer it over single-source options (Demirel & Enginarlar, 2016; Loan, 2017). However, some studies suggested that the practicality of English writing feedback largely depends on the verified contexts (Phothongsunan, 2023); hence, it is important to apply various feedback to cater to different environments. Additionally, students distrust peers involved in the "teacher +peer" feedback, due to perceived disrespect or incompetence (Zhou et al., 2020), feeling stressed by peer assessment (Stančić, 2021), or questioning peers' language skills, threatening the face-saving norms in Chinese culture (Xue, 2023), and insufficient specialized knowledge and literacy about the topic (Valero Haro et al.,2023). Therefore, this study is conducted to verify its effectiveness in boosting learners' writing engagement.

Writing Engagement

The definition of engagement could be traced back to Ralph Tyler in the 1930s, who viewed it as active participation in learning (Reeve, 2012), evolving through Pace's (1984, P.54) emphasis on "quality of effort". Fredricks et al. (2004) categorized it into behavioral (effort), emotional (affect), and cognitive (strategy) dimensions, which offer a conductible measurement for it. Behavioral engagement manifests learners' behavioral responses to feedback (Tian & Zhou, 2020), mainly reflected by the length of revision and feedback uptake rate (Ranalli, 2021); emotional engagement refers to learners' emotional and attitudinal responses to feedback (Tian & Zhou, 2020), covering interest, value, and affect (Fan & Xu, 2020) and cognitive engagement pertains to feedback awareness, cognitive strategies, and metacognitive strategies (Han & Hyland, 2015). In this study, behavioral engagement encompasses time spent revising and frequency of feedback uptake; cognitive engagement covers the depth of processing feedback, and self-reported effort in understanding and applying it, and emotional engagement includes attitudes toward feedback, motivation, and confidence. Since most existing literature targets single feedback types: teacher, automated, or peer, whereas not sufficient mixed feedback studies (Tian & Zhou, 2020), this study intends to explore the impact of the combination of teacher and student feedback in the context of Chinese Tertiary Education.

Writing Engagement and "Teacher+peer" Mixed Feedback

Research on English writing feedback explores peer training, influencing factors, strategies, interaction, rubrics, and focus, with rising interest in learning engagement's role in English writing effectiveness (Xu & Fan, 2019; Zhang & Hyland, 2018). Studies also examine revision behaviors

(Goldstein, 2016), feedback processing depth (Kim & Bowles, 2019), and emotional responses (Mahfoodh, 2017). However, a few studies have comprehensively integrated emotional, behavioral, and cognitive engagement (e.g., Han & Xu, 2021). Moreover, most research focuses on single feedback types—such as teacher feedback (Zheng et al., 2020), automated feedback (Koltovskaia, 2020), or peer feedback (Geng & Yu, 2023)—while mixed feedback patterns remain underexplored. Although studies on mixed feedback (e.g., Tian & Zhou, 2020) highlight the benefits of combining teacher and peer feedback, the relevance of feedback approaches to writing engagement, particularly comparisons between single feedback (e.g.,"teacher-only") and mixed feedback (e.g., teacher and peer) in higher education, is still largely unexamined.

Methodology

The research design was elaborated in terms of participants, sampling, instrumentation, and procedure.

Design of the Study

This study employed mixed methods (quantitative and qualitative), which allow for a comprehensive analysis of engagement dimensions and student perceptions, aligning with the student-centered framework (MoE, 2020). The use of different data collection methods (questionnaire and focus group) will lead to deeper insight into the understanding of the phenomenon under investigation. As it is a quasi-experimental research, non-random allocation of subjects may potentially introduce selection bias; nevertheless, the convenience sample conducted in the study fits existing classroom structures (42-56 students per class), providing real-world applicability.

Participants

In the experiment, a total of 85 first-year undergraduates in two classes from a Chinese tertiary university with an age span of 18 to 21 participated in this study by obtaining their consent after informing the research purpose and procedures, and the rules of participant autonomy and confidentiality. In social science research, sample sizes of $N \geq 30$ are generally sufficient to produce a normal distribution of sample means (Uttley, 2019). The two classes are from the same school to avoid the difference due to their majors. The control class consisted of 42 students, while the experimental class included 43 students. The total sample of 85 participants exhibited a similar gender ratio of male to female students across both classes, as presented in Table 1. Pre-intervention English writing proficiency was assessed via the College Entrance Examination results in writing, yielding mean scores of 12.33 for the control class and 11.80 for the experimental class (out of a possible maximum score of 25). To confirm equivalence in writing ability, an independent samples t-test was conducted, resulting in a p-value of 0.418, which is greater than the conventional threshold of 0.05. This indicates no statistically significant difference, confirming that the English writing proficiency levels of students in both classes were closely aligned before the experiment. There is another t-test conducted to ensure similar baseline engagements between the two classes.

Table 1 *Demography of research participants*

Class	Number of Students	Gender (Male/Female)	Mean
control class	42	16/26	12.33
experimental class	43	17/26	11.80

Instruments

To explore the research questions and address the aims, two instruments are employed, namely, a questionnaire and a focus group protocol.

a. Questionnaires

This study administered surveys to measure self-reported engagement concerning three dimensions: behavioral (e.g., effort in revising), cognitive (e.g., focus on improvement), and emotional (e.g., enjoyment). The questionnaires were administered to measure these dimensions of participants in response to different feedback patterns. Among them, the questionnaire on writing engagement in the teacher's feedback mode served as the baseline for the subsequent questionnaire on writing engagement in the 'teacher + peer' feedback mode. The latter was adapted from the former by modifying specific wording to reflect the mixed feedback approach. Both questionnaires featured three main similar components: a behavioral, emotional, and cognitive engagement scale with 18 questions.

The behavioral engagement scale contains frequency and time spending, and the emotional engagement scale was constructed by revising existing frameworks, drawing on the three sub-dimensions of emotional engagement by Tsui and Ng (2000), Fan and Xu (2020), and Zhao (2024). They are interest, value, and feeling. Similarly, the cognitive engagement scale was developed regarding established research, including O'Malley and Chamot's (1990) classification of metacognitive strategies, Zhao's (2024) self-developed questionnaire for senior middle schools, and Lv's "Feedback Engagement Questionnaire for College English Composition," which contains three sub-dimensions of feedback awareness, cognitive strategies, and metacognitive strategies (2019). All items in the scales were positively worded and presented in a 5-point Likert format, with response options ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). Higher scores on these scales indicate greater levels of emotional and cognitive engagement among learners. Before the distribution to the participants, a pilot study with 30 undergraduates who were also first-year and with similar writing scores in the College Entrance Examination with the mean scores of control and experimental groups was conducted, the reliability and validity of the questionnaires were tested as 0.812, 0.853.and 0.802 per dimension (The Cronbach's alpha value) and 0.801, 0.871, and 0.782 per dimension (The KMO value) separately, both >0.7, indicating they are reliable and valid. The questionnaires were distributed anonymously to students in both the experimental group and the control group.

b. Focus group protocol

To gather in-depth insights, a semi-structured focus group discussion was conducted after the intervention with six students from the experimental class, selected based on their writing performance in the College Entrance Examination (evenly distributed across beginning, intermediate, and advanced levels), which reflexes the real-world diversity in Chinese context, which involves mixed-proficiency level learners in the same English class in most cases. The discussion aimed to explore students' perceptions of the mixed feedback patterns and their impact on engagement. A focus group protocol was developed, comprising three questions aligned with engagement dimensions, addressing students' perceptions, attitudes, reactions, and judgments.

Procedures

After obtaining the necessary consent from the participants, the study was conducted step by step in a control group and an experimental group with the same course instructor to ensure standardization across groups, as shown in Figure One.

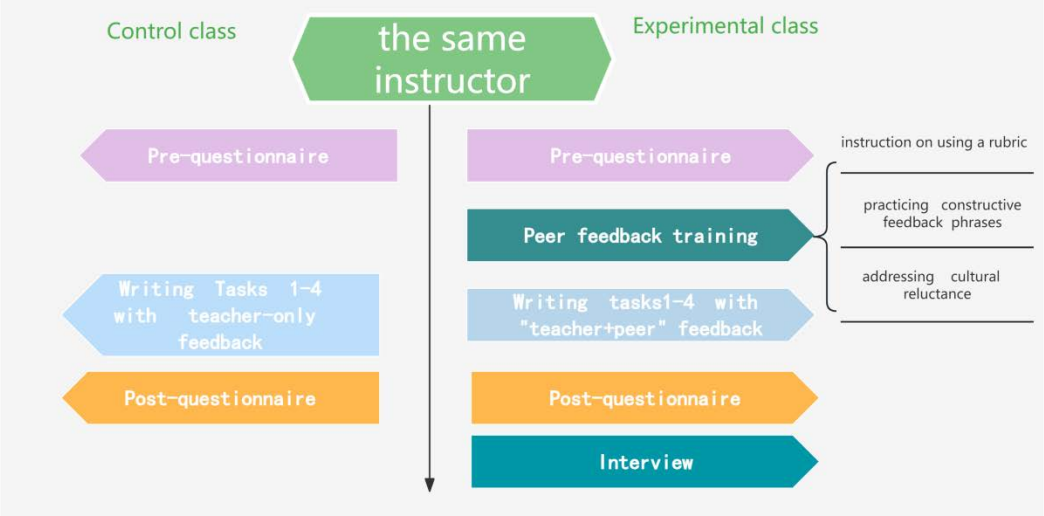


Figure 1. Research procedures

The experiment spanned 12 weeks, involving four writing tasks across groups. Initially, the researcher conducted a pre-survey with a questionnaire for both the control and experimental groups to establish baseline writing engagement levels and to find whether the classes have similar writing engagement levels for a decision about what data analysis approaches would be applied.

Next, before the intervention, a one-hour training session is conducted for the experimental class to equip students with peer feedback skills. This includes instruction on using a rubric (e.g., assessing clarity, organization, and language use), practicing constructive feedback phrasing (e.g., "I suggest adding more examples here"), and addressing cultural reluctance (e.g., role-playing to build confidence in critiquing peers respectfully). Training emphasizes actionable, specific feedback to complement teacher input.

The grouping strategy is employed. Students in the experimental class were grouped into small teams of 3-4 based on mixed proficiency levels (e.g., pairing stronger and weaker writers) to encourage balanced peer input, thus avoiding the aforementioned discrepancy between peers, which may result in ineffective peer feedback. The instructor assigns groups to ensure diversity through the intervention, thereby broadening exposure to different perspectives. This specific grouping strategy caters to the context, as Webb and Farivar (1994) argue that less-skilled students benefit from explanations by more-skilled peers, while higher-ability students deepen their understanding by teaching.

In the implementation phase, spanning over a semester (12 weeks), students in the experimental class draft and revise multiple writing assignments, receiving both types of feedback. Specifically, to optimize learner writing engagement, teacher and peer feedback were balanced by assigning the teacher the role of guiding the whole process, providing authoritative feedback in some critical aspects, such as controversial problems in peer feedback, complex grammar, and structure. Meanwhile, Peers were required to conduct the initial proofreading and feedback by the rubric through pre-intervention training, emphasizing words, general grammar, clarity, and idea development. The intervention was administered in a balanced sequence as shown in Figure 2.



Figure 2. Procedure of teacher-peer mixed feedback for experimental class

In contrast, the control group received traditional "teacher-only" feedback, where the same instructor provided written comments on drafts focusing on grammar, structure, and content. They submitted drafts directly to the teacher for feedback and revision without peer involvement. Their writing maintained consistency in task types and frequency (4 writing tasks) with the experimental group.

In the post-intervention phase, questionnaires to analyze changes in engagement levels across both classes were administered following the twelve-week intervention. The same questionnaires from the pre-intervention phase were administered to the control and experimental classes, allowing for direct comparison of pre-and-post scores. Firstly, paired t-tests within the experimental group explore how engagement shifts over time, addressing RQ1 for the experimental group. Then, the investigation utilizes statistical methods, specifically independent t-tests, to compare post-intervention engagement means between groups, assessing whether the "teacher+peer" feedback yielded higher engagement over "teacher-only" feedback (RQ2).

A qualitative research approach is employed through post-intervention focus group discussion conducted with a subset of the experimental class (6 students). This semi-structured session, lasting approximately sixty minutes, probed students' experiences with the teacher-peer mixed feedback pattern, using questions like "How did peer feedback affect your motivation to write?" (emotional), "Did you revise more after peer or teacher feedback?" (behavioral), and "How did combining feedback types influence your idea of writing?" (cognitive).

Data Analysis

In this section, the results and analysis of data collected to investigate the impact of the "teacher+peer" mixed feedback pattern on engagement in English writing are presented, addressing both its impact across behavioral, emotional, and cognitive dimensions and comparative effectiveness against "teacher-only" feedback.

Impact of "Teacher+peer" Mixed Feedback on Learning Engagement

To address the first research question—how the "teacher+peer" mixed feedback pattern impacts engagement—within-group changes were analyzed. The experimental class exhibited a significant increase, indicating that the mixed feedback approach positively enhances engagement, likely due to the synergistic interplay of peer collaboration and teacher expertise. These changes are presented in Table Two.

Table 2. *The descriptive analysis of the data of the experimental group*

Experimental Group	Phases	Mean	Sd.	Md.	t.	p.
Emotional Engagement	Pre-intervention	3.56	0.431	0.42	5.46	<0.0001
	Post-intervention	3.98	0.428			
Cognitive Engagement	Pre-intervention	3.41	0.423	0.43	5.89	<0.0001
	Post-intervention	3.84	0.376			

Behavioral Engagement	Pre-intervention	3.46	0.301	0.57	9.36	<0.0001
	Post-intervention	4.03	0.366			

According to the table, emotional engagement increased from 3.56 to 3.98 and remains statistically significant ($t = 5.46$, $p < 0.0001$), indicating a meaningful improvement post-intervention. The rise from 3.41 to 3.84 in Cognitive Engagement is also significant ($t = 5.89$, $p < 0.0001$), suggesting a strong intervention effect. As for Behavioral Engagement, the increase from 3.46 to 4.03 shows the most considerable change and remains highly significant ($t = 9.36$, $p < 0.0001$). Therefore, this paired t-test indicates a robust improvement post-intervention.

Comparative Impact between "Teacher-peer" Mixed and "Teacher-only" Feedback

Independent t-tests were conducted to confirm the groups' similarity in engagement before the intervention, which is a critical step to ensure that subsequent differences are not due to baseline differences. It could be attributed to the feedback conditions rather than pre-existing disparities (Cohen et al., 2018). Drawing on Fredricks et al.'s (2004) multidimensional engagement framework, the analysis focuses on pre-intervention scores derived from a Likert-scale questionnaire, setting the stage for post-intervention comparisons and qualitative insights from focus groups. This initial examination of baseline data provides a foundation for understanding how engagement evolves in response to the differing feedback approaches implemented over the 12-week, as shown in Table three.

Table 3. *Independent sample t-test for writing engagement of two classes in the pre-intervention questionnaire*

Class	Domains	N.	Mean	Sd.	Md.	T	p.
control class	Emotional Engagement	42	3.67	0.417	0.11	1.2	0.23
experimental class		43	3.56	0.431			
control class	Cognitive Engagement	42	3.57	0.420	0.16	1.76	0.08
experimental class		43	3.41	0.423			
control class	Behavioral	42	3.39	0.288	-0.07	-1.09	0.28
experimental class	Engagement	43	3.46	0.301			

Both the control and experimental classes share several key features in their pre-intervention engagement levels across the behavioral, emotional, and cognitive dimensions, as assessed before implementing the "teacher+peer" mixed feedback intervention. There are several similarities between the control class ($n = 42$) and the experimental class ($n = 43$) in their baseline engagement, as evidenced by the independent t-test results. For emotional engagement, the control class ($M = 3.67$, $SD = 0.417$) and the experimental class ($M = 3.56$, $SD = 0.431$) showed a mean difference of 0.11, with a t-value of 1.20 ($p = 0.23$), indicating no significant disparity. Similarly, cognitive engagement levels in the control class ($M = 3.57$, $SD = 0.420$) and experimental class ($M = 3.41$, $SD = 0.423$) yielded a mean difference of 0.16 ($t = 1.76$, $p = 0.08$), a result that, while approaching significance, does not cross the conventional threshold of $p < 0.05$. Behavioral engagement further underscores this equivalence, with the control class ($M = 3.39$, $SD = 0.288$) and experimental class ($M = 3.46$, $SD = 0.301$) differing by only -0.07 ($t = -1.09$, $p = 0.28$), a statistically non-significant outcome. These results are similar to those reported by Fredricks et al. (2004), who found that engagement scores across student groups often exhibit consistency at baseline in educational research settings.

Another independent sample t-test was conducted after the intervention; the results are shown in Table 4.

Table 4. *Independent Samples T-test for Post-test of Emotional Engagement in Two Feedback Approaches*

	"teacher-only" Feedback		"Teacher + Peer" mixed Feedback				
	Mean	Sd.	Mean	Sd.	Md.	t	p
Emotional Engagement	3.76	0.411	3.98	0.428	0.22	2.42	0.018
Cognitive Engagement	3.59	0.476	3.84	0.376	0.25	2.69	0.009
Behavioral Engagement	3.35	0.502	4.03	0.366	0.68	8.72	<0.001

The results present the findings from post-intervention engagement data collected from the control group and the experimental group. Independent t-tests assessed post-intervention differences to find out whether "teacher+peer "outperforms "teacher-only" feedback.

Post-intervention results revealed significant differences favoring the experimental group across all engagement dimensions. For emotional engagement, the experimental class ($M = 3.98$, $SD = 0.428$) outperformed the control class ($M = 3.76$, $SD = 0.411$), with a mean difference of 0.22 ($t(83) = 2.42$, $p = 0.018$). Within-group changes showed a modest increase for the control group (+0.09) compared to a substantial rise for the experimental group (+0.42), suggesting that the mixed feedback pattern markedly enhanced emotional engagement, such as motivation or positive affect toward writing. Cognitive engagement followed a similar trend, with the experimental group ($M = 3.84$, $SD = 0.376$) exceeding the control group ($M = 3.59$, $SD = 0.476$) by 0.25 ($t = 2.69$, $p = 0.009$). The experimental group's cognitive engagement increased by 0.43, while the control group rose slightly (+0.02), indicating deeper cognitive processing (e.g., reflection on feedback) under the mixed feedback condition.

The most remarkable effect emerged in behavioral engagement, where the experimental group ($M = 4.03$, $SD = 0.366$) significantly surpassed the control group ($M = 3.35$, $SD = 0.502$), with a mean difference of 0.68 ($t = 8.72$, $p < 0.001$). The experimental group's behavioral engagement surged by 0.57, reflecting increased revision or effort, whereas the control group slightly declined (-0.04), potentially due to fatigue or lack of peer stimulation.

These results indicate that the "teacher+peer" mixed feedback approach significantly enhances writing engagement across all three dimensions compared to "teacher-only" feedback, addressing the second research question affirmatively. The experimental group's greater improvements highlight the intervention's efficacy in fostering motivation, reflection, and effort, aligning with prior research on peer-assisted learning (Webb & Farivar, 1994).

Besides the quantitative approach, the current study used qualitative analysis to gain in-depth insights into the influences of the intervention, thus complementing quantitative findings. Three questions were conveyed to the focus group discussion: How did peer feedback affect your motivation to write?" (concerning emotional engagement), "Did you revise more after teacher+peer feedback?" (concerning behavioral engagement), and "How did combining feedback types help you think about your writing?" (concerning cognitive engagement). The data are audibly recorded throughout, transcribed, and thematically analyzed to identify themes. The following is the feedback and the corresponding themes drawn from the focus group:

The feedback for Question 1 "How did peer feedback affect your motivation to write?" (emotional engagement) revealed that peer feedback significantly boosted students' motivation to write, fostering a sense of emotional engagement (four out of six interviewees reported). The theme

extracted manually revealed increased motivation through peer concern, encouragement, and self-esteem（three codes）, which aligns with the quantitative finding of a substantial increase in emotional engagement scores for the experimental group, as shown below:

Student 1: "I like this mixed pattern, especially when my classmate said my ideas were interesting and just needed clearer examples. I felt encouraged. It made me want to write better next time because I knew someone liked me and saw my potential in it."

Student 2"I feel teachers and my classmates are my good helpers in writing."

Student 5: "Teacher feedback is helpful, but peer feedback felt more personal and interesting. Like, my friend wrote, It is great, go on!' and drew a "thumb-up" symbol. That pushed me to enjoy writing more."

Student 6: "I was nervous at first, but after getting positive peer comments, I felt less stressed and more excited to improve my essay. If my teachers' remarks align with my peers' comments, I will be impressed and trust my peers more than before. "

The feedback for Question 2: "Did you revise more after 'peer+teacher feedback?" (behavioral engagement) revealed that peer feedback drives greater and more frequent revision, as shown below (four out of six interviewees reported):

Student 2: "I revised a lot more after peer feedback because they pointed out things I did not notice, like confusing parts. I would fix those before showing the teacher."

Student 3: "Teacher feedback was good, but it felt final. With peers, I would rewrite whole sections after our discussion because I wanted it to be clearer for them next time."

Student 4: "Yes. More than before..."

Student 6: "My group gave me ideas—like adding examples—and I kept revising until it made sense. I probably changed my draft three times after peers, then I felt confident to submit the draft to teachers, but only once after the teacher."

The original excerpts indicate that "teacher+peer" feedback prompted more frequent and iterative revisions, likely due to its collaborative and immediate nature (two codes), aligning with the experimental group's significant behavioral engagement increase.

The feedback for Question 3 "How did teacher+peer feedback pattern affect your idea about writing?" (cognitive engagement) demonstrated that peer feedback enhanced reflection (three out of six interviewees reported) utilizing "thinking harder, seeing my writing differently and two-step procedure" (3 codes) as shown below:

Student 2: "Peers gave me simple ideas, like 'this part is unclear,' and then the teacher explained why—like sentence structure. Together, it made me think harder about how to fix it."

Student 4: "Combining them helped me see my writing differently. Peers showed me what a reader might like to read, and the teacher gave me rules to follow, so I planned my revisions better."

Student 5: "I would think about peer comments first, then check the teacher's feedback to make sure I got it right. It was like two steps of the way."

In the third question, students reported that combining teacher and peer feedback deepened their cognitive engagement by encouraging critical reflection and strategic improvement in their writing. This theme also supports the quantitative data, where the experimental group's cognitive engagement improved markedly.

Discussions

In addressing RQ 1, the present study confirms that the "teacher+ peer" mixed feedback pattern is associated with significant enhancements in students' emotional, cognitive, and

behavioral engagement in English writing, which is consistent with the findings that students prefer it over single-source options (Demirel & Enginarlar, 2016; Loan, 2017; Yu, 2020). A similar finding was also reported by Nelson and Schunn (2009), whose study on peer feedback in writing highlighted its role in prompting iterative revisions, particularly when peers provide specific, actionable suggestions.

However, the results contradict some scholars' claim that learners prefer teachers' written feedback over other forms of feedback and that teachers' feedback is the most helpful in improving writing enhancement (Ferris & Roberts, 2001; Mustafa et al., 2022).

The above-mentioned contradiction entirely supports Phothongsunan's claim that the practicality of English writing feedback largely depends on the verified contexts (2023). Varied and specific English writing teaching environments, such as school policies and backgrounds, teacher professional capabilities, and students' English proficiency levels, would lead to significant diversity across schools and even classrooms.

In answering RQ 2, a comparison indicated that the mixed feedback yielded significantly higher engagement than teacher-only feedback. It aligns with Xing and Wang (2023), as they rank it as the most effective among feedback sources. Interestingly, with the integration of peer feedback to the writing outcome, the mixed method completely outperformed the "teacher-only" feedback, indicating the peer feedback somehow enhances the writing engagement, other than weaken it by students' distrust in peer due to perceived disrespect or incompetence (Zhou et al., 2020), feeling stressed by peer assessment (Stančić, 2021), or questioning peers' language skills, threatening the face-saving norms in Chinese culture (Xue, 2023), and scarce specialized knowledge and literacy about the topic (Valero Haro et al.,2023). There is no negative feedback on peer feedback in English writing in the current study. Surprisingly, Zhao (2024) conducted a related study in Chinese secondary education, finding that the emotional engagement in the two feedback modes is essentially identical, which may be caused by different study phases and motivations of participants.

Pedagogical Implications

The current study provides several pedagogical implications for educators aiming to enhance writing instruction, addressing logistical and cultural factors: using rotating mixed-ability groups instead of pairs, overcoming cultural hesitancy among reserved Chinese students, accommodating variations in English proficiency and engagement, and managing time constraints within a semester.

- Effective use of grouping techniques, specifically mixed-ability groups

The study's success in boosting engagement and qualitative reports of increased revision suggest that pairing stronger and weaker writers fosters diverse peer feedback and sustained effort. Educators should adopt this strategy to encourage balanced input and expose students to varied perspectives, enhancing collaboration and engagement over time. The consideration of rotating groups periodically may prevent stagnation and ensure that students benefit from multiple peer interactions.

- Training on cultural reluctance

A noted challenge in Chinese educational contexts, where students may hesitate to critique peers due to face-saving norms, also carries significant pedagogical weight. To overcome the difficulty, teachers should implement pre-intervention training sessions, as done in this study, to address the bias towards criticism and build students' confidence in providing constructive

critique. Since the training requires continuous efforts which may not be practical in some contexts, alternatively, offering anonymous feedback options—such as written comments submitted without names—can further reduce social pressure, ensuring that peer feedback remains productive without compromising cultural norms.

- Careful group composition

As for proficiency variability, where differences in English ability could worsen writing engagement outcomes, it underscores the need for careful group composition, statistical control, and the significance of teachers' involvement in feedback. Although this study assumed similar baseline proficiency, the significant cognitive engagement suggests a further study on groups with different proficiency or writing engagement levels, whose comparison may still be conducted by application of a covariate in analysis, as recommended in quasi-experimental designs (Cohen et al., 2018). This approach ensures that feedback interactions remain equitable.

- A manageable number of cycles

Time constraints within a semester pose a practical way for its employment. With only 12 weeks, the intervention achieved significant results across four writing cycles, suggesting that a manageable number of cycles is sufficient to yield engagement benefits.

Collectively, these implications suggest that integrating a "teacher+peer" mixed feedback approach into English writing curricula can substantially enhance student engagement, provided educators address grouping dynamics, cultural sensitivities, proficiency differences, and temporal limits.

Conclusion

This study has identified that the "teacher+peer" mixed feedback pattern significantly enhances engagement in English writing among Chinese tertiary students across emotional, cognitive, and behavioral dimensions, as evidenced by substantial post-intervention mean increases in the experimental group compared to minimal changes in the control group. The research has also shown that integrating peer feedback with teacher input outperforms "teacher-only" feedback, with statistically significant between-group differences and qualitative themes highlighting increased motivation, revision, and reflection, offering a robust framework for enhancing writing engagement in EFL settings. It enriched the existing literature on writing feedback and engagement theory in the ELT domain by incorporating empirical data from the Chinese context, and also advanced peer involvement through strategies to overcome the cultural reluctance of "hurting face" in collectivist cultures, such as China.

Admittedly, there are some limitations, such as the sampling technique and sample size, which probably prevent a general finding, and the short duration. Even the significant increase in writing engagement could partially validate the effectiveness of the mixed feedback, a further and longer intervention is deemed important and necessary.

Funding: This research was partially funded by Universiti Kebangsaan Malaysia through the Faculty of Education (grant code: TAP-K017971), whose support the authors gratefully acknowledge.
Acknowledgments: Not applicable.
Conflicts of Interest: The authors declare no conflict of interest.
Authenticity: This manuscript is an original work
Artificial Intelligence Statement: AI and AI-assisted technologies, specifically ChatGPT-

4.0,were used in this study for grammar and language refinement. The authors take full responsibility for the accuracy and integrity of the content.

About the Authors:

Min Deng is currently a doctorate candidate in TESL at the Faculty of Education, Universiti Kebangsaan Malaysia (UKM). She works at China West Normal University, China. She specializing in English textbook evaluation, intercultural communication, AI-empowered writing teaching. ORCID: https://orcid.org/0009-0005-0996-9277

Dr. Nur Ehsan Mohd Said is a senior lecturer at the Faculty of Education, Universiti Kebangsaan Malaysia (UKM). His main research interest lies in the field of ESL classroom intervention, curriculum design, language assessment, and educational technology integration. His research explores the intersection of applied linguistics, pedagogy, and digital learning environments with influential publications in internationally recognized journals.
ORCID: https://orcid.org/0000-0002-2891-327X

Dr. Nur Ainil Sulaiman is a senior lecturer at the Centre of Innovation in Teaching and Learning, Faculty of Education, Universiti Kebangsaan Malaysia (UKM). Her research focuses on advancing online English instruction and second language acquisition, with influential publications in internationally recognized journals.ORCID: https://orcid.org/0000-0001-6212-7494

Yunyun Zhao is an EFL teacher at Jiujiang New Town Middle School, Chengdu, China. She is also an ELT researcher with expertise in peer feedback and English writing. She has conducted studies on language curriculum design for Secondary Education learners.
ORCID:https://orcid.org/0009-0002-8546-1625

References

Bai, H. Q., & Li, T. (2020). Tendencies in adopting comments in online peer review: A perspective based on emotional experience. *Open Education Research,* 26(04), 95–101.

Cai, K. G. (2011). A Comparative Study of Online Peer Feedback and Teacher Feedback on Chinese College Students' English Writing, *Foreign Language Community,* (02), 65-72.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2018). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Routledge.

Demirel, E. T., & Enginarlar, H. (2016). Effects of Combined Peer-Teacher Feedback on Second Language Writing Development. *Hacettepe University Journal of Education,* 31(4), 657-675. https://doi.org/10.16986/huje.2016015701

Fan, Y., & Xu, J. (2020). Exploring Student Engagement with Peer Feedback on L2 Writing. *Journal of Second Language Writing,* 50, 100775. https://doi.org/10.1016/j.jslw.2020.100775

Ferris, D., & Roberts, B. (2001). Error Feedback in L2 Writing Classes: How Explicit does it Need to be? *Journal of Second Language Writing,* 10(3), 161–184. https://doi.org/10.1016/S1060-3743(01)00039-X

Flower, L. S. (1979). Writer-Based Prose: A Cognitive Basis for Problems in Writing. *College English*, 41, 19–37.

Fredricks, J., Blumenfeld, P., & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research,* 74, 109–159.

https://doi.org/10.3102/00346543074001059
Gao, E., Wang, Y., & Schunn, C. D. (2019). A study of English writing peer feedback rubric adoption and its influencing factors. *Foreign Language e-Learning*, (02), 17–24.
Geng F. &Yu S. (2023). A study of learners' input to peer feedback in English argumentative essay writing. *Frontiers of Foreign Language Education Research*, 6(01), 67-74+95.
Goldstein, L. M. (2016). *Handbook of Second and Foreign Language Writing.* Berlin: De Gruyter Mouton.
Han, Y.& Hyland, F. (2015). Exploring Learner Engagement with Written Corrective Feedback in a Chinese Tertiary EFL Classroom. *Journal of Second Language Writing*, 30, 31-44. https://doi.org/10.1016/j.jslw.2015.08.002
Han, Y., & Xu, Y. (2021). Student Feedback Literacy and Engagement with Feedback: A Case Study of Chinese Undergraduate Students. *Teaching in Higher Education*, 26(2), 181–196.
Hyland, K., & Hyland, F. (2019). *Feedback in Second Language Writing.* Cambridge: Cambridge University Press.
Kim, H. R., & Bowles, M. A. (2019). How Deeply Do Second Language Learners Process Written Corrective Feedback? Insights Gained From Think-Alouds. *TESOL Quarterly,* 53(4), 913-938. https://doi.org/10.1002/tesq.522
Koltovskaia, S. (2020). Student Engagement with Automated Written Corrective Feedback (AWCF) Provided by Grammarly: A Multiple Case Study. *Assessing Writing*, 44, 100450. https://doi.org/10.1016/j.asw.2020.100450
Liu, X.H., & Ji, X.L. (2018). A study on the feasibility and effectiveness of peer scoring in college English writing. *Foreign Language Community*, (05), 63-70.
Loan, N. T. (2017). A Case Study of Combined Peer-Teacher Feedback on Paragraph Writing at a University in Thailand. *Indonesian Journal of Applied Linguistics,* 7, 253-262. https://doi: dx.doi.org/10.17509/ijal.v7i2.8128
Lv, Y. Z. (2019). *A study on English learners' engagement with written corrective feedback and its influencing factors* [Master's thesis]. Beijing Foreign Studies University, Beijing, China.
Mahfoodh, O. H. A. (2017). "I Feel Disappointed": EFL University Students' Emotional Responses towards Teacher Written Feedback. *Assessing Writing*, 31, 53-72. https://doi.org/10.1016/j.asw.2016.07.001
MOE of the People's Republic of China. (2020). *College English Curriculum Guides. Beijing.* Beijing: Higher Education Press.
Mustafa, A., Arbab, A. N., & El Sayed, A. A. (2022). Challenges in Academic Writing in English as a Second/Foreign Language from the Perspective of Undergraduate Students in Higher Education Institutions in Oman. *Arab World English Journal*, 13(3), 41-53. https://dx.doi.org/10.24093/awej/vol13no3.3
Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance.*Instructional Science*, 37(4), 375-401. https://doi.org/10.1007/s11251-008-9053-x
Noroozi, O., Banihashem, S. K., Biemans, H. J., Smits, M., Vervoort, M. T., & Verbaan, C. L. (2023).
Design, implementation, and evaluation of an online supported peer feedback module to enhance
students' argumentative essay quality. *Education and Information Technologies, 1–28.*

https://doi.org/10.1007/s10639-023-11683-y

O'Malley, J. M., & Chamot, A. U. (1990). *Learning Strategies in Second Language Acquisition.* Cambridge: Cambridge University Press.

Pace, C. (1984). *Measuring the Quality of College Student Experiences. An Account of the Development and Use of the College Student Experiences Questionnaire.* Los Angeles: Higher Education Research Institute.

Phothongsunan, S. (2023). EFL Student-Directed Feedback for Improving Academic Writing Skills in Thailand. *Arab World English Journal (AWEJ) Volume,* 14,223-231. https://dx.doi.org/10.24093/awej/vol14no1.14

Ranalli, J. (2021). L2 Student Engagement with Automated Feedback on Writing: Potential for Learning and Issues of Trust. *Journal of Second Language Writing,* 52, 100816. https://doi.org/10.1016/j.jslw.2021.100816

Reeve, J. (2012). A self-determination theory perspective on student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on Student Engagement* (pp. 149-172). Springer.

Ren, Q.M. (2023). The predictive effect of learning engagement in the English classroom on satisfaction with learning gains. *Modern Foreign Languages,* 46(04), 540-551.

Shi, M. (2019). The effects of class size and instructional technology on student learning performance. *The International Journal of Management Education, 17(1),* 130–138. https://doi.org/10.1016/j.ijme.2019.01.004

Stančić, M. (2021). Peer Assessment as a Learning and Self-assessment Tool: A Look Inside the Black Box. *Assessment & Evaluation in Higher Education,* 46(6), 852-864. https://doi.org/10.1080/02602938.2020.1828267

Tian, L., & Zhou, Y. (2020). Learner Engagement with Automated Feedback, Peer Feedback, and Teacher Feedback in an Online EFL Writing Context. *System,* 91, 102247. https://doi.org/10.1016/j.system.2020.102247

Tsui, A. B. M., & Ng, M. (2000). Do Secondary L2 Writers Benefit from Peer Comments? *Journal of Second Language Writing,* 9(2), 147-170. https://doi.org/10.1016/S1060-3743(00)00022-9

Uttley, J. (2019). *Power analysis, sample size, and assessment of statistical assumptions—Improving the evidential value of lighting research*. Leukos.

Valero Haro, A., Noroozi, O., Biemans, H. J., Mulder, M., & Banihashem, S. K. (2024). How does the type of online peer feedback influence feedback quality, argumentative essay writing quality, and domain-specific learning? *Interactive Learning Environments,* 32(9), 5459-5478. https://doi.org/10.1080/10494820.2023.2215822

Webb, N. M., & Farivar, S. (1994). Promoting helping behavior in cooperative small groups in middle school mathematics. *American Educational Research Journal,*31(2), 369-395. https://doi.org/10.3102/00028312031002369

Xing, J.X.& Wang, H.H. (2023). The effects of different sources of written feedback on college students' English writing performance. *Modern Foreign Languages,* 46(02), 213-223.

Xu, J.F., & Fan, Y.M. (2019). A study of foreign language learners' engagement from a social cognitive perspective. *Foreign Language Teaching,* 40(05), 39-43+56.

Xue, S., Yang, Y., Du, J., & Liu, F. (2023). Multi-layered e-feedback anxiety: An action research study among Chinese learners using peer feedback activities in an academic writing course. *Frontiers in Psychology, 14*, 1062517.https://doi.org/10.3389/fpsyg.2023.1062517

Yu, H. (2020). An empirical study on the effects of teacher feedback and peer review on English major students' writing. *Overseas English,* (20), 31–33.

Yu, S., & Lee, I. (2016). Peer Feedback in Second Language Writing (2005–2014). *Language Teaching,* 49(4), 461-493. https://doi.org/10.1017/S0261444816000161

Zhang, Z., & Hyland, K. (2018). Student Engagement with Teacher and Automated Feedback on L2 writing. *Assessing Writing,* 36, 90-102. https://doi.org/10.1016/j.asw.2018.02.004

Zhao, Y.Y. (2024). *A Study on the Effects of "teacher+peer" Feedback on Senior High School Students' English Writing Proficiency,* (Unpublished Master's thesis). China West Normal University.China.

Zheng, Y., Yu, S., Wang, B., & Zhang, Y. (2020). Exploring student engagement with supervisor feedback on Master's thesis: insights from a case study. *Innovations in Education and Teaching International,* 57(2), 186-197. https://doi.org/10.1080/14703297.2019.1617181

Zhou, J., Zheng, Y., & Tai, J. H. M. (2020). Grudges and Gratitude: the Social-affective Impacts of Peer Assessment. *Assessment & Evaluation in Higher Education,* 45(3), 345-358. https://doi.org/10.1080/02602938.2019.1643449

Appendix
Rubrics for Peer Feedback

Feedback types	Categories	Description
Form	**Words** **Singular/plural** **Word form** **Spelling** **Punctuation** **Sentence structure** **Grammatical errors** **Word count**	**Pron./prep./conj./articles** **Verb tense and voice** **Missing or adding** **Run-on/subject-verb agreement…** **Insufficient word counts**
Content	**Main-idea** **Content score points** **Genera** **Structure** **Redundancy**	**Clearly & accurately stated** **Complete** **Letter/diary/narration…** **Well-organized**

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.175- 187
DOI: https://dx.doi.org/10.24093/awej/vol16no3.10

A Comparative Study of Human and AI-Assisted Assessment using ChatGPT: The Case of Moroccan EFL Learners

Hajar Ennadir
Department of English, Faculty of Letters and Humanities Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, Fes, Morocco
Corresponding Author: hajar.ennadir@usmba.ac.ma

Salma Habiballah
Department of English, Faculty of Letters and Humanities Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, Fes, Morocco
Email : salma.habiballah@usmba.ac.ma

Received:06/02/2025 Accepted: 09/12/2025 Published:09/20/2025

Abstract
Artificial intelligence (AI) is capable of performing various tasks and has been applied across multiple fields, including higher education. AI is increasingly used in teaching, learning, and classroom assessment. The primary aim of this study is to compare human assessments and AI-assisted assessments in Moroccan higher education, focusing on summative e-assessments. The primary question of the study is: How does AI-assisted assessment differ from human assessment in evaluating assignments in Moroccan higher education? The study yields data specific to the Moroccan context, filling the gap in this concern; it also provides educators and policymakers with a clear view of the similarities and differences between human and AI-assisted assessment. Writing assignments from EFL classes in two departments at the Faculty of Letters and Humanities, Dhar El Mahraz, were graded by their respective instructors and ChatGPT-4. The results of this study showed that AI grading produced higher scores than human grading, with differences that are both statistically significant in the department of French language studies, while in the department of psychology, the differences between human grading and AI grading are neither statistically significant nor practically large, suggesting that AI grading is comparable to human grading. The study's findings imply that the grading system for AI might be different and is, therefore, incapable of replacing the human aspect of the grading process in Moroccan higher education, opening the possibility to enhance AI models of assessing EFL and align them with human standards of assessing EFL.

Keywords: Artificial intelligence, assessment, ChatGPT, e-assessment, Moroccan higher education, writing assignment

How to cite: Ennadir, H., & Habiballah, S. (2025). A Comparative Study of Human and AI-Assisted Assessment using ChatGPT: The Case of Moroccan EFL Learners *Arab World English Journal, 16*(3): 175-187. https://dx.doi.org/10.24093/awej/vol16no3.10

Introduction

Artificial Intelligence (AI), since its emergence, has been developing to be able to perform multiple actions. It has been used in a number of activities in different fields such as speech recognition, autonomous learning and scheduling, game playing, spam fighting, logistics planning, machine translation, etc. (Stuart & Peter, 2010). With its massive advancement impacting several fields and domains, AI has also come to affect the field of education in general and the field of higher education in particular. Research predicted that in the upcoming years, one of the most useful technologies will be artificial intelligence (AI) (Chai et al., 2021). No field was exempt from the AI invasion; teaching and learning are a great example of that, notably in the aspect of classroom assessment (Jurane-Bremane, 2023).

In Moroccan educational contexts, research shows that concerns were raised about the use of AI in the classroom, particularly in early childhood education. Yet, using AI as technical support seems to have optimistic implications (Douali et al., 2022). Moreover, Moroccan higher education has institutionalized the adoption of AI, with Mohammed VI Polytechnic University (UM6P) being the first university in Africa to formally implement ChatGPT Edu as an assistant in teaching, research, and administration (Hespress EN, 2024). Statistically, Moroccan educators are eager to use AI in the classroom, with 64% of them thinking it would be simple to learn the required skills. Nonetheless, 92% of educators acknowledge that users' intentions may have an impact on the adoption of AI (Moukhliss et al., 2024).

This study focuses on summative e-assessment in the context of Moroccan higher education. It aims to compare human and AI grading in EFL classes to unravel the accuracy of each one. Writing assignments of two departments at the faculty of letters and humanities Dhar El Mahraz were graded by their respective instructors and ChatGPT-4 abiding by five criteria: grammar, spelling, punctuation, clarity, and structure.

The paper raises the research question “to what extent is AI summative e-assessment accurate compared to the human one in the context of Moroccan EFL education?”. The significance of this study addresses the issue of the impact of AI tools on summative e-assessment in Moroccan universities. By focusing on comparing AI and human grades, this paper discusses practical issues of the use of AI in education and offers empirical insights into the Moroccan breakthrough regarding educational technology. Most studies tackle the issue of AI and education by simply highlighting its advantages and pitfalls when used as a learning support tool rather than an instrument for assessing academic assignments. This creates a research gap that this study addresses to enhance the understanding of AI-assisted assessment accuracy.

Literature Review

Artificial Intelligence

Artificial intelligence can be defined as a field of study and the resulting innovations and developments that have culminated in computers, machines, and other artifacts having human-like intelligence characterized by cognitive abilities, learning, adaptability, and decision-making capabilities (Chen, 2020). The term Artificial Intelligence was first coined by John McCarthy in 1956 (Sumakul et al., 2022) and can be defined as the existence of thinking and intelligent reasoning, which can go into intelligent machines (Turing, 2009 as cited in Jen & Salam, 2024). Thanks to its impact on the field of higher education, AI has eventually led to a new field of research named Artificial Intelligence in Higher Education (AIED). Zawacki-Richter et al. (2019) classify research in AIED into four broad areas, which are Profiling and Prediction, Intelligent

Tutoring Systems (ITS), Adaptive Systems, and Personalisation and Assessment and Evaluation. The latter is the focus of this paper.

Profiling and prediction

This area of research focuses on how AI is used to make informed decisions about students' academic journey (Zawacki-Richter et al., 2019). The use of AI prediction models to determine student profiles has increased dramatically. These models focus on the result of student performance to draw a line of action that allows the student to strengthen their deficiencies and gradually achieve their performance objectives (Y. Chen, 2018 as cited in Rincon-Flores, 2022). AI allows for the analysis of vast amounts of data, including academic records, attendance patterns, in addition to social and emotional factors, which provide a comprehensive view of the academic profile of students (Katarya, R. 2019 as cited in Maulana, A., Idroes, 2023).

Intelligent Tutoring Systems (ITS)

This domain is concerned with the use of AI to provide tailored instructional interventions. This domain is concerned with the use of AI to provide tailored instructional interventions (Zawacki-Richter et al., 2019). An ITS provides computer-based tutorials on topics in well-defined structured subjects. ITS delivers a sequence of information, activities, and quizzes adapted to each individual student. While the student engages with a particular activity, the system captures thousands of data points such as clicks, typing, correct answers, and misconceptions. These data are analysed in order to determine the next information, activity, and quiz to be delivered (Holmes and Tuomi, 2022). An ITS can monitor students' progress, provide feedback, hints, and scaffolding, and select appropriate problems or tasks for students to practice. Besides, they can model the psychological state of students, such as motivation, emotion, and cognition, as well as their prior knowledge, skills, and preferences (Lin et al., 2023).

Adaptive Systems and Personalisation

Adaptive systems and personalisation focus on the use of AI in personalising the educational experience of learners (Zawacki-Richter et al., 2019). AI provides specialized support and raises knowledge-gap awareness, enabling instructors to teach effectively and efficiently through personalized and adaptive instruction (Guan et al., 2020 as cited in Chen, 2021). AI is a strong toolkit for expanding the adaptivity provided to students. It can develop a technology that meets students where they are, builds on their strengths, and grows their knowledge (U.S. Department of Education, 2023).

Assessment and Evaluation

The final area is AI-based assessment and evaluation, which is the focus of this study. AI in assessment and evaluation focuses on the automation of evaluation using AI (Zawacki-Richter et al., 2019). According to Mirchi et al., 2020 as cited in González-Calatayud, 2021, the researchers used AI for simulation-based training in medicine, and they created a Virtual Operative Assistant in order to provide automatic feedback to students based on metrics of performance. From a formative educational perspective, virtual reality and AI are integrated to classify students in relation to proficiency performance benchmarks, and the system gives feedback to help them improve (Janpla, 2020 as cited in González-Calatayud, 2021). Some autograders are claimed to diagnose the type of error and suggest corrections, while others, depending on the domain, are

claimed to score student answers correctly with an accuracy of about 90 percent (Hsu et al., 2021 as cited in Holmes and Tuomi, 2022).

AI and Assessment

There is a strong association between AI and human grades, according to research on AI-based essay grading systems in higher education. However, its successful incorporation depends on ethical considerations, appropriate instruction, and the improvement of AI-generated material. Enhancing AI writing tools and investigating new avenues are necessary to improve writing in higher education (Aljuaid, 2024).

Assessment is one of the most affected areas by AI in higher education. AI assessment tools provide personalized feedback through the analysis of students' responses to assignments, helping in identifying their areas of strength and weakness. Eventually, providing educators with important insights into instructional effectiveness, student performance, and learning outcomes (Nazaretsky et al., 2022).

Similarly, Huang et al. (2023) state that AI tools can provide insights about the effectiveness of the instruction and identify areas where teaching strategies may need to be adjusted. Besides, the study adds that AI-powered tools can help in automating different aspects of the assessment process, saving time and reducing the burden on teachers. AI-powered grading tools are an example; these tools can analyse essays of students and provide feedback on grammar, structure, and content; hence, reducing teachers' time grading assignments.

Assessment and E-assessment

According to Gierl et al. (2014), the process of assessment is deemed challenging by teachers, especially when it comes to hand-marking essays. Manual grading requires careful consideration, critical analysis, and tailored criticism, which adds to the lengthy and highly resource-intensive grading process as a whole. Furthermore, the difficulties of manual assessment are made worse by the vast number of sheets, particularly in light of the expanding student population.

Particularly in higher education, e-assessment is becoming a daily need in the educational process. Applications, learning management systems, or specialist software can provide it. Appiah and Van Tonder (2018, p. 1454) assess that "e-assessment involves the use of any technological device to create, deliver, store and/or report students' assessment marks and feedback".

Formative vs Summative Assessment

In most research studies, the focus lies on formative assessment. The concept of employing AI for assessment is founded on the assistance that this technology may offer teachers with a large class size (Bai et al., 2022) or on the notion that learning is enhanced by the rapid feedback that students get (Podsiad and Havard, 2020). When the learning process is finished, summative assessment takes place. Tests, projects, presentations, and performance assignments are a few types of summative assessments. The goal behind summative assessment is to demonstrate how well a student has grasped the unit's concepts, knowledge, and abilities. The latter is the focus of this paper.

AI in EFL Writing Assessment

Research emphasizes the importance of automated scoring, particularly in EFL contexts. Globally, Uyar and Büyükahıska (2025) argue that despite ChatGPT's potential as a scoring tool, the discrepancies indicate that it is still not proficient enough for practical use. Furthermore, Yavuz et al. (2024) conclude that although ChatGPT and Google's Bard showed a high degree of consistency and grading proficiency, further work and modification are required to gain a more sophisticated comprehension of non-objective essay requirements. In Morocco, generative AI is not only used for assessment but also for test generation. A study conducted by Far-hat and Ouchouid (2024) revealed that Moroccan EFL instructors' effectiveness when utilizing ChatGPT to create formal exams was positively impacted by providing them with a guide that included prompt designing and test construction guidelines. Among other things, it was suggested that their professional training incorporate AI's application in planning and teaching.

Method

The study compares human and AI-assisted assessment in English writing assignments. The objective of the study is to examine the extent to which human assessment is similar to or different from AI-assisted assessment. In order to meet the objective of the study, the paper used ChatGPT-4. This version of ChatGPT is designed for detailed analysis, nuanced understanding, and structured responses, making it well-suited for language correction and teaching. Besides, the study used the standard ChatGPT interface. A writing assignment was assigned to the participant in which they were asked to describe a person of their choice, like a friend or a family member. The pieces of writing were then corrected by both ChatGPT-4 and two English language teachers to assess the difference between the two methods of assessment.

Participants

The sampling approach adopted was convenience sampling. The sample was composed of 30 participants from the Faculty of Letters and Humanities Dhar El Mahraz at the University of Sidi Mohamed Ben Abdellah in Fez, Morocco. Eighteen participants were from the department of French language studies; eleven were female participants, and seven were male. Twelve participants were from the department of psychology; eleven were female participants, with one male.

Procedures

To conduct the comparison, participants were asked to write a paragraph in which they describe a person of their choice, like a friend or a family member. They had twenty minutes to do the assignment on a Google Form. First, the assignment was graded by two teachers focusing on five criteria, which are grammar, spelling, punctuation, clarity, and structure. Then, the same pieces of writing were graded using ChatGPT-4. ChatGPT-4 was given the following prompt: "Grade the following paragraph out of 10 based on the following criteria: Grammar, spelling, punctuation, clarity, and structure. Each criterion should be graded out of 2." The total grade was 10 out of 10, with 2 grades for each criterion. The grammar criterion focused on correct sentence structure, subject-verb agreement, proper use of articles, and pronoun accuracy. The criterion of spelling checked the correct spelling of each word. For punctuation, it evaluated capitalization at the beginning of sentences and for proper nouns, as well as the use of periods to separate ideas. The clarity criterion examined whether the intended meaning could be understood by the reader

without confusion. The final criterion was structure; it focused on the organization of ideas into logical, separate sentences and paragraphs.

Results

This section lays out the data in tables and charts and provides descriptive analysis accounting for observations, patterns, or crossovers. The section presents an interpretation for both the French and psychology departments.

Findings of the French department

Table 1. *French department's students' human and AI grades*

Human Grading	AI Grading	Difference	Degree of Difference
4	8	Different	-4
6	7.5	Different	-1.5
6	8	Different	-2
10	9.5	Different	0.5
7	9.5	Different	-2.5
4	7.5	Different	-3.5
6	8	Different	-2
9	10	Different	-1
5	8	Different	-3
4	8	Different	-4
9	10	Different	-1
7	9	Different	-2
3	10	Different	-7
7	10	Different	-3
5	8	Different	-3
5	10	Different	-5
9	10	Different	-1
7	10	Different	-3

The table represents the grades provided by the teacher (Column 1) and the grades provided by ChatGPT (Column 2) to the 18 students of the French department. The third column accounts for whether the grades were different or similar among the assessors while the fourth column shows the difference scores.

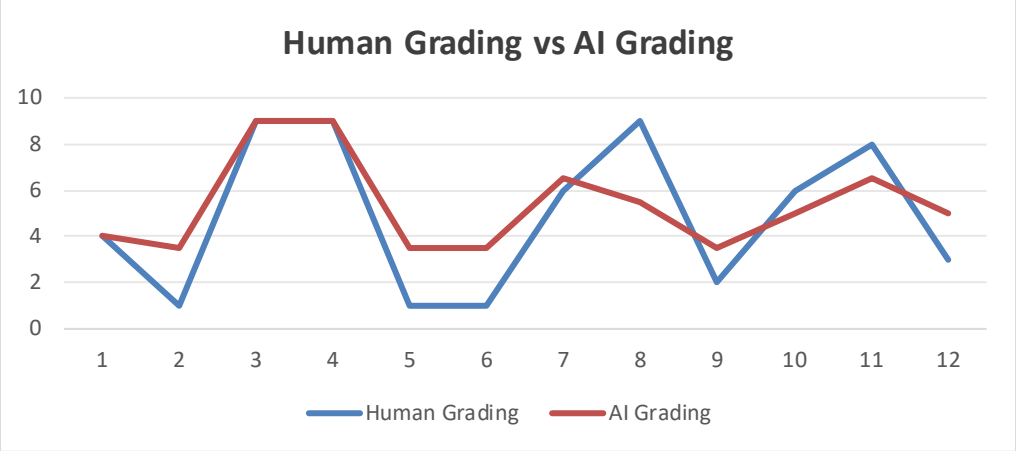


Figure 1. Comparison of human vs. AI grading

The figure illustrates a line chart comparing the human and AI grades among the 18 students of the French department. The x-axis represents the students, and the y-axis indicates the grades out of 10. The chart clearly shows that the AI grading is, overall, higher than the human one. The lowest AI grade is 7.5 while the lowest human grade is 3. Grading paragraphs consistently as 7.5 and higher raises questions of credibility of AI tools.

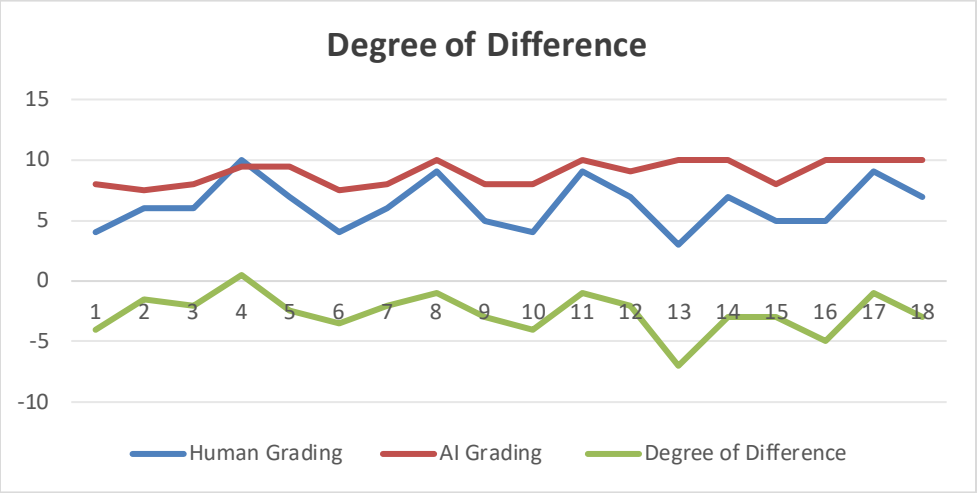


Figure 2. Degree of difference between human and AI assessors

The line chart reflects the degree of difference between human and AI assessment in the context of French department students. The difference scores range between -7 and 0.5, representing a wide variance among assessors. For example, student n14 was given a 3 by the human assessor and a 10 by the AI tool. This suggests that ChatGPT can, sometimes, overlook the assessment criteria provided in the prompt. In other cases, student n4 and student n8 were graded nearly similarly by the two assessors, shrinking the grade gap by 0.5 and 1.

To point out the significant differences between human and AI grading in the French department, the researchers ran several statistical tests. The paired-samples t-test gave a p-value of (> .001). Since the value is below the standard 0.05 significance level, it shows that the difference in average grades (6.39 for human grading and 9.07 for AI grading) is statistically significant. This means the difference is due to there being a consistent gap between the two grading methods.

Effect size was also measured using Cohen's d, which was calculated at −1.55, indicating a very large and substantial difference. The negative value shows that ChatGPT routinely assigned higher grades than the human instructors.

These results show that AI grading produces consistently higher scores than human grading, with differences that are both statistically significant. This sheds light on the impact of AI on assessment outcomes as well as the urgent need for further research regarding grading accuracy.

Findings of the psychology department

Table 2. *Psychology department's students' human and AI grades*

Human Grading	AI Grading	Difference	Degree of Difference
4	4	Match	0
1	3.5	Different	-2.5
9	9	Match	0
9	9	Match	0
1	3.5	Different	-2.5
1	3.5	Different	-2.5
6	6.5	Different	-0.5
9	5.5	Different	3.5
2	3.5	Different	-1.5
6	5	Different	1
8	6.5	Different	1.5
3	5	Different	-2

This chart summarises the results of the Department of Psychology. 12 students were given a writing task, the pieces of writing were graded by a human assessor, who was the teacher of the subject, referred to as Human Grading (Column 1), they were then graded by an AI assessor, which was ChatGPT, referred to as AI Grading (Column 2). Three grades were similar while the other grades were not (column 3). Column 4 represents the degree of difference between Human grading and AI grading, which varies from 3.5 to -2.5.

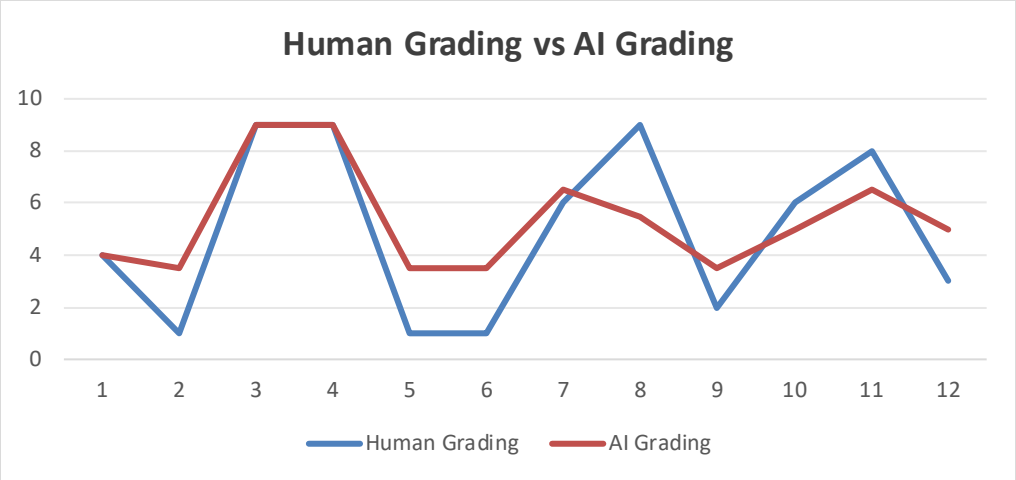


Figure 3. Comparison of human vs. AI grading

The graph represents both Human Grading and AI grading. Two instances from the data show that human grading and AI grading may match and may not. Both the teacher and ChatGPT gave 4 grades for the first writing, which shows that human grading and AI grading may match. The second piece of writing got 1 grade from the teacher and 3.5 grades were given by the AI tool (ChatGPT), which shows that there might be a difference between the grading of teachers and AI tools.

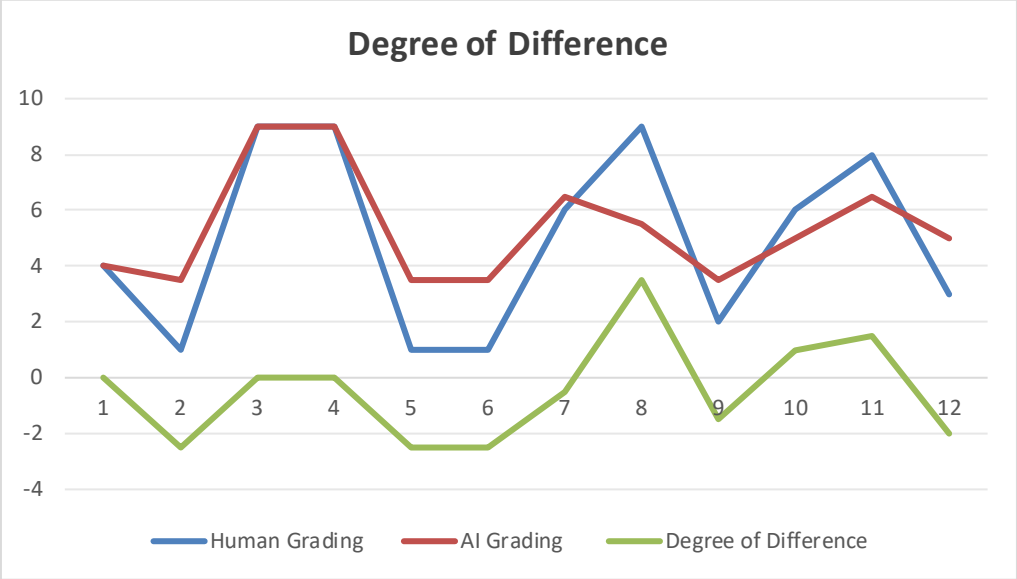


Figure 4. Degree of difference between human and AI assessors

Graph number two is a reflection of the degree of difference between the two grades; the difference ranges between 3.5 and -2.5. For students number 1,3, and 4, the degree of difference is 0 as the grades that were given by the teacher and ChatGPT were, respectively, 4 and 9,9. For student number 8 with a grade of 9 from the teacher and a grade of 5.5 from ChatGPT, the degree of difference is 3.5 grades; the teacher's grade was 3.5 grades higher. On the other hand, student number 2 was given 1 grade by the teacher and 3.5 grades by the AI tool, resulting in a -2.5 degree of difference; the teacher's grading was lower by 2.5 grades compared to the AI tool.

In order to examine whether a significant difference exists between human and AI-assessment in the department of psychology, the researchers conducted a series of statistical analyses. In the first place, a paired t-test was performed to determine significance; the test yielded a p-value of 0.41. Given that this value is greater than the standard 0.05 significance level, the difference in the mean grades, 4.92 for human grading and 5.38 for AI-grading, is not statistically significant. This shows that there is a high probability that the observed difference is due to random chance and not a systematic variation between the two assessment methods.

In addition to significance, the study used Cohen's d to measure the effect size and to quantify the magnitude of the difference. The calculated Cohen's d was −0.25, which is considered a small effect size. This shows that despite the fact that the AI grades were slightly higher on average, the practical difference is minimal. The negative sign simply reflects that the AI's mean grade was higher than the human's. Eventually, while the human grading and AI-grading are not identical, their differences are neither statistically significant nor practically large, suggesting that the AI-grading is comparable to human grading.

Based on the comprehensive analysis of the data, the AI grading system appears to be a reliable alternative to human grading. While there are some minor differences in the average grades and the overall score spread, these are not statistically significant and have a small practical effect. This suggests that ChatGPT can provide grades that are largely consistent with a human grader. Therefore, the AI grading system is a comparable and potentially valuable tool.

Discussion

According to this study's comparison of AI and human grading in the context of Moroccan higher education, AI grades tend to be marginally higher than human-assigned grades. It's interesting to see, in some cases, that ChatGPT provided higher grades, especially when human graders gave lower ones. This aligns with findings of the mixed-method study conducted by El Messaoudi et al. (2025), which shows that students appreciate ChatGPT for its efficiency (73.2% reported time-saving) and idea generation capabilities (68.3% reported frequent benefit), reflecting that ChatGPT is rewarding different qualities than human graders. This study's findings imply that the grading system for AI might be different and is, therefore, incapable of replacing the human aspect of the grading process. This opens the possibility to enhance AI models to be aligned with the human standards of assessment. These concerns were highlighted by Khoual et al. (2025) who used predictive modeling to report that 395 students from various Moroccan faculties reported issues in regard to ethics of AI use in education.

To verify the correctness of AI technology in academic assessment, similar studies were conducted. For instance, Grivokostopoulou et al.'s (2017) work compared the outcomes of AI-assisted assessment with those of manual assessments. In their study, the findings showed a strong correlation and were very comparable. It was found that the results were quite similar when compared to the grade levels provided by the teachers (very low, low, medium, good, and excellent). The instructors, however, identified the highest scores as outstanding, but the AI tool used did not. Moreover, research indicates that AI-assisted assessment can be particularly effective for grammatical accuracy, which might explain some of the discrepancies observed. For instance, a comparison between the two modes of assessment made by Bouziane and Bouziane (2024) indicated that a balanced approach to effective essay correction is suggested by the fact that while ChatGPT performs better in grammar, spelling, sentence structure, relevance, and supporting evidence, human evaluators do better in thematic consistency.

Conclusion

AI assessment tools such as ChatGPT are one of the most remarkable advances that have come to influence the educational field, particularly at the level of assessment. This study was an attempt to compare human assessment with AI assessment. Pieces of writing of two groups of students from the faculty of Letters and Humanities, Dhar El Mahraz, Fes, one group belongs to the department of French language studies, and the other to the department of psychology. Teachers of the two groups, in addition to ChatGPT-4 assessed and graded the pieces of writing provided by the participants. According to the results, the grading of human assessors and AI grading in the department of French language studies are not different, while in the department of psychology, they are similar. In the department of French language studies, ChatGPT produced higher scores than human grading, with differences that are both statistically substantial, while in the department of psychology, the differences between human grading and ChatGPT are neither statistically significant nor practically large, suggesting that AI-grading is comparable to human grading.

The results of this study indicated that AI-assessment tools such as ChatGPT might provide grades that are significantly different in some contexts, while in other contexts, the difference is not statistically significant. Therefore, any use of AI-assessment tools should be implemented with some precaution. Hence, Moroccan teachers, educators, and tool developers should work hand in hand to develop AI-assessment tools that can provide comparable grades even in different

contexts. For a reliable use of AI-assessment tools, they should provide grades that align with human standards of assessment. Besides, teachers and educators are encouraged to dive into the art of prompting, as it is a crucial aspect for tools such as ChatGPT, as it might affect the whole process of assessment and grading. Moreover, stakeholders in Moroccan higher education, including teachers, educators, administrative staff, and tool developers, should benefit from initiatives such as the National Strategy: Digital Morocco 2030. One of the aims of this initiative is to use AI to enhance educational outcomes, foster digital literacy, and improve the efficiency of academic institutions.

The focus of the study was to compare human and AI assessment tools in grading writing, with special focus on grammar, spelling, punctuation, clarity, and structure. Future research could address this comparison, concentrating on aspects such as cohesion, coherence, vocabulary, and conciseness. Researchers are recommended to address research questions, such as to what extent AI and human assessment are similar or different in assessing cohesion in writing? To what extent is AI similar or different from teachers in assessing conciseness in writing? Besides, researchers are encouraged to address different aspects of EFL, like comprehension and grammar, in their comparison between AI and human assessment. They could address research questions similar to the following: In what way are AI and human assessment tools similar or different in assessing comprehension? To what extent does AI assess grammar similarly or differently from human assessment?

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 was used to enhance the readability and quality of the paper.

Authors Biography

Hajar Ennadir is a second year PhD candidate at the Faculty of Letters and Human Sciences, Dhar El Mahraz, Fez, Morocco, with a focus on education, EFL teaching and learning, AI in education, and Academic writing. She serves as a part time EFL instructor. ORCid ID: https://orcid.org/0009-0003-8121-4640

Salma Habiballah is a PhD holder in applied linguistics. Her research interests include examining the effectiveness of online education as perceived by both teachers and students, as well as discovering the ways different members of educational boards use to enhance their digital learning experiences. ORCid ID: https://orcid.org/0000-0002-4122-7477

References

Aljuaid, H. (2024). The impact of artificial intelligence tools on academic writing instruction in higher education: A systematic review. *Arab World English Journal (AWEJ) Special Issue on ChatGPT*, 26-55. https://dx.doi.org/10.24093/awej/ChatGPT.2

Appiah, M., & Van Tonder, F. (2018). E-assessment in higher education: A review. *International Journal of Business Management & Economic Research*, *9*(6), 1454-1460.

Bai, J. Y., Zawacki-Richter, O., Bozkurt, A., Lee, K., Fanguy, M., Cefa Sari, B., & Marín, V. I. (2022). Automated essay scoring (AES) systems: Opportunities and challenges for open and distance education. *In Proceedings of The Tenth Pan-Commonwealth Forum on Open Learning (PCF10)*. https://doi.org/10.56059/pcf10.8339

Chen, X., Zou, D., Cheng, G., & Xie, H. (2020). Detecting latent topics and trends in educational technologies over four decades using structural topic modeling: A retrospective of all volumes of Computers & Education. *Computers & Education*, *151*, 103855. https://doi.org/10.1016/j.compedu.2020.103855

Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two Decades of Artificial Intelligence in Education: Contributors, Collaborations, Research Topics, Challenges, and Future Directions. *Educational Technology & Society*, 25 (1), 28-47.

Douali, L., Selmaoui, S., & Bouab, W. (2022). Artificial intelligence in education: Fears and faiths. *International Journal of Information and Education Technology*, 12(7), 650-657. doi: 10.18178/ijiet.2022.12.7.1666

El Messaoudi, M., Lamiae, A., & Daoudi. R. (2025). Student Satisfaction with ChatGPT as a Writing Tool in Moroccan Higher Education: A Mixed Methods Study. *Arab World English Journal (AWEJ) Special Issue on Artificial Intelligence:*193-219. https://dx.doi.org/10.24093/awej/AI.11

Gierl, M. J., Latifi, S., Lai, H., Boulais, A. P., & De Champlain, A. (2014). Automated essay scoring and the future of educational assessment in medical education. *Medical education*, *48*(10), 950-962. https://doi.org/10.1111/medu.12517

González-Calatayud, V., Prendes-Espinosa, P., & Roig-Vila, R. (2021). Artificial intelligence for student assessment: A systematic review. *Applied sciences*, *11*(12), 5467. https://doi.org/10.3390/app11125467

Hespress En. (2024, May 31). "Study: 75% of students see AI as an asset for learning". https://en.hespress.com/85975-study-75-of-students-see-ai-as-an-asset-for-learning.html

Huang, A. Y., Lu, O. H., & Yang, S. J. (2023). Effects of Artificial Intelligence–Personalized Recommendations on Learners' Learning Engagement, Motivation, and Outcomes in a Flipped Classroom. *Computers & Education*, *194*, 104684. https://doi.org/10.1016/j.compedu.2022.104684

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European journal of education*, *57*(4), 542-570. https://doi.org/10.1111/ejed.12533

Jen, S. L., & Salam, A. R. (2024). Using Artificial Intelligence for Essay Writing. *Arab World English Journal (AWEJ) Special Issue on ChatGPT*, April 2024: 90-99. https://dx.doi.org/10.24093/awej/ChatGPT.5

Khoual, M., Elkaimbillah, Z., Mcharfi, Z., & El Asri, B. (2025). AI in Moroccan education: Evaluating student acceptance using machine learning classification models. *Indonesian Journal of Electrical Engineering and Computer Science*, 37(1), 452-462. http://doi.org/10.11591/ijeecs.v37.i1.pp452-462

Lin, CC., Huang, A.Y.Q. & Lu, O.H.T. Artificial intelligence in intelligent tutoring systems toward sustainable education: a systematic review. *Smart Learn. Environ.* 10, 41 (2023). https://doi.org/10.1186/s40561-023-00260-y

Maulana, A., Idroes, G. M., Kemala, P., Maulydia, N. B., Sasmita, N. R., Tallei, T. E., ... & Rusyana, A. (2023). Leveraging Artificial Intelligence to Predict Student Performance: A

Comparative Machine Learning Approach. *Journal of Educational Management and Learning, 1*(2), 64-70. https://doi.org/10.60084/jeml.v1i2.132

Maryam, F., Jamaa, O. (2025). An Empirical Assessment of Moroccan EFL Teachers' Use of Generative AI for EFL Formal Assessment: An Intervention Study. *Social Sciences*, 14(1), 8-21. https://doi.org/10.11648/j.ss.20251401.12

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

Nazaretsky, T., Ariely, M., Cukurova, M., & Alexandron, G. (2022). Teachers' trust in AI-powered educational technology and a professional development program to improve it. *British Journal of Educational Technology, 53*(4), 914-931. https://doi.org/10.1111/bjet.13232

Podsiad, M., & Havard, B. (2020). Faculty acceptance of the peer assessment collaboration evaluation tool: A quantitative study. *Educational Technology Research and Development, 68*, 1381-1407. https://doi.org/10.1007/s11423-020-09742-z

Rincon-Flores, E. G., Lopez-Camacho, E., Mena, J., & Olmos, O. (2022). Teaching through Learning Analytics: Predicting Student Learning Profiles in a Physics Course at a Higher Education Institution, 82-89. https://doi.org/10.9781/ijimai.2022.01.005

Sumakul, D. T. Y., Hamied, F. A., & Sukyadi, D. (2022). Artificial intelligence in EFL classrooms: Friend or foe?. *LEARN Journal: Language Education and Acquisition Research Network, 15*(1), 232-256.

U.S. Department of Education, Office of Educational Technology, Artificial Intelligence and Future of Teaching and Learning: Insights and Recommendations, Washington, DC, 2023.

Uyar, A. C., & Büyükahıska, D. (2025). Artificial intelligence as an automated essay scoring tool: A focus on ChatGPT. *International Journal of Assessment Tools in Education*, 12(1), 20-32. https://doi.org/10.21449/ijate.1517994

Yavuz, F., Çelik, Ö., & Yavaş Çelik, G. (2025). Utilizing large language models for EFL essay grading: An examination of reliability and validity in rubric-based assessments. *British Journal of Educational Technology*, 56(1), 150-166. https://doi.org/10.1111/bjet.13494

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic Review of Research on Artificial Intelligence Applications in Higher Education–Where are the Educators?. *International journal of educational technology in higher education, 16*(1), 1-27. https://doi.org/10.1186/s41239-019-0171-0

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.188-210
DOI: https://dx.doi.org/10.24093/awej/vol16no3.11

Global Trends in Language for Specific Purposes Curriculum Development: A Systematic Review

Mior Syazril Mohamed Sapawi
Faculty of Education, The National University of Malaysia, Bangi, Malaysia
Corresponding Author: p121100@siswa.ukm.edu.my

Nik Mohd Rahimi Nik Yusoff
Faculty of Education, The National University of Malaysia, Bangi, Malaysia
Email: nik@ukm.edu.my

Received:06/06/2025 Accepted: 08/20/2025 Published: 09/20/2025

Abstract
This systematic review investigates how language curriculum design is evolving to address the increasing need for domain-specific communication skills in education and professional practice. It examines the question of how educational systems worldwide adapt curriculum frameworks to meet contextual and disciplinary requirements while maintaining linguistic accuracy and pedagogical quality. The study also highlights the significance of this shift in enhancing equity, relevance, and global competitiveness in language for specific purposes education. Guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses protocol, 24 peer-reviewed articles published between 2021 and 2025 were systematically retrieved from Scopus, Web of Science, and the Education Resources Information Center databases. Thematic synthesis of the selected studies identified three major trends: the contextual localisation of curriculum models and frameworks, the adoption of learner-centred teaching strategies to strengthen skill transfer, and the integration of technology and contextual awareness in curriculum delivery. These findings point to a global movement toward flexible and interdisciplinary approaches in language curriculum development. The review concludes with recommendations for policymakers and educators to align teacher preparation and technological infrastructure with international benchmarks while addressing local learning priorities.
Keywords: Curriculum design, educational technology, language for specific purposes, pedagogical innovation, systematic review, teaching strategies

How to cite: Sapawi, M.S.M., & Yusoff, N. M.R.N. (2025). Global Trends in Language for Specific Purposes Curriculum Development: A Systematic Review. *Arab World English Journal, 16*(3): 188-210. https://dx.doi.org/10.24093/awej/vol16no3.11

Introduction

In today's globalised academic and professional landscape, the need for learners to communicate effectively within domain-specific contexts has become a critical focus in language education (Fitria, 2020; Purwanto & Nurhamidah, 2021). This demand has accelerated the growth of Language for Specific Purposes (LSP) as a major subfield in applied linguistics and language pedagogy. LSP encompasses a range of instructional domains, including English for Specific Purposes (ESP), Arabic for Specific Purposes (ASP) and other tailored language models that address the communicative needs of specialised communities (Enríquez et al., 2023; Kurniawan & Sumani, 2022; Morell, 2020; Reynolds et al., 2023; Rus, 2020). As professional and industrial domains become increasingly complex, the ability to navigate technical discourse in fields such as medicine, engineering, and law reinforces the urgency of developing robust, context-sensitive language curricula. Effective LSP curriculum design must therefore transcend linguistic form and incorporate pragmatic, sociocultural, and institutional factors that shape communication within each discipline (Burston et al., 2017; Cobb & Bower, 2021; Finch, 2023; Katsaounis & Steinmüller, 2021).

Despite this momentum, research on LSP curriculum remains fragmented and inconsistent across educational contexts and regional boundaries. Many studies emphasise localised teaching practices or narrowly scoped innovations without providing a broader curricular perspective that could guide educational policy and instructional coherence (Cha et al., 2019; Meijer & Bekkers, 2015; Oravcová et al., 2022). Instructional methodologies in LSP are diverse, ranging from task-based and genre-informed approaches to corpus-based and content-integrated frameworks, which often leads to conceptual ambiguity and varied pedagogical implementations (Bessadet, 2022; Czeczulin & Greydak, 2020; Freeman, 2017; Kazakova, 2015; Piri & Haxhillazi, 2023). This variability highlights the lack of a unified framework capable of aligning policy goals with effective classroom practice in LSP curriculum development (Cerezo Herrero et al., 2021; Dursun, 2023; Haque, 2019; Tirizite & Marjane, 2023).

This systematic review addresses this gap by consolidating recent international research to inform policymakers, educators, and curriculum developers on strategies for enhancing relevance, equity, and innovation in LSP curriculum design. This includes examining how curriculum models are adapted to different regional and disciplinary contexts, identifying teaching practices that enhance learner outcomes and assessing how technological and contextual factors influence curriculum implementation. These aims were operationalised through the following research questions:

1. How are curriculum models and frameworks adapted to address disciplinary and regional needs in LSP education?
2. What pedagogical strategies and teaching practices enhance learner engagement and skill development in LSP contexts?
3. In what ways do technological and contextual factors influence the design, relevance, and implementation of LSP curricula across educational systems?

Literature Review

Curriculum development in Language for Specific Purposes (LSP) has received increasing attention as language educators respond to the highly specialised communicative demands of academic and professional domains. This aligns with broader educational paradigms that call for curricula addressing not only linguistic competencies but also critical literacies such as sustainability, interdisciplinarity, and digital fluency (Arslan & Curle, 2021; Sanosi et al., 2025). Sanosi et al. (2025) observed that "curricula must

extend beyond language skills to embed future-oriented competencies such as sustainability and digital literacy," and demonstrated how English for Specific Purposes (ESP) syllabi are being reformulated to align with the Sustainable Development Goals (SDGs), signalling a curricular shift towards future-oriented competencies. Similarly, Arslan and Curle (2021) identified the presence of sustainability-related content in Turkish English language textbooks, though implementation remains limited. Recent evidence from Bessadet (2022) further illustrates how innovative pedagogical strategies, such as drama-based teaching, can create more engaging and context-specific learning experiences, which is a priority in LSP curriculum design.

Alongside global agendas, the relevance of learner needs has become central in shaping curriculum design. Generoso and Arbon (2020) reported significant discrepancies between syllabus expectations and students' real-world proficiency in their review of English for Academic Purposes (EAP) curricula in the Philippines. This concern was echoed by Carreira and Kagan (2011), who highlighted the need for community-responsive curricula for heritage language learners in the United States. Haque (2019) emphasised that "learner autonomy is at the heart of effective language curriculum design," reinforcing the importance of equipping students with transferable skills that extend beyond linguistic accuracy, a core consideration in the design of effective LSP curricula. The tension between global curriculum standards and local realities is further exemplified by Zare-ee and Hejazi (2018), who found that internationally produced textbooks consistently outperformed local materials in terms of clarity and cohesion, raising concerns about content alignment and instructional coherence.

Teacher agency has also emerged as a critical dimension in LSP curriculum development. Jurkovič et al. (2023) designed a professional development framework for LSP educators across multiple European contexts, underscoring the need to strengthen teachers' curricular design skills in addition to language proficiency. However, institutional constraints and top-down reforms often hinder such efforts. Lei and Medwell (2020) observed that English teachers in China experienced conceptual ambiguity regarding curriculum reforms, limiting their ability to adapt or innovate in classroom practice. Similarly, Sierra-Piedrahita and Echeverri-Sucerquia (2020) documented how the exclusion of teachers from curriculum planning processes in Colombia weakened implementation and reduced contextual fit.

Technological advancement has further transformed the delivery of LSP curricula, particularly in response to remote learning needs and digital literacy imperatives. Wang (2014) advocated for internet-based tools to enhance student autonomy and encourage constructivist learning environments in tertiary English instruction. These findings can be further contextualised using established technology-integration frameworks. Rogers et al. (2019) in Diffusion of Innovations describe how educators adopt new technologies through a sequence of awareness, interest, evaluation, trial, and adoption. Similarly, Granić (2024) highlights that Levels of Adoption are critical in understanding how individuals integrate technology effectively in different contexts. The Five-Stage Model by Moule (2007) provides a developmental lens for understanding how teachers evolve from familiarisation to innovation when embedding technology in LSP curricula. Rogers et al. (2019) note that "the rate of adoption is highly influenced by perceived value and institutional support," which explains why many institutions struggle to sustain effective digital integration despite widespread availability of resources. The TUFS project in Japan (Kigoshi et al., 2003), which introduced multilingual e-learning modules, exemplifies how digital platforms can scale across linguistic boundaries. Kelly and Hou (2022) further demonstrated how machine translation can empower learners in English-as-an-additional-language contexts, highlighting the potential for such tools to support autonomy and translanguaging practices in LSP education. However, Zucker et al. (2019) cautioned that the scaling of digital interventions often results in uneven outcomes, primarily due to contextual variability and inadequate teacher preparation.

Cultural content and curriculum authenticity also remain central to effective LSP design. Syamsi et al. (2024) reported positive learning outcomes from using a culture-based Indonesian textbook for non-native speakers, suggesting that culturally relevant materials enhance learner engagement and motivation. In Mexico, Balderas et al. (2024) called for more explicit literacy instruction in English and

Spanish within higher education curricula, emphasising the importance of linguistic equity. Yet structural limitations persist. Bardi (2013) criticised the lack of coherence between textbook content and evaluation criteria in curriculum policy, while Koglbauer (2022) noted how prescriptive reforms in the UK marginalised teacher voices in curriculum planning.

Despite the diversity of research on LSP curricula, significant gaps remain. Fragmentation continues to characterise the field, particularly with respect to cross-contextual analysis and the development of transferable curriculum models. While individual studies offer valuable insights, they often remain disconnected, limiting the formation of a coherent knowledge base that can guide policy or practice. This review addresses this gap by synthesising recent international findings to identify dominant trends and propose more inclusive and pedagogically sound approaches to LSP curriculum design. To frame this review, concepts from technology integration theory (Granić, 2024; Rogers et al., 2019) and curricular contextualisation (Jurkovič et al., 2023; Zare-ee & Hejazi, 2018) provide the analytical scaffolding through which studies were synthesised.

Method

This study adopted a Systematic Literature Review (SLR) approach to examine global trends in curriculum development for Arabic language and medical terminology integration. The review process was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) protocol as outlined by Page et al. (2021), ensuring transparency and replicability at each stage. The review was conducted across three major databases: Scopus, WoS and ERIC, selected for their comprehensive indexing of high-quality research in education and language studies. The full screening process is visually presented in the PRISMA flow diagram as Figure One, covering the phases of identification, screening, eligibility and final inclusion.

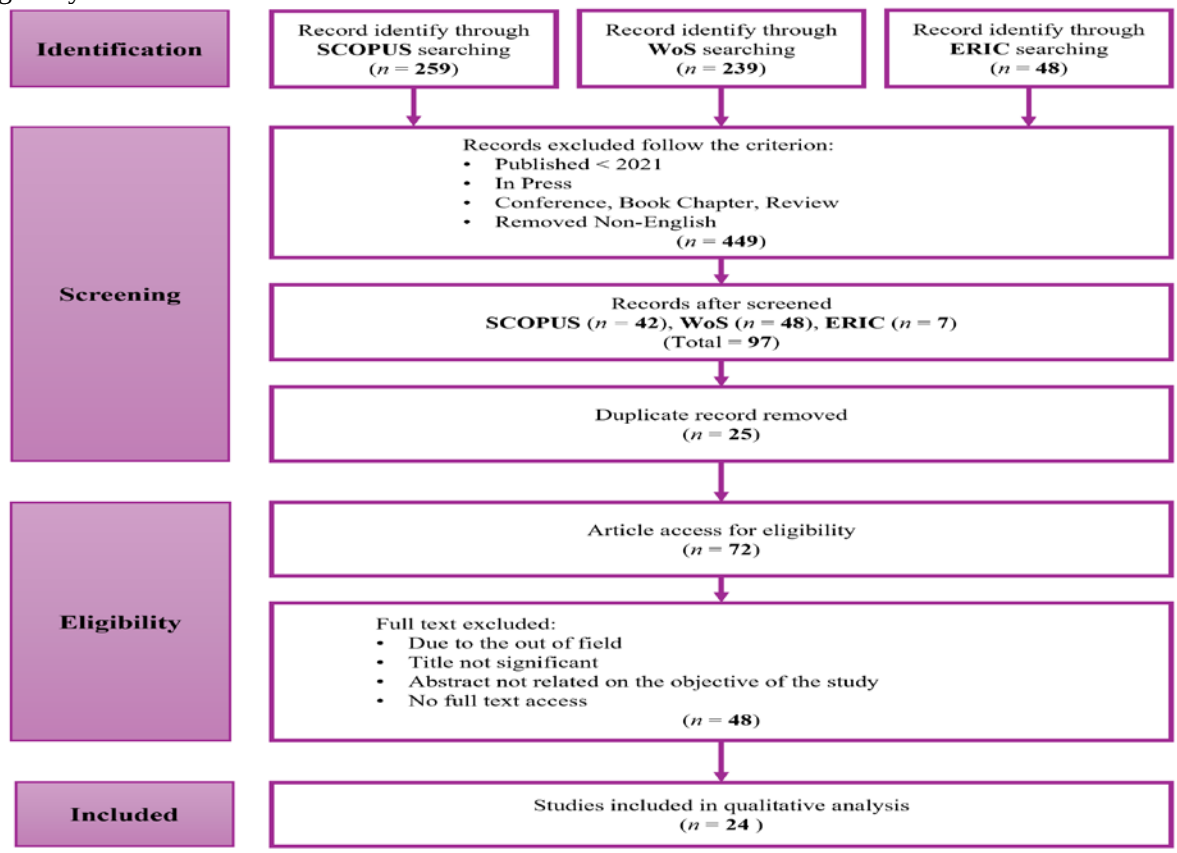


Figure 1. Flow diagram of the proposed searching study

Participants

In the context of this review, the term "participants" refers to the peer-reviewed journal articles that were selected and analysed as primary data sources. A total of 546 articles were initially identified through systematic searches of three reputable databases: Scopus (259), WoS (239) and ERIC (48). Following this, a rigorous screening process was conducted to remove duplicates and exclude irrelevant or low-quality studies. After applying the inclusion and exclusion criteria, only 72 unique articles remained. Further evaluation based on thematic relevance and availability of full text resulted in a final selection of 24 articles for qualitative synthesis. The breakdown of articles by source is presented in Appendix A (Table A1).

Instruments

Several tools and frameworks were employed to ensure systematic and replicable results. First, a structured search string using Boolean operators was developed to retrieve relevant studies from each database. The search terms included combinations of keywords such as "curriculum," "syllabus," "LSP," "ESP," "language," "development," and "trends." Although terms like "TPACK" and "blended learning" are relevant in broader LSP-technology literature, they were excluded in this review to maintain a focused scope on curriculum-specific terminology and frameworks. The exact search strings used for each database are displayed in Table One.

Table 1. *The search string*

DATABASE	SEARCH STRING
SCOPUS	TITLE-ABS-KEY ((curriculum OR syllabus OR textbook OR module) AND develop* AND (language OR lsp OR asp OR esp) AND (purpose* OR goal OR idea) AND trends) AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2022) OR LIMIT-TO (PUBYEAR , 2023) OR LIMIT-TO (PUBYEAR , 2024) OR LIMIT-TO (PUBYEAR , 2025)) AND (LIMIT-TO (LANGUAGE , "English")) AND (LIMIT-TO (SUBJAREA , "SOCI") OR LIMIT-TO (SUBJAREA , "ARTS")) AND (LIMIT-TO (SRCTYPE , "j")) Date of Access: May 2025
WoS	(curriculum OR syllabus OR textbook OR module) AND develop* AND (language OR lsp OR asp OR esp) AND (purpose* OR goal OR idea) AND trends (All Fields) and 2025 or 2024 or 2023 or 2022 or 2021 (Publication Years) and Article (Document Types) and English (Languages) and Education Educational Research or Linguistics (Web of Science Categories) and Education Educational Research or Linguistics (Research Areas) Date of Access: May 2025
ERIC	(curriculum OR syllabus OR textbook OR module) AND develop* AND (language OR lsp OR asp OR esp) AND (purpose* OR goal OR idea) AND trends naepubyearmin:2021 pubyearmax:2025 Date of Access: May 2025

To determine which articles would be included in the final analysis, specific selection criteria were established prior to the screening phase. These criteria were based on publication date, language, document type and thematic relevance. The full set of inclusion and exclusion criteria is outlined in Table Two.

Table 2. *The selection criterion is searching*

CRITERION	INCLUSION	EXCLUSION
Language	English	Non-English
Timeline	2021 – 2025	< 2021
Literature Type	Journal (Article)	Conference, Book, Review
Publication Stage	Final	In Press

ISSN: 2229-9327

To ensure quality and mitigate methodological bias, each selected study was appraised using a six-criterion framework adapted from Abouzahra et al. (2020). These criteria (QA1–QA6) included: (1) clarity of purpose; (2) relevance and usefulness; (3) methodological transparency; (4) conceptual definition; (5) comparative analysis; and (6) acknowledgement of limitations. A three-point scale was used (3 = strong, 2 = moderate, 1 = weak). In certain cases, partial scores (e.g., 0.5) were assigned when an article demonstrated partial fulfilment of a criterion, such as when comparative insights were mentioned without full empirical grounding.

The scoring rubric was applied independently by two reviewers, followed by reconciliation through discussion. This process served as the study's risk of bias assessment, reducing evaluator subjectivity and increasing inter-rater reliability. The quality scores for each study are provided in Appendix B (Table B1) and only a summary reference is made in the main text.

Procedures

The procedures followed in this review comprised four main phases as outlined by the PRISMA protocol: identification, screening, eligibility assessment and final inclusion. During the identification stage, database-specific search queries were executed using the pre-established search strings. The date of access for all searches was May 2025. During the screening phase, duplicate articles were removed, and titles and abstracts were reviewed to eliminate studies that did not meet the basic inclusion criteria. In the eligibility phase, full-text articles were retrieved and examined in detail to assess their direct relevance to LSP curriculum development. Articles that lacked empirical or theoretical significance or were inaccessible in full text were excluded. The final phase involved extracting key information from the 24 included studies and organising their findings into emergent thematic categories. These themes subsequently structured the analytical framework for the results and discussion sections of the paper.

Result

This review analysed 24 peer-reviewed articles published between 2021 and 2025 that addressed curriculum development in LSP. The thematic synthesis identified three core domains across the selected studies: (1) curriculum models and LSP frameworks; (2) teaching practices and pedagogical strategies in LSP; and (3) technological and contextual adaptations of LSP curricula. The distribution of studies under each theme is provided in Appendix C (Table C1).

Curriculum Models and LSP Frameworks

Efforts to develop effective LSP curricula consistently highlighted the need to adapt global models to local linguistic, disciplinary and cultural contexts. In Japan, the revised national Model Core Curriculum underscored a competency-based approach rooted in local educational values requiring interpretive reconstruction of global terminology (Urushibara-Miyachi et al., 2024). In South Africa, van Rooyen et al. (2024) examined the integration of multilingual skills into medical education revealing student concerns about the cultural relevance of isiXhosa instruction and underscoring the importance of curricular responsiveness. Similarly, Valle Coto et al. (2024) showed that early and inclusive equity-based instruction significantly improved students' language awareness and application in pre-clinical settings.

Technological integration within curriculum design appeared across several studies. Hegrenæs et al. (2022) redesigned a legal translation course in Norway using translation competence theory signalling a shift toward outcome-based education. In China, interpreting

programs analysed through the GETAMEL model revealed how tablet-based notetaking adoption hinged on instructor readiness suggesting the need for formalised digital literacy components in curriculum (Tian & Wang, 2024). Tyron et al. (2021) applied psycho-pedagogical principles in a modular course for ship engineers aligning technical literacy with cognitive development and industrial demands.

The alignment of professional expectations with LSP curricular frameworks was further demonstrated by Hegrenæs et al. (2022) whose postgraduate legal translation modules fostered critical thinking and accuracy through authentic tasks. A recurring challenge, however, involved aligning theoretical goals with classroom practice. Namanyane and Shaoan (2025) showed how teachers in China adapted their pedagogical content knowledge to implement communicative approaches within the ELCS. Valle Coto et al. (2024) similarly reported that equity-focused interventions enhanced learner preparedness for multilingual settings. Namanyane and Shaoan (2025) offered additional evidence from China showing that despite policy support, teachers faced difficulties executing role-play activities effectively due to a mismatch between curriculum expectations and classroom realities.

In summary, three major directions emerged: (1) localisation of global curriculum models through cultural and linguistic contextualisation, (2) integration of technological tools to support instructional outcomes and (3) alignment between educator competence and curricular innovation.

Teaching Practices and Pedagogical Strategies in LSP

Pedagogical strategies in LSP are evolving rapidly with the rise of digital technologies and artificial intelligence prompting a shift toward more learner-responsive designs. In China, EFL students' difficulties in blended learning revealed limitations in learner autonomy, engagement and instructional coherence (Cao et al., 2024). Complementing this, Meniado (2023) highlighted ChatGPT's potential in personalising writing instruction though concerns over ethical risks and dependency remained. Similarly, F. F. Wang et al. (2025) found that interactive AI scaffolding could boost motivation and self-regulation when instruction matched learner developmental needs.

Digital pedagogies are also redefining teacher competencies in LSP. Meniado (2023) introduced Digital Language Teaching 5.0 aligning with Industry 5.0's focus on adaptive technologies and human-centred innovation. Huynh and Nguyen (2021) reported that Vietnamese teachers responded positively to CALL training but often struggled to apply it effectively indicating gaps in institutional support. Medvedeva et al. (2022) added to this discourse by proposing a system to monitor soft skills development showing how LSP teaching is expanding to integrate transversal competencies like leadership and teamwork.

Teacher readiness remains a crucial factor in LSP delivery. Huynh and Nguyen (2021) found that although in-service educators felt confident using technology, structured training remained scarce. This highlighted the value of targeted professional development to build instructional confidence and curriculum alignment.

Contextualised strategies also play a vital role in enhancing language performance and professional identity. In legal education, Chovancová (2024) demonstrated how mediation tasks supported linguistic negotiation through role-play. Likewise, Savula et al. (2024) documented the benefits of CLIL, flipped classrooms and virtual simulations in improving cognitive engagement and specialised communication. Beyond technology, motivation and learner diversity emerged as key themes. Zhou et al. (2023) showed that EMI learners with lower language proficiency often compensated through metacognitive preparation indicating motivation was more predictive of

success than proficiency alone. Porto et al. (2021) similarly revealed tensions between national policy and classroom realities in Argentina reinforcing the importance of differentiated instruction and curriculum decentralisation.

In summary, LSP pedagogy is shifting toward integration of AI, learner-centred tasks and digital literacy while challenges persist in teacher preparation, ethical usage and contextual curriculum adaptation.

Technological and Contextual Adaptation in LSP Curriculum

This theme highlights the intersection between technological advancement and local context in shaping LSP curriculum design. In Chile, a Business English course was developed based on needs analysis revealing student preferences for smaller groups, interactive delivery and globally relevant content (Lasekan et al., 2023). Similarly, Shen and Lu (2024) introduced a multimedia animation-based platform for mechanical drawing in China which improved learners' conceptual clarity and engagement in technical subjects. In Taiwan, Chung and Chen (2021) applied text mining to identify competency demands in the HR industry enabling curriculum designers to align course content with labour market expectations. These studies support the idea that technology in LSP should be both pedagogically grounded and market responsive. However, the successful implementation of such innovations often depends on systemic factors such as institutional leadership, availability of digital infrastructure and administrative willingness to support curricular change.

Contextual adaptation was also crucial in community-based and inclusive learning contexts. Diller and Fraiser (2022) documented how mother tongue preservation among the Prai community in Thailand flourished when integrated into locally designed curricula. In Kenya, Adhiambo et al. (2025) illustrated how visual symbol strategies enhanced expressive communication for hearing-impaired learners demonstrating how curriculum can meet diverse learner needs. Similarly, Moscovici et al. (2022) found that biomedical engineering students in Argentina perceived higher motivation and identity development when engaging in Wikipedia writing tasks in local languages suggesting the power of public-facing curriculum activities. Yet, these practices also rely on systemic enablers including school policies that recognise diversity, support from programme coordinators and allocation of time and space for alternative instructional models. Without institutional backing, innovative efforts risk becoming isolated practices rather than scalable solutions.

The social and emotional dimensions of curriculum design were emphasised in multiple studies. For instance, Perlado Lomo de Espinosa and Trujillo Vargas (2024) highlighted that assertive communication training could serve as a preventive measure against student violence while enhancing critical thinking. Their empirical findings further demonstrated that embedding social communication skills within the curriculum helped address behavioural challenges in educational settings. Complementary perspectives were offered by Adhiambo et al. (2025) and Diller with Fraiser (2022) who illustrated how inclusive and culturally grounded practices can reshape learners' perceptions of literacy and empower local knowledge systems. Collectively, these findings suggest that while technological and contextual strategies are promising, their long-term impact is contingent on addressing systemic barriers such as policy inertia, underfunding and inconsistent leadership engagement.

In summary, these studies affirm that LSP curriculum must embrace both technological innovation and contextual responsiveness. A multidimensional approach that incorporates

leadership support, resource accessibility, multimedia integration, community-based learning, emotional intelligence and real-world alignment enables LSP instruction to be more inclusive, relevant and future-ready.

Discussion

This review set out to examine three core areas: the frameworks guiding LSP curriculum design (RQ1), the pedagogical practices employed (RQ2) and the role of technology and context in shaping curriculum development (RQ3). The findings revealed that curriculum models are increasingly contextualised to disciplinary and cultural needs; pedagogical strategies are shifting towards learner-responsive, technology-enhanced methods; and curricular innovations are deeply influenced by both institutional leadership and sociocultural realities. These insights provide a multi-layered understanding of current LSP practices and inform the discussion that follows.

This systematic review synthesised 24 peer-reviewed studies to offer a global overview of recent developments in LSP curriculum design. The thematic analysis revealed three prominent and interconnected domains: (1) Curriculum Models and LSP Frameworks, (2) Teaching Practices and Pedagogical Strategies and (3) Technological and Contextual Adaptation. These themes collectively reflect the evolving demands placed on curriculum designers, educators and institutions to ensure relevance, inclusivity and innovation in LSP education.

Findings under the first theme highlight a global shift toward contextualising curriculum frameworks to fit specific disciplinary, linguistic and regional needs. For example, the Japanese Model Core Curriculum restructured medical education by integrating locally informed competencies while aligning with international standards (Urushibara-Miyachi et al., 2024). In South Africa, language offerings such as isiXhosa were integrated into undergraduate medical training, yet student feedback revealed limitations in cultural and pedagogical relevance (Rooyen et al., 2024). This underlines the need for curricular adaptation that not only reflects societal needs but also accommodates learner expectations. Similarly, the work of Valle Coto et al. (2024) in integrating equity-focused instruction into pre-clinical education demonstrated measurable gains in language awareness affirming the value of inclusive and structured curriculum design.

The second theme illustrated a growing pedagogical transition toward learner-centred practices facilitated by digital technologies and personalised instruction. Studies on artificial intelligence-based tools such as ChatGPT (Asad et al., 2024) and interactive scaffolding systems (Wang et al., 2025) revealed increased learner motivation and self-regulation when instructional design was aligned with student needs. However, findings from Huynh and Nguyen (2021) indicated that while in-service teachers expressed openness toward Computer-Assisted Language Learning (CALL), a lack of structured professional development hindered classroom application. These challenges reflect limited application of the Technological Pedagogical Content Knowledge (TPACK) framework, which emphasises the intersection of content expertise, pedagogical skills and technological fluency. Without comprehensive TPACK training, teacher integration of digital tools remains superficial and inconsistent. Medvedeva et al. (2022) expanded on this by introducing digital tracking of soft skills suggesting a broader shift in LSP pedagogy toward transversal competencies such as teamwork and leadership. Meniado (2023) advocated for Digital Language Teaching 5.0 rooted in Industry 5.0 which calls for adaptive and human-centred technology integration in teacher preparation. However, as with Huynh and Nguyen, Meniado's findings also point to the absence of systemic enablers, particularly school leadership and institutional planning, as critical barriers to implementation. Leadership support, especially from

principals and programme coordinators, is central to facilitating teacher readiness and sustaining innovation in LSP instruction.

The third theme explored how LSP curricula respond to contextual variables especially in marginalised or diverse learning environments. Adhiambo et al. (2025) demonstrated that symbol-based instruction enhanced communication in students with hearing impairments. In Thailand, Diller and Fraiser (2022) found that curriculum grounded in local language and identity supported minority literacy development. Moscovici et al. (2022) observed that task-based assignments such as Wikipedia editing in biomedical contexts, improved learner self-efficacy and academic motivation. In a broader socio-emotional context, Perlado Lomo de Espinosa and Trujillo Vargas (2024) argued for embedding assertive communication into curriculum as a preventive mechanism against student violence as further elaborated in their own empirical findings. Despite policy endorsement for learner-centred models in China, educators faced barriers translating those policies into practice due to systemic constraints (Namanyane & Shaoan, 2025). This tension reinforces the need for decentralised planning and institutional flexibility. These findings mirror earlier research that situates LSP reform within a broader ecosystem of governance, school culture and administrative responsiveness. Without institutional frameworks that promote equity, resource access and curricular freedom, contextually driven innovations remain fragmented and unsustainable.

These findings collectively reveal that the three themes are interconnected and form a cohesive response to modern curricular challenges in LSP. The emphasis on curriculum localisation demonstrates that meaningful instruction is grounded in disciplinary logic, institutional identity and learner reality (Urushibara-Miyachi et al., 2024; Valle Coto et al., 2024). Similarly, the widespread use of digital platforms and learner-responsive methods underscores the growing necessity for flexible instructional environments (Medvedeva et al., 2022; F. F. Wang et al., 2025). Contextual sensitivity is not merely a pedagogical advantage but a core component of equity and social impact in curriculum design (Adhiambo et al., 2025; Diller & Fraiser, 2022). Additionally, as shown in Porto et al. (2021), decentralisation was essential for empowering teachers to meet learner diversity across different systems. The present review thus aligns with and extends previous research by highlighting the systemic and institutional prerequisites, such as leadership engagement, infrastructure provision and interdepartmental collaboration, that shape the success of curriculum innovation in LSP settings.

The implications of this review are threefold. First, curriculum development should be informed by interdisciplinary collaboration, involving language specialists, content experts, and local stakeholders. Second, teacher education programmes must include digital fluency, critical pedagogy, and adaptive design to prepare educators for changing learning environments. Third, educational policies should promote equitable access to instructional technologies, particularly those proven to support marginalised learners such as tablet-supported notetaking (Tian & Wang, 2024) and animation-enhanced instruction (Shen & Lu, 2024). Beyond access, the sustainability of these innovations depends on school leadership buy-in, institutional incentives, and supportive monitoring mechanisms. Continuous curriculum evaluation, informed by learner feedback and performance data, remains essential to ensure pedagogical relevance and long-term sustainability.

Nevertheless, several limitations should be acknowledged. The review was limited to English-language sources from Scopus, WoS, and ERIC, potentially excluding valuable non-English or region-specific perspectives. Additionally, the synthesis was based primarily on abstract-level content, which may not fully reflect methodological intricacies or empirical richness.

Furthermore, limited benchmarking across contexts (QA5) suggests a need for comparative frameworks that situate findings within broader global trends. Future research should focus on longitudinal studies exploring LSP curriculum reform, especially those integrating artificial intelligence, multilingual education, and inclusive teaching practices. Greater scholarly attention is also needed in underrepresented regions such as Sub-Saharan Africa, Central Asia, and the Middle East to ensure a more balanced and equitable understanding of global LSP curriculum trajectories.

Conclusion

This systematic review aimed to examine how Arabic LSP curricula can be designed to meet institutional priorities, learner diversity, and evolving educational demands, particularly in scientific and medical contexts. The review responds to curricular fragmentation by emphasising the need for coherent, context-aware design grounded in cultural, disciplinary, and technological relevance.

Insights emerging from this review provide a nuanced understanding of how LSP curricula can be restructured to meet evolving institutional, and learner demands. Rather than prescribing one-size-fits-all models, the review underscores the necessity of integrating cultural, disciplinary, and technological considerations into curriculum planning. These conclusions highlight the importance of context-aware approaches that align pedagogical practice with the evolving demands of learners and institutions.

To translate these insights into action, curriculum developers should begin with clear stakeholder mapping that includes not only educators and content experts but also school leaders, ICT teams and policy advisors. Interdisciplinary collaboration must be guided by needs assessments and structured implementation plans. Policymakers must prioritise targeted funding for teacher training programmes that integrate TPACK principles and ensure that professional development initiatives are supported by school leadership to improve uptake and sustainability. Investment in equitable technology infrastructure should also be aligned with pedagogical goals and include provisions for underserved communities.

This review also paves the way for future research. Longitudinal investigations are needed to examine the sustained effects of AI integration, multilingual practices, and socio-emotional skill training within LSP curricula. Researchers should also expand their scope to include underrepresented geographical regions and non-English language scholarship. Such expansion would enhance the global LSP discourse through more balanced, inclusive, and evidence-based curriculum innovation.

Funding: Faculty of Education, The National University of Malaysia.

Acknowledgments: This study was supported by the Faculty of Education, The National University of Malaysia.

Conflicts of Interest: The authors declare no conflict of interest.

Authenticity: This manuscript is an original work.

Artificial Intelligence Statement: ChatGPT (OpenAI) was used only for grammar and language editing during manuscript preparation. All AI-assisted outputs were reviewed and revised by the authors, who take full responsibility for the accuracy and integrity of the content.

About the Authors:
Mior Syazril Mohamed Sapawi is a PhD candidate in Curriculum Development at The National University of Malaysia. His research focuses on Arabic Language Curriculum for Specific Purposes (ASP) in science secondary schools. He has worked in curriculum development roles at Malaysia's Ministry of Education. Email: miorsyazril@gmail.com
ORCID ID: https://orcid.org/0009-0007-4711-0636

Nik Mohd Rahimi Nik Yusoff is a Professor of Arabic Language Education at The National University of Malaysia. His scholarship covers Arabic pedagogy, curriculum design and policy; he has authored widely and led national textbook and module projects. Current interests include technology-enhanced Arabic teaching. Email: nikmrahimi@gmail.com
ORCID ID: https://orcid.org/0000-0003-0605-4010

References

Abouzahra, A., Sabraoui, A., & Afdel, K. (2020). Model composition in Model Driven Engineering: A systematic literature review. *Information and Software Technology*, *125*(May), 106316. https://doi.org/10.1016/j.infsof.2020.106316

Adhiambo, O. E., Sharon, O., & Alfred, O. (2025). Leveraging symbolic systems to foster expressive communication in students with hearing impairments. *Social Sciences and Humanities Open*, *11*(101376). https://doi.org/10.1016/j.ssaho.2025.101376

Arslan, S., & Curle, S. (2021). Sustainable development goals in the English language high school curriculum in Turkey. *European Journal of Education*, *56*(4), 681–695. https://doi.org/10.1111/ejed.12473

Asad, M. M., Shahzad, S., Shah, S. H. A., Sherwani, F., & Almusharraf, N. M. (2024). ChatGPT as artificial intelligence-based generative multimedia for English writing pedagogy: challenges and opportunities from an educator's perspective. In *International Journal of Information and Learning Technology* (Vol. 41, Issue 5, pp. 490–506).

Balderas, I. R., Moore, P., de Lourdes Gutiérrez Aceves, M., Irigoyen, E. N. P., & Vara, R. C. (2024). Curriculum design and discourses of Ivanic in literacy development in five English language teaching degrees. *Mextesol Journal*, *48*(1). https://www.scopus.com/inward/record.uri?eid=2-s2.0-85193355974&partnerID=40&md5=75b3a3c911ac5c8f9a47969683d33358

Bardi, M. (2013). Developing public managers' English language communication skills - Proposal for a textbook design and evaluation model. *Administratie Si Management Public*, *20*, 6–24. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84884372650&partnerID=40&md5=4cf62b155f11f68a872edb1bfeafbc79

Bessadet, L. (2022). Drama-based approach in English language teaching. *Arab World English Journal*, *13*(1), 525–533. https://doi.org/10.24093/awej/vol13no1.34

Burston, J., Athanasiou, A., & Yiokari-Neophytou, M. (2017). Achieving independent language learning through the mobilization of ubiquitous instructional technology resources. *Languages*, *2*(3), 1–14. https://doi.org/10.3390/languages2030016

Cao, Y., Jeyaraj, J. J., & Razali, A. B. (2024). Attitudes and perceptions of English as a foreign language students on blended learning and learner autonomy in a private university of China. *International Journal of Learning, Teaching and Educational Research*, *23*(1), 549–571. https://doi.org/10.26803/ijlter.23.1.26

Carreira, M., & Kagan, O. (2011). The results of the national heritage language survey: Implications for teaching, curriculum design, and professional development. *Foreign Language Annals*, *44*(1), 40–64. https://doi.org/10.1111/j.1944-9720.2010.01118.x

Cerezo Herrero, E., Schmidhofer, A., & Koletnik, M. (2021). An LSP framework for translation and interpreting pedagogy. *Revista de Lenguas Para Fines Específicos*, *27*(2), 140–156. https://doi.org/10.20420/rlfe.2021.441

Cha, S. E., Hewlin, P. F., Roberts, L. M., Buckman, B. R., Leroy, H., Steckler, E. L., Ostermeier, K., & Cooper, D. (2019). Being your true self at work: Integrating the fragmented research on authenticity in organizations. *Academy of Management Annals*, *13*(2), 633–671. https://doi.org/10.5465/annals.2016.0108

Chovancová, B. (2024). Training legal skills in the ESP classroom: Mediation activities and student self-reflection. *Hermes (Denmark)*, *64*, 137–148. https://doi.org/10.7146/hjlcb.vi64.147478

Chung, C.-H., & Chen, L.-J. (2021). Text mining for human resources competencies: Taiwan example. In *European Journal of Training and Development* (Vol. 45, pp. 588–602).

Cobb, W., & Bower, V. (2021). Languages across the curriculum. In *Language Learning and Intercultural Understanding in the Primary School* (pp. 99–118). https://doi.org/10.4324/9781003129738-7

Czeczulin, A., & Greydak, C. (2020). Language for a specific purpose: Business Russian. *Global Business Languages*, *20*(4), 55–63. https://doi.org/10.4079/gbl.v20.4

Diller, F. J., & Fraiser, D. M. (2022). Dear new reader: Factors affecting interest in reading prai. *Journal of Mekong Societies*, *18*(2), 125–144. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85137237667&partnerID=40&md5=0dd07348c0a6b236df8600c63cad1c02

Dursun, A. (2023). Domain analysis as a multidimensional research framework: Evidence-based alignment for LSP research, assessment, and curricula. *Global Business Languages*, *23*(1), 1–13. https://doi.org/10.4079/gbl.v23.2

Enríquez, T., Alonso-Stuyck, P., & Martínez-Villaseñor, L. (2023). The language of nature and artificial intelligence in patient care. *International Journal of Environmental Research and Public Health*, *20*(15). https://doi.org/10.3390/ijerph20156499

Finch, H. (2023). Dialogic co-creation in English language teaching and learning: a personal experience. *Language Learning in Higher Education*, *13*(1), 341–348. https://doi.org/10.1515/cercles-2023-2010

Fitria, T. N. (2020). Teaching English for specific purposes (ESP) to the students in English language teaching (ELT). *JET Adi Buana*, *5*(1), 55–66. https://doi.org/10.36456/jet.v5.n01.2020.2276

Freeman, D. (2017). The case for teachers' classroom English proficiency. *RELC Journal*, *48*(1), 31–52. https://doi.org/10.1177/0033688217691073

Generoso, J. C., & Arbon, A. M. M. (2020). Language needs analysis: An EAP curriculum design to develop foreign students' English skills. *Journal of Asia TEFL*, *17*(2), 428–445. https://doi.org/10.18823/asiatefl.2020.17.2.8.428

Granić, A. (2024). Technology adoption at individual level: Toward an integrated overview. *Universal Access in the Information Society*, *23*(2), 843–858. https://doi.org/10.1007/s10209-023-00974-3

Haque, Md. M. (2019). From cognition, metacognition to autonomy: A framework for understanding language learning dynamics. *Arab World English Journal*, *1*, 207–222. https://doi.org/10.24093/awej/efl1.15

Hegrenæs, C. F., Roald, J., Sandvei, B., & Simonnæs, I. (2022). Teaching specialized translation: Curriculum design of an online master course in legal translation. *Current Trends in Translation Teaching and Learning E*, *9*, 258–289. https://doi.org/10.51287/cttl20228 WE - Emerging Sources Citation Index (ESCI)

Huynh, T. N., & Nguyen, U. N. T. (2021). In-service language teachers' attitudes towards technology use and the development of their perceived call competencies. *Taiwan Journal of TESOL*, *18*(2), 29–62. https://doi.org/10.30397/TJTESOL.202110_18(2).0002 WE - Emerging Sources Citation Index (ESCI)

Jurkovič, V., Dostal, M., Podgoršek, S., & Mertelj, D. (2023). LSP teacher professional development-needs-based curriculum design. In *Linguistic and Extralinguistic in Interaction* (pp. 207–222).

Katsaounis, N., & Steinmüller, U. (2021). Languages for specific purposes in medicine and healthcare in times of the COVID-19 pandemic: Reflections on usage-based teaching. *Training, Language and Culture*, *5*(3), 29–40. https://doi.org/10.22363/2521-442X-2021-5-3-29-40

Kazakova, O. (2015). Language for specific purposes: Methodological problems, trends and perspectives. *Procedia - Social and Behavioral Sciences*, *214*(2015), 977–982. https://doi.org/10.1016/j.sbspro.2015.11.687

Kelly, R., & Hou, H. (2022). Empowering learners of English as an additional language: Translanguaging with machine translation. *Language and Education*, *36*(6), 544–559. https://doi.org/10.1080/09500782.2021.1958834

Kigoshi, T., Nakata, S., Abe, S., & Mochizuki, H. (2003). Design and development of multilingual e-learning materials TUFS language modules - pronunciation. *Proceedings of the IASTED International Conference on Computers and Advanced Technology in Education*, 591–596. https://www.scopus.com/inward/record.uri?eid=2-s2.0-1542726605&partnerID=40&md5=76fed89160296cafd21a8ed820b362c4

Koglbauer, R. (2022). Ofsted's curriculum research review for languages - what does this mean for language teachers, department leaders, curriculum design and professional development? *Language Learning Journal*, *50*(2), 262–267. https://doi.org/10.1080/09571736.2022.2045681

Kurniawan, A. B., & Sumani. (2022). Communicative language teaching approach potential for English language teaching. *KnE Social Sciences*, 37–44. https://doi.org/10.18502/kss.v7i7.10648

Lasekan, O. A., Moraga-Pumarino, A. F., & Pachava, V. (2023). Using needs analysis to foster sustainability of business English courses: A case study of a university in the south of Chile. *Sustainability (Switzerland)*, *15*(22). https://doi.org/10.3390/su152216074

Lei, M., & Medwell, J. (2020). How do English language teachers understand the idea of professional development in the recent curriculum reforms in China? *Asia Pacific Journal of Education*, *40*(3), 401–417. https://doi.org/10.1080/02188791.2020.1717440

Medvedeva, O. D., Rubtsova, A. V, Vilkova, A. V, & Ischenko, V. V. (2022). Digital monitoring of students' soft skills development as an interactive method of foreign language learning. *Education Sciences*, *12*(8). https://doi.org/10.3390/educsci12080506

Meijer, A., & Bekkers, V. (2015). A metatheory of e-government: Creating some order in a fragmented research field. *Government Information Quarterly*, *32*(3), 237–245. https://doi.org/10.1016/j.giq.2015.04.006

Meniado, J. C. (2023). Digital language teaching 5.0: Technologies, trends and competencies. *RELC Journal*, *54*(2), 461–473. https://doi.org/10.1177/00336882231160610

Morell, T. (2020). EMI teacher training with a multimodal and interactive approach: A new horizon for LSP specialists. *Language Value*, *12*(1), 56–87. https://doi.org/10.6035/LanguageV.2020.12.4

Moscovici, K., Ben Arie, G., Shelef, I., & Kalet, A. (2022). Creating Wikipedia articles on health and technology topics can empower the writers and benefit the community. *BMC Medical Education*, *22*(1). https://doi.org/10.1186/s12909-022-03389-5 WE - Science Citation Index Expanded (SCI-EXPANDED) WE - Social Science Citation Index (SSCI)

Moule, P. (2007). Challenging the five-stage model for e-learning: A new approach. *ALT-J, Research in Learning Technology*, *15*(1), 37–50. https://doi.org/10.1080/09687760601129588

Namanyane, T., & Shaoan, M. M. R. (2025). English teachers' pedagogical content knowledge in the implementation of learner-centered role-play: A case study of curriculum and policy reforms in China. *Cogent Education*, *12*(1). https://doi.org/10.1080/2331186X.2025.2488073 WE - Emerging Sources Citation Index (ESCI)

Oravcová, V., Yakovenko, K., & Mišík, M. (2022). Complete but fragmented: Research on energy in Central and eastern Europe. *Energies*, *15*(17), 1–16. https://doi.org/10.3390/en15176185

Page, M. J., McKenzie, J. E., Bossuyt, P., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Medicina Fluminensis*, *57*(4), 444–465. https://doi.org/10.21860/medflum2021_264903

Perlado Lomo de Espinosa, I., & Trujillo Vargas, J. J. (2024). Research analysis on social communicative skills as a tool to prevent violence in the educational context. *Revista Latina de Comunicacion Social*, *2024*(82). https://doi.org/10.4185/RLCS-2024-2302

Piri, S., & Haxhillazi, A. (2023). Medical Italian language: Application of terminology and specific didactic materials in the teaching process. *XLinguae*, *16*(2), 129–139. https://doi.org/10.18355/XL.2023.16.02.10

Porto, M., López-Barrios, M., & Banegas, D. L. (2021). Research on English language teaching and learning in Argentina (2014-2018). *Language Teaching*, *54*(3), 355–387. https://doi.org/10.1017/S0261444821000082 WE - Social Science Citation Index (SSCI) WE - Arts & Humanities Citation Index (A&HCI)

Purwanto, S., & Nurhamidah, I. (2021). Digitizing English for specific purposes in the era of COVID-19 pandemic. *PAROLE: Journal of Linguistics and Education*, *11*(1), 57–72. https://doi.org/10.14710/parole.v11i1.57-72

Reynolds, B. L., Yu, M. H., & Ha, X. Van. (2023). Administrative staff learning and using English to communicate with international students. *Innovation in Language Learning and Teaching*, *17*(2), 424–439. https://doi.org/10.1080/17501229.2022.2058513

Rogers, E. M., Singhal, A., & Quinlan, M. M. (2019). Diffusion of innovations. In *An Integrated Approach to Communication Theory and Research, Third Edition* (pp. 415–433). https://doi.org/10.4324/9780203710753-35

Rooyen, I. van, Claassen, J., Moodaley, N., Doyle, G., Skade, T., Nash, R., Gxilishe, S., & Hellenberg, D. A. (2024). Multilingual education: Medical interns perceptions regarding the usefulness of non-mother tongue communications skills taught during the undergraduate curriculum. *BMC Medical Education*, *24*(1). https://doi.org/10.1186/s12909-024-05414-1

Rus, D. (2020). Creative methodologies in teaching English for engineering students. *Procedia Manufacturing*, 337–343. https://doi.org/10.1016/j.promfg.2020.03.049

Sanosi, A. B., Nur, H. S., Arafat, Y., & Muhammad, D. (2025). Developing an ESP syllabus to promote sustainable development goals: A Delphi study. *World Journal of English Language*, *15*(2), 381–393. https://doi.org/10.5430/wjel.v15n2p381

Savula, A., Mylyk, O., Skiba, N., Podoliak, M., & Dzyubynska, K. (2024). Effective methods of teaching a foreign language in the modern educational environment. *Scientific Herald of Uzhhorod University. Series Physics*, *55*, 2588–2597. https://doi.org/10.54919/physics/55.2024.258qy8

Shen, Y., & Lu, W. (2024). Technology of animation in multimedia courseware of mechanical drawing. *International Journal of Electrical Engineering and Education*, *61*(4), 438–451. https://doi.org/10.1177/0020720920986085

Sierra-Piedrahita, A. M., & Echeverri-Sucerquia, P. A. (2020). Governmental professional development initiatives for the implementation of language policies and curriculum guidelines: Secondary school teachers' experiences, challenges and views. *Ikala*, *25*(1), 137–152. https://doi.org/10.17533/udea.ikala.v25n01a13

Syamsi, K., Zuchdi, D., Kusmiatun, A., Purbani, W., & Harun, A. (2024). Developing a culture-based Indonesian language for academic purposes textbook for non-native speakers. *Cakrawala Pendidikan*, *43*(1), 115–126. https://doi.org/10.21831/cp.v43i1.60321

Tian, Y., & Wang, Y. (2024). Understanding instructors' tablet adoption for note-taking in interpreting: Insights from the GETAMEL model. *Frontiers in Education*, *9*. https://doi.org/10.3389/feduc.2024.1456770

Tirizite, Y., & Marjane, D. (2023). Use of differentiated instruction in LSP courses in Moroccan higher education. *EIKI Journal of Effective Teaching Methods*, *1*(3), 1–8. https://doi.org/10.59652/jetm.v1i3.30

Tyron, O., Korieshkova, S., Didenko, M., & Kaminska, S. (2021). Psycho-pedagogical approaches to the design of English learning course for ship engineers. *International Journal of Engineering Pedagogy*, *10*(6), 21–35. https://doi.org/10.3991/IJEP.V10I6.14107

Urushibara-Miyachi, Y., Matsushita, K., & Nishigori, H. (2024). Development of a competency-based national model core curriculum in Japan: A case study. *Medical Teacher*, *46*(sup1), S11–S20. https://doi.org/10.1080/0142159X.2024.2388906

Valle Coto, M. G., Iñiguez, R. X., Lentskevich, M. A., Ally, S. A., Farfan, J. F., Park, Y. S., Gangopadhyaya, A. G., & Ortega, P. (2024). Teaching foundational language equity concepts in the pre-clinical curriculum. *BMC Medical Education*, *24*(1). https://doi.org/10.1186/s12909-024-05474-3

Wang, F. F., Zhou, X. H., Li, K. X., Cheung, A. C. K., & Tian, M. (2025). The effects of artificial intelligence-based interactive scaffolding on secondary students' speaking performance, goal setting, self-evaluation, and motivation in informal digital learning of

English. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2025.2470319

Wang, L. G. (2014). Research on design & development for multimedia curriculum resources in foreign language and literature. *Advanced Materials Research, 971–973*, 2524–2527. https://doi.org/10.4028/www.scientific.net/AMR.971-973.2524

Zare-ee, A., & Hejazi, S. Y. (2018). Iranian university learners' and teachers' views on adopted and locally-developed English language teaching textbooks. *International Journal of Instruction, 11*(3), 291–308. https://doi.org/10.12973/iji.2018.11321a

Zhou, S. H., Fung, D., & Thomas, N. (2023). Towards deeper learning in EMI lectures: The role of English proficiency and motivation in students' deep processing of content knowledge. *Journal of Multilingual and Multicultural Development*. https://doi.org/10.1080/01434632.2023.2248078 WE - Social Science Citation Index (SSCI) WE - Arts & Humanities Citation Index (A&HCI)

Zucker, T. A., Carlo, M. S., Landry, S. H., Masood-Saleem, S. S., Williams, J. M., & Bhavsar, V. (2019). Iterative design and pilot testing of the Developing Talkers tiered academic language curriculum for pre-kindergarten and kindergarten. *Journal of Research on Educational Effectiveness*, *12*(2), 274–306. https://doi.org/10.1080/19345747.2018.1519623

APPENDIX A. Number and Details of Primary Studies Database

Table A1. ***Number and details of primary studies database***

No.	AUTHOR	TITLE	YEAR	JOURNAL	SCOPUS	WoS	ERIC
1.	Chovancová B.	Training Legal Skills in the ESP Classroom: Mediation Activities and Student Self-Reflection	2024	Hermes (Denmark)	/		
2.	Valle Coto M.G.; Iñiguez R.X.; Lentskevich M.A.; Ally S.A.; Farfan J.F.; Park Y.S.; Gangopadhyaya A.G.; Ortega P.	Teaching Foundational Language Equity Concepts in The Pre-Clinical Curriculum	2024	BMC Medical Education	/	/	
3.	Urushibara-Miyachi Y.; Matsushita K.; Nishigori H.	Development of a Competency-Based National Model Core Curriculum in Japan: A Case Study	2024	Medical Teacher	/		
4.	Van Rooyen I.; Claassen J.; Moodaley N.; Doyle G.; Skade T.; Nash R.; Gxilishe S.; Hellenberg D.A.	Multilingual Education: Medical Interns Perceptions Regarding the Usefulness of Non-Mother Tongue Communications Skills Taught During the Undergraduate Curriculum	2024	BMC Medical Education	/	/	
5.	Tyron O.; Korieshkova S.; Didenko M.; Kaminska S.	Psycho-Pedagogical Approaches to the Design of English Learning Course for Ship Engineers	2021	International Journal of Engineering Pedagogy	/		
6.	Perlado Lomo de Espinosa I.; Trujillo Vargas J.J.	Research Analysis on Social Communicative Skills as a Tool to Prevent Violence in the Educational Context	2024	Revista Latina de Comunicacion Social	/		
7.	Medvedeva O.D.; Rubtsova A.V.; Vilkova A.V.; Ischenko V.V.	Digital Monitoring of Students' Soft Skills Development as an Interactive Method of Foreign Language Learning	2022	Education Sciences	/	/	
8.	Savula A.; Mylyk O.; Skiba N.; Podoliak M.; Dzyubynska K.	Effective Methods of Teaching a Foreign Language in the Modern Educational Environment	2024	Scientific Herald of Uzhhorod University. Series Physics	/		
9.	Lasekan O.A.; Moraga-Pumarino A.F.; Pachava V.	Using Needs Analysis to Foster Sustainability of Business English Courses: A Case Study of a University in the South of Chile	2023	Sustainability (Switzerland)	/		
10.	Diller F.J.; Fraiser D.M.	Dear New Reader: Factors Affecting Interest in Reading Prai	2022	Journal of Mekong Societies	/		

No.	AUTHOR	TITLE	YEAR	JOURNAL	SCOPUS	WoS	ERIC
11.	Meniado J.C.	Digital Language Teaching 5.0: Technologies, Trends and Competencies	2023	RELC Journal	/	/	/
12.	Cao Y.; Jeyaraj J.J.; Razali A.B.	Attitudes and Perceptions of English as a Foreign Language Students on Blended Learning and Learner Autonomy in a Private University of China	2024	International Journal of Learning, Teaching and Educational Research	/		
13.	Tian Y.; Wang Y.	Understanding Instructors' Tablet Adoption for Note-Taking in Interpreting: Insights from the GETAMEL Model	2024	Frontiers in Education	/	/	
14.	Shen Y.; Lu W.	Technology of Animation in Multimedia Courseware of Mechanical Drawing	2024	International Journal of Electrical Engineering and Education	/		
15.	Adhiambo O.E.; Sharon O.; Alfred O.	Leveraging Symbolic Systems to Foster Expressive Communication in Students with Hearing Impairments	2025	Social Sciences and Humanities Open	/		
16.	Moscovici, Khen; Ben Arie, Gal; Shelef, Ilan; Kalet, Adina	Creating Wikipedia Articles on Health and Technology Topics Can Empower the Writers and Benefit the Community	2022	BMC Medical Education		/	
17.	Zhou, Sihan; Fung, Daniel; Thomas, Nathan	Towards Deeper Learning in EMI Lectures: The Role of English Proficiency and Motivation in Students' Deep Processing of Content Knowledge	2023	Journal of Multilingual and Multicultural Development		/	
18.	Hegrenaes, Claudia Forster; Roald, Jan; Sandvei, Beate; Simonnaes, Ingrid	Teaching Specialized Translation: Curriculum Design of An Online Master Course in Legal Translation	2022	Current Trends in Translation Teaching and Learning E		/	
19.	Thu-Nguyet Huynh; Uyen Nu Thuy Nguyen	In-Service Language Teachers' Attitudes Towards Technology Use and the Development of Their Perceived Call Competencies	2021	Taiwan Journal of Tesol		/	
20.	Porto, Melina; Lopez-Barrios, Mario; Banegas, Dario Luis	Research on English Language Teaching and Learning in Argentina (2014-2018)	2021	Language Teaching		/	

No.	AUTHOR	TITLE	YEAR	JOURNAL	SCOPUS	WoS	ERIC
21.	Namanyane, Tebatso; Shaoan, Md Mirajur Rhaman	English Teachers' Pedagogical Content Knowledge in the Implementation of Learner-Centered Role-Play: A Case Study of Curriculum and Policy Reforms in China	2025	Cogent Education		/	
22.	Wang, Feifei; Zhou, Xiaohua; Li, Kangxin; Cheung, Alan C. K.; Tian, Ming	The Effects of Artificial Intelligence-Based Interactive Scaffolding on Secondary Students' Speaking Performance, Goal Setting, Self-Evaluation, And Motivation in Informal Digital Learning of English	2025	Interactive Learning Environments		/	
23.	Chung, C.-H., & Chen, Lu-Jia	Text Mining for Human Resources Competencies: Taiwan Example	2021	European Journal of Training and Development			/
24.	Asad, M. M., Shahzad, S., Shah, S. H. A., Sherwani, F., & Almusharraf, N. M.	ChatGPT as Artificial Intelligence-Based Generative Multimedia for English Writing Pedagogy: Challenges and Opportunities from an Educator's Perspective	2024	International Journal of Information and Learning Technology			/

APPENDIX B. Quality Assessment of Selected Studies Based on Systematic Review Criteria (QA1–QA6)

Table B1. ***Quality assessment of selected studies based on systematic review criteria (QA1–QA6)***

PRIMARY STUDY	TITLE	QA1	QA2	QA3	QA4	QA5	QA6	TOTAL MARK	PERCENTAGE (%)
PS1	Training Legal Skills in the ESP Classroom: Mediation Activities and Student Self-Reflection	1	1	1	1	0.5	0.5	5	83.33
PS2	Teaching Foundational Language Equity Concepts in the Pre-Clinical Curriculum	1	1	1	1	0.5	0.5	5	83.33
PS3	Development of a Competency-Based National Model Core Curriculum in Japan: A Case Study	1	1	1	1	1	1	6	100.00
PS4	Multilingual Education: Medical Interns Perceptions Regarding the Usefulness of Non-Mother Tongue Communications Skills Taught During the Undergraduate Curriculum	1	1	1	1	1	1	6	100.00
PS5	Psycho-Pedagogical Approaches to the Design of English Learning Course for Ship Engineers	1	1	1	1	1	0.5	5.5	91.67
PS6	Research Analysis on Social Communicative Skills as a Tool to Prevent Violence in the Educational Context	1	1	1	1	1	0.5	5.5	91.67
PS7	Digital Monitoring of Students' Soft Skills Development as an Interactive Method of Foreign Language Learning	1	1	1	1	0.5	0.5	5	83.33

PRIMARY STUDY	TITLE	QA1	QA2	QA3	QA4	QA5	QA6	TOTAL MARK	PERCENTAGE (%)
PS8	Effective Methods of Teaching a Foreign Language in The Modern Educational Environment	1	1	1	1	1	1	6	100.00
PS9	Using Needs Analysis to Foster Sustainability of Business English Courses: A Case Study of a University in the South of Chile	1	1	1	1	0.5	0.5	5	83.33
PS10	Dear New Reader: Factors Affecting Interest in Reading Prai	1	1	1	1	1	0.5	5.5	91.67
PS11	Digital Language Teaching 5.0: Technologies, Trends and Competencies	1	1	1	1	1	1	6	100.00
PS12	Attitudes and Perceptions of English as a Foreign Language Students on Blended Learning and Learner Autonomy in a Private University of China	1	1	1	1	0.5	0.5	5	83.33
PS13	Understanding Instructors' Tablet Adoption for Note-Taking in Interpreting: Insights from The GETAMEL Model	1	1	1	1	0.5	1	5.5	91.67
PS14	Technology of Animation in Multimedia Courseware of Mechanical Drawing	1	1	1	1	1	0.5	5.5	91.67
PS15	Leveraging Symbolic Systems to Foster Expressive Communication in Students with Hearing Impairments	1	1	1	1	0.5	0.5	5	83.33
PS16	Creating Wikipedia Articles on Health and Technology Topics Can Empower the Writers and Benefit the Community	1	1	1	1	1	0.5	5.5	91.67
PS17	Towards Deeper Learning in EMI Lectures: The Role of English Proficiency and Motivation in Students' Deep Processing of Content Knowledge	1	1	1	1	1	1	6	100.00
PS18	Teaching Specialized Translation: Curriculum Design of an Online Master Course in Legal Translation	1	1	1	1	0.5	0.5	5	83.33
PS19	In-Service Language Teachers' Attitudes Towards Technology Use and the Development of Their Perceived Call Competencies	1	1	1	1	1	0.5	5.5	91.67
PS20	Research on English Language Teaching and Learning in Argentina (2014-2018)	1	1	1	1	1	0.5	5.5	91.67
PS21	English Teachers' Pedagogical Content Knowledge in the Implementation of Learner-Centered Role-Play: A Case Study of Curriculum and Policy Reforms in China	1	1	1	1	0.5	1	5.5	91.67
PS22	The Effects of Artificial Intelligence-Based Interactive Scaffolding on Secondary Students' Speaking Performance, Goal Setting, Self-Evaluation, and Motivation in Informal Digital Learning of English	1	1	1	1	1	0.5	5.5	91.67
PS23	Text Mining for Human Resources Competencies: Taiwan Example	1	1	1	1	0.5	0.5	5	83.33

PRIMARY STUDY	TITLE	QA1	QA2	QA3	QA4	QA5	QA6	TOTAL MARK	PERCENTAGE (%)
PS24	ChatGPT as Artificial Intelligence-Based Generative Multimedia for English Writing Pedagogy: Challenges and Opportunities from an Educator's Perspective	1	1	1	1	1	0.5	5.5	91.67

APPENDIX C. Thematic Classification of Studies in LSP Curriculum Development

Table C1. ***Thematic classification of studies in LSP curriculum development***

THEME	PS CODE	TITLE
Curriculum Models and LSP Frameworks	PS2	Teaching Foundational Language Equity Concepts in the Pre-Clinical Curriculum (Valle Coto et al., 2024)
	PS3	Development of a Competency-Based National Model Core Curriculum in Japan: A Case Study (Urushibara-Miyachi et al., 2024)
	PS4	Multilingual Education: Medical Interns Perceptions Regarding the Usefulness of Non-Mother Tongue Communications Skills Taught During the Undergraduate Curriculum (Rooyen et al., 2024)
	PS5	Psycho-Pedagogical Approaches to the Design of English Learning Course for Ship Engineers (Tyron et al., 2021)
	PS13	Understanding Instructors' Tablet Adoption for Note-Taking in Interpreting: Insights from the GETAMEL Model (Tian & Wang, 2024)
	PS18	Teaching Specialized Translation: Curriculum Design of an Online Master Course in Legal Translation (Hegrenæs et al., 2022)
	PS21	English Teachers' Pedagogical Content Knowledge in Role-Play Implementation of Learner-Centered Role-Play: A Case Study of Curriculum and Policy Reforms in China (Namanyane & Shaoan, 2025)
Teaching Practices and Pedagogical Strategies in LSP	PS1	Training Legal Skills in the ESP Classroom: Mediation Activities and Student Self-Reflection (Chovancová, 2024)
	PS7	Digital Monitoring of Students' Soft Skills Development as an Interactive Method of Foreign Language Learning (Medvedeva et al., 2022)
	PS8	Effective methods of teaching a foreign language in the modern educational environment (Savula et al., 2024)
	PS11	Digital Language Teaching 5.0: Technologies, Trends and Competencies (Meniado, 2023)
	PS12	Attitudes and Perceptions of English as a Foreign Language Students on Blended Learning and Learner Autonomy in a Private University of China (Cao et al., 2024)
	PS17	Towards Deeper Learning in EMI Lectures: The Role of English Proficiency and Motivation in Students' Deep Processing of Content Knowledge (Zhou et al., 2023)
	PS19	In-Service Language Teachers' Attitudes Towards Technology Use and the Development of Their Perceived Call Competencies (Huynh & Nguyen, 2021)
	PS20	Research on English Language Teaching and Learning in Argentina (2014-2018) (Porto et al., 2021)
	PS22	The Effects of Artificial Intelligence-Based Interactive Scaffolding on Secondary Students' Speaking Performance, Goal Setting, Self-Evaluation, and Motivation in Informal Digital Learning of English (F. F. Wang et al., 2025)

THEME	PS CODE	TITLE
	PS24	ChatGPT as Artificial Intelligence-Based Generative Multimedia for English Writing Pedagogy: Challenges and Opportunities from an Educator's Perspective (Asad et al., 2024)
Technological and Contextual Adaptation in LSP Curriculum	PS6	Research Analysis on Social Communicative Skills to Prevent Violence in the Educational Context (Perlado Lomo de Espinosa & Trujillo Vargas, 2024)
	PS9	Using Needs Analysis to Foster Sustainability of Business English Courses: A Case Study of a University in the South of Chile (Lasekan et al., 2023)
	PS10	Dear New Reader: Factors Affecting Interest in Reading Prai (Diller & Fraiser, 2022)
	PS14	Technology of Animation in Multimedia Courseware of Mechanical Drawing (Shen & Lu, 2024)
	PS15	Leveraging Symbolic Systems to Foster Expressive Communication in Students with Hearing Impairments (Adhiambo et al., 2025)
	PS16	Creating Wikipedia Articles on Health and Technology Topics Can Empower the Writers and Benefit the Community (Moscovici et al., 2022)
	PS23	Text Mining for Human Resources Competencies: Taiwan Example (Chung & Chen, 2021)

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.211-226
DOI: https://dx.doi.org/10.24093/awej/vol16no3.12

From Fragmentation to Coherence: A Knowledge Graph-Driven Integration Framework for Blended Learning

Lingling Lou
School of English Studies, Zhejiang Yuexiu University, Shaoxing, China
Email: 237341893@qq.com

Received:06/23/2025 Accepted: 08/03/2025 Published: 09/20/2025

Abstract
Blended learning in higher education often struggles with disconnected resources across various platforms, causing cognitive fragmentation for students and heavy workloads for instructors. This study addresses these issues by proposing a knowledge graph-driven integration framework, with its core being a model that organizes key course concepts into a hierarchical knowledge graph and connects them to resources from online courses, digital textbooks, and virtual simulations. The model resolves long-standing resource integration inefficiencies, boosting teaching efficiency and learner engagement. A quasi-experimental study with 180 undergraduates (evenly split into experimental and control groups) showed that the experimental group using the model achieved higher final scores (82.1 vs. 76.2, $p<0.001$), stronger perceived resource coherence (4.0/5 vs. 3.2/5, $p<0.001$), and reduced cognitive load (4.9/10 vs. 6.1/10, $p<0.001$). The instructor also saw a 37.8% weekly workload reduction, particularly in grading and resource coordination. These findings demonstrate that the knowledge graph-based model enhances student learning outcomes and eases instructor burdens by improving resource coherence. Universities using multiple digital platforms are recommended to adopt this approach, especially for courses requiring strong concept linkage. Future research will test the model in other subjects and over longer periods to verify its scalability.
Keywords: blended learning, cognitive load reduction, educational resource integration, instructional efficiency, knowledge graph-driven integration, learning outcomes

How to cite: Lou, L. (2025). From Fragmentation to Coherence: A Knowledge Graph-Driven Integration Framework for Blended Learning. *Arab World English Journal, 16*(3): 211-226. https://dx.doi.org/10.24093/awej/vol16no3.12

Introduction

The rapid growth of digital learning tools—from MOOCs and e-textbooks to virtual simulations—has transformed education by providing unprecedented access to diverse resources (Alharbi & Basiouni, 2025). However, this expansion has created a significant unintended consequence: fragmented resources. When course materials scatter across multiple unconnected platforms, "both students and teachers pay a heavy cognitive price" (Graham, 2013, p. 339). Like trying to assemble a puzzle with pieces stored in different rooms, this fragmentation makes it extraordinarily difficult to form a coherent understanding of concepts.

This problem is particularly severe for subjects requiring a deep understanding. A good example is linguistics: to understand vowel sounds, students must integrate knowledge about sound physics, articulatory mechanisms, and regional accent variations. However, in practice, these learning resources are usually scattered across various platforms. Students might access a lecture video on dialectology, a textbook chapter on articulatory phonetics, and a pronunciation simulation program as separate, unlinked resources (Wang & Li, 2021). Even on integrated platforms like Moodle or XuetangX, materials lack semantic connections, forcing students to waste cognitive effort on resource integration rather than on the actual learning content (Huang et al., 2021).

To address this fragmentation, the study aims to develop a Knowledge Graph-Driven Integration (KGDI) framework, with three core objectives: first, to semantically unify multimodal resources across platforms through a structured knowledge graph; second, to evaluate the framework's impact on student learning outcomes and cognitive experiences; and third, to assess its potential for reducing instructor workload in blended learning environments. Guiding the in-depth exploration of these objectives are the following three research questions:

Question 1: How does the KGDI model semantically integrate multimodal resources to create coherent learning experiences?

Question 2: To what extent does the KGDI model improve student learning outcomes compared to conventional methods?

Question 3: How significantly does the KGDI model reduce instructor workload compared to traditional platforms?

The significance of this study lies in its direct response to UNESCO's (2023) call for the more innovative integration of technology in education. Theoretically, it bridges the gap between cognitive science and educational practice, assessing whether structured knowledge organization enhances real-world teaching effectiveness. Practically, it offers institutions a scalable solution to balance digital pedagogy demands with sustainable teaching workflows.

Notably, existing educational knowledge graphs focus mainly on structuring domain concepts hierarchically (Ji et al., 2021). The KGDI model innovates by linking scattered cross-platform resources into cohesive sequences, addressing their disconnected utilization. This structured unification aligns with cognitive load theory, which posits that learning efficiency decreases when learners waste cognitive resources on integrating disjointed information (Sweller, 2010). By pre-establishing semantic connections between resources, the model reduces unnecessary cognitive load, allowing students to focus on deep learning rather than navigating resources.

To empirically test the model, three hypotheses have been formulated:

Hypothesis 1: KGDI model users will show significantly higher academic performance and learning satisfaction compared to those in traditional blended learning environments.

Hypothesis 2: KGDI model users will report significantly lower cognitive load and stronger perceived resource coherence than peers in traditional blended learning environments.

Hypothesis 3: The instructor will exhibit a significant reduction in time spent on routine tasks compared to standard practices.

By adopting a quasi-experimental design involving 180 undergraduates, the research implemented the KGDI model in a linguistics course to evaluate the model's effectiveness. Methodology and analysis details are provided in subsequent sections.

Literature Review

Blended learning, which combines in-person classroom teaching with online resources, has become a central approach in modern education. Research confirms that "effective blended learning makes education more flexible and improves student learning outcomes by combining the best parts of different teaching methods" (Aljuaid, 2024, p. 45). Landmark studies by Garrison and Kanuka (2004) emphasized its value in building active learning communities by combining real-time interactive activities and self-paced tasks. This dual-mode advantage is further supported by studies by Alsalihi (2025) and Mir et al. (2023), which found that well-designed blended courses consistently yield better learning outcomes than either fully online or traditional classes alone.

The benefits extend beyond academic performance. When courses successfully integrate theoretical knowledge with practical tasks, students demonstrate a stronger ability to apply what they learn in real-world situations (Tabassum & Mohd Saad, 2025). This is particularly evident in complex problem-solving scenarios, where well-sequenced blended learning activities show clear advantages (Tambunan et al., 2021).

The successful implementation of blended learning requires careful instructional design. Research globally highlights key success factors (Bernard et al., 2014). At the institutional level, administrative support, including investments in technology and systems, is critical (Abirami et al., 2023). This is particularly evident in China, where national "Internet Plus Education" policies have promoted platforms like XuetangX and Zhihuishu. At the teaching level, teacher training covering blended learning strategies is also of great importance (Ni & He, 2024).

However, this rapid expansion has revealed new challenges. Teachers report spending nearly one-third of their preparation time managing technical issues across different platforms rather than focusing on personalized instruction (Denbel, 2023), and recent research confirms that "increased workload" has already been a primary concern for educators in blended learning settings (Al Hajri et al., 2025, p. 35). Meanwhile, students increasingly complain about the mental fatigue caused by constantly switching between disconnected systems (Skulmowski & Xu, 2022).

Cognitive science now provides measurable evidence for this burden of fragmentation. Recent brain activity studies indicate that constant platform switching imposes a significant cognitive load, thereby reducing the mental resources available for actual learning (Shaya et al., 2025). Specifically, Wang et al. (2024) found that when students alternate between lecture videos, e-textbooks, and simulation tools, "their working memory capacity for conceptual integration decreases by approximately 30%" (p. 632). These findings objectively confirm the mental exhaustion reported by learners and help explain why resource fragmentation remains a critical barrier to blended learning's full potential.

While various technological solutions have been tried, most current approaches offer only superficial integration. Basic system linking methods fail to establish meaningful conceptual relationships or adapt to individual learning paths (Zawacki-Richter et al., 2019). In light of these

circumstances, knowledge graphs (KGs)—which visually map relationships between concepts and enable dynamic information linking—have emerged as a promising solution to address such integration challenges. Their core strength lies in structuring dispersed information into semantic networks, making them well-suited to connect fragmented resources across platforms.

Prior research has demonstrated the potential of KGs in solving complex integration problems across domains. In construction safety, for example, KGs have reduced on-site accidents through real-time integration of regulatory standards, equipment manuals, and training materials (Xiang et al., 2025). Similarly, in manufacturing, cross-platform KG integration has improved production efficiency by 31% (Su et al., 2025).

Yet in educational contexts, unlike other fields, KG research has not reached this potential. This aligns with the observation that "educational KG research remains fragmented" (Fan, 2024, p. 1256). Indeed, current research primarily focuses on separate applications, including organizing course content, tracking individual learning paths, or suggesting theoretical approaches (Fedorenko & Kravchenko, 2023). Besides, these investigations rely heavily on qualitative methods with limited experimental testing. Few studies have achieved full multi-resource integration or assessed the practical impact of knowledge graph systems on teaching and learning.

In summary, there are three key research gaps. First, existing educational KG approaches usually handle isolated aspects without integrating cross-platform resources. Second, methodologies favor small cases and theoretical discussions over controlled trials with sufficient participants. Third, practical implementation questions remain, particularly how to effectively connect diverse resources like videos, textbooks, and simulations.

To address these research gaps, this study introduces the Knowledge Graph-Driven Integration (KGDI) framework, which embeds educational theories into graph systems to enable automatic linking of cross-platform materials. Using a quasi-experimental approach, the research measures the combined impact of the model on learning outcomes and teaching processes, thereby strengthening both the theoretical and practical foundations for intelligent educational technology.

KGDI Model Construction

Building on the previously identified literature gaps, this section presents the KGDI model's architectural design and its operational teaching workflow.

Knowledge Graph-Driven Resource Integration Framework

The KGDI model is a systematic blended learning approach addressing resource fragmentation in digital education. As shown in Figure 1 (original creation for this study), this model establishes a unified learning ecosystem centered on a knowledge graph that integrates three core educational resources: the digital textbook, the MOOC platform, and the virtual simulation system. This integration creates a cohesive network where key course concepts serve as connection points, dynamically coordinating materials across different platforms. Automated tagging and linking are supported by large language models (LLMs) and neural network validation.

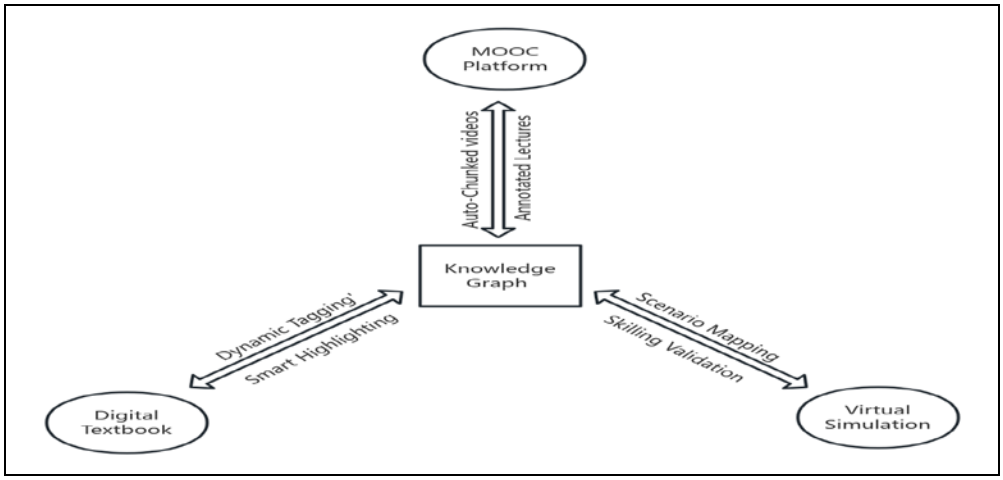


Figure 1. Knowledge graph-driven multimodal resource integration framework

Consider the *Introduction to Linguistics* course as an example. On the Zhihuishu MOOC platform, this course contains 1,282 learning resources, including 47 instructional videos. Each video is divided into several labeled sections that match knowledge graph concepts—tags generated automatically via the LLM-NER pipeline—enabling students to find relevant video segments when studying specific topics quickly. Besides videos, the platform's 366 quizzes and discussions are connected in the knowledge graph system, ensuring all materials work seamlessly together.

The digital textbook *Introduction to Linguistics* covers six chapters and contains approximately 60,000 words. Its content integrates 56 videos, 136 diagrams, and 118 interactive exercises. Each section features specialized tags that connect directly to the knowledge graph. For instance, when students study "phonemic variation", the system automatically displays related MOOC video clips showing dialect differences alongside matching pronunciation simulations. In this way, manual cross-platform searching is eliminated.

Adapted from the provincial multicultural communication project, the virtual simulation system is used for pragmatics instruction through three progressive stages: preparatory warm-ups, simulated cultural interactions, and applied practice tasks. Using 3D models and 360-degree scenarios, every simulation module is directly linked to knowledge graph concepts, supporting the chapter's real-world communication skill objectives.

The KGDI system integrates these three core teaching platforms through an intelligent coordination mechanism powered by the knowledge graph. Each resource—whether textbook content, MOOC videos, or virtual simulations—is tagged with concept tags and location identifiers. This architecture eliminates platform-switching burdens. When students explore a concept like "speech act theory", the knowledge graph automatically assembles relevant materials from all three platforms, presenting them in a logically sequenced workflow: theoretical explanations from the textbook first, followed by video demonstrations, and finally hands-on practice in virtual scenarios. The system preserves each resource's pedagogical strengths while ensuring seamless thematic continuity across platforms. By actively coordinating rather than simply combining resources, the system eliminates the need for platform switching and ensures that all components function as a unified instructional tool.

The Three-Phase Teaching Model Based on Knowledge Graph-Driven Integration

The KGDI model operates through three interconnected phases, as visually summarized in Figure 2 (original creation for this study): pre-class preparation, in-class instruction, and post-class consolidation.

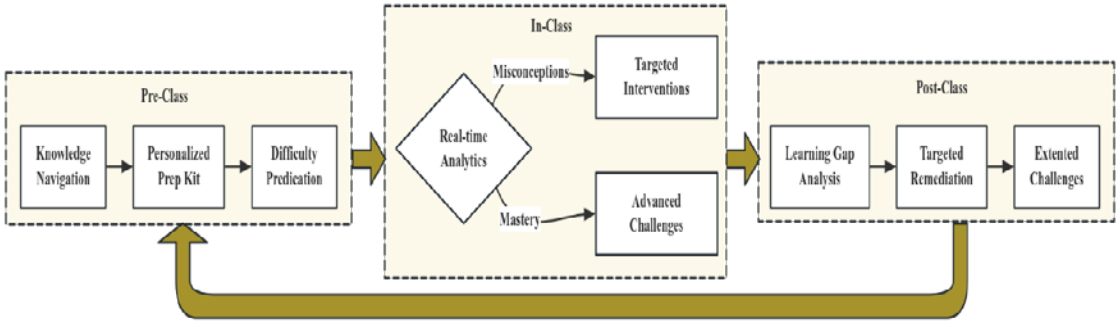


Figure 2. KGDI three-phase teaching model

During the pre-class phase, students interact with the knowledge graph interface, which displays core concepts and their relationships. The system recommends personalized resources—including specific MOOC video clips, relevant textbook sections, and preparatory exercises—based on individual learning patterns. This targeted preparation enables students to build foundational knowledge efficiently before face-to-face sessions.

During classroom instruction, the knowledge graph serves as both a teaching aid and a learning scaffold. Instructors utilize the visualized concept network to demonstrate relationships between ideas, identify common misconceptions through real-time data analysis, and facilitate more in-depth discussions. Students engage in active learning activities suited to their pre-class preparation level, with the system suggesting appropriate case studies, group tasks, or simulation exercises based on their demonstrated competencies.

The post-class phase reinforces learning through adaptive practice. The system generates personalized review materials by identifying knowledge gaps from classroom interactions and previous assessments. Virtual simulation tasks are recommended automatically for practical application, with the difficulty level adjusting dynamically based on performance. Simultaneously, instructors receive automated class-wide understanding reports, allowing them to focus support on challenging concepts.

This three-phase model creates an integrated learning cycle where each stage informs the next. The knowledge graph acts as the central coordination system, continuously updating recommendations based on learner interactions while maintaining resource coherence across platforms. By structurally unifying traditionally separate learning components, the KGDI model achieves authentic integration of blended learning.

The model's effectiveness stems from four key innovations: (1) its dynamic resource-linking mechanism that maintains contextual relationships between materials; (2) the dual-layer adaptation system that responds to both individual progress and group learning patterns; (3) the automated data analysis process that reduces instructor workload while improving feedback precision; and (4) the consistent user experience across all platforms through unified navigation via the knowledge graph interface. Together, these features transform the typically fragmented blended learning environment into a cohesive, intelligently adaptive educational environment.

Methodology

Building directly on the framework established in previous chapters, this section outlines the research methodology employed to evaluate the efficacy of the KGDI model.

Research Participants

The study involved 180 second-year undergraduates (102 females, 78 males; aged 19-21) from six intact classes (30 students each) of the mandatory *Introduction to Linguistics* course at Zhejiang Yuexiu University. Conducted during the 2024 spring semester, the sample included English majors from the School of English Studies. Classes were purposively selected based on equivalent pre-course language proficiency and were assigned to either the experimental group (KGDI model, n=90) or the control group (traditional blended learning, n=90). Pretests confirmed equivalent baseline knowledge ($p>0.5$) and similar demographics.

Research Instruments

Tests

Academic achievement was assessed via a comprehensive final examination with five question types evaluating theoretical understanding and applied competence across core linguistic categories. Developed by departmental faculty and piloted for validity, this 100-point instrument had strong reliability (Cronbach's $\alpha=0.87$). The identical test was administered under standardized conditions to both groups, allowing for a direct comparison of learning outcomes.

Questionnaires

A questionnaire evaluated students' learning experience by assessing three dimensions: perceived coherence of integrated resources (5-point Likert scale: 1=Very poor, 5=Excellent), cognitive load levels (10-point scale: 1=Low load, 10=High load; reverse-scored in analysis), and overall satisfaction (5-point Likert scale: 1=Very dissatisfied, 5=Very satisfied). The instrument underwent pilot testing and refinement before main study implementation.

Time Logs

The same instructor documented weekly time expenditure for both conditions using standardized digital logs, recording hours exclusively for three defined tasks: assignment grading, resource coordination, and personalized feedback. Tasks excluded lesson planning and Q&A sessions to isolate the efficiency impact of KGDI. System logs were validated through two methods: random reviews of 20% of timestamp entries (records of when specific actions happened in the system) and analysis of screen recording samples collected during grading sessions.

Interviews

Semi-structured interviews were conducted to gain deeper insights into students' experience with the KGDI system and the instructor's perspective on efficiency gains. Both students and the instructor participated in interviews to capture their respective authentic perceptions. Purposive sampling was used to ensure representation across different performance levels. All interviews were audio-recorded (averaging approximately 25 minutes each) and later transcribed word-for-word for subsequent thematic analysis.

Research Procedures

Over a 16-week term, both groups accessed identical learning resources, but through distinct instructional approaches: the experimental group (n=90) used the KGDI model, which organically connects all resources via knowledge graphs; the control group (n=90) adopted traditional blended learning with unconnected resources. The same instructor taught all sessions. All participants completed identical pre- and post-tests, with final exam scores serving as the primary outcome metric. Following the intervention, students completed questionnaires that measured cognitive load, perceived resource coherence, and overall satisfaction, while the instructor maintained weekly time logs to record instructional activities. After that, semi-structured interviews were conducted with purposively sampled students and the instructor.

Quantitative data were analyzed using independent samples *t*-tests for student outcome measures (final examination performance and survey responses), paired samples *t*-tests for instructor time efficiency data based on 16 weekly observations, Cohen's *d* for effect sizes (reported with 95% confidence intervals), and χ^2 tests for categorical comparisons (excellence rate). Qualitative data from interview transcripts underwent thematic analysis following Braun and Clarke's (2006) six-phase framework. All statistical tests used a two-tailed α threshold of 0.05, with Bonferroni correction applied for multiple comparisons where appropriate.

Results

The following findings are derived from the quasi-experimental study data, which cover multiple dimensions, including academic performance, cognitive load, resource coherence, overall satisfaction, and instructional efficiency.

Academic Performance

The standardized final examination served as the primary measure for evaluating the impact of the KGDI model on learning outcomes. As detailed in Table 1, a comprehensive analysis of examination results revealed consistent, statistically significant advantages for the experimental group across all assessment dimensions.

Table 1. *Final examination performance comparison*

Assessment Metric	KGDI Group (SD)	Control Group (SD)	Statistical Significance (95% CI)
Mean Total Score	82.1 (6.9)	76.2 (8.1)	$t(178)=4.37$, $p<0.001$, $d=0.78$ [0.45, 1.11]
Excellence Rate (≥80)	68%	41%	$\chi^2(1)= 15.3$, $p<0.001$
Essay Questions	82.3% (9.1)	71.9% (12.1)	$t(178)=4.88$, $p<0.001$, $d=0.85$ [0.51, 1.19]
Syntactic Analysis	74.5% (11.3)	67.5% (13.4)	$t(178)=3.98$, $p<0.001$, $d=0.62$ [0.29, 0.95]

The examination data reveal compelling evidence for the KGDI model's efficacy. Students using the knowledge graph system achieved markedly higher overall scores, with their mean performance (82.1±6.9) exceeding the control group (76.2±8.1) by 5.9 points. The distributional pattern further reinforced this substantial difference: 68% of KGDI learners achieved high scores (≥80), compared to only 41% in the traditional instruction group ($\chi^2(1)=15.3$, $p<0.001$).

Beyond overall performance, the benefits extended consistently across different question types. In essay questions requiring theoretical synthesis, the experimental group's accuracy rate (82.3±9.1%) significantly outpaced that of the controls (71.9±12.1%) by 10.4 percentage points ($t(178)=4.88$, $p<0.001$). Similar superiority emerged in syntactic analysis tasks, with KGDI participants achieving 74.5±11.3% accuracy versus 67.5±13.4% for conventional learners

($t(178)=3.98$, $p<0.001$). These parallel improvements suggest the new teaching model equally enhances competency in both conceptual understanding and applied analysis.

Questionnaire Outcomes

All 180 participants completed the post-intervention survey to systematically compare their learning experiences between the experimental and control conditions. The questionnaire consisted of 15 items measuring three theoretically grounded dimensions: perceived coherence of integrated resources (5 items, Cronbach's $\alpha = 0.88$), cognitive load levels (5 reverse-coded items, $\alpha = 0.82$), and overall satisfaction (5 items, $\alpha = 0.85$). The survey data provide complementary evidence of the KGDI model's effectiveness through direct feedback from learners. As summarized in Table 2, students using the integrated system reported significantly improved learning experiences across all measured dimensions compared to the conventional blended learning approach.

Table 2. *Learning experience survey results*

Evaluation Dimension	KGDI Group (SD)	Control Group (SD)	Statistical Test
Resource Coherence	4.0 (0.7)	3.2 (0.9)	$t(178)= 7.2$, $p<0.001$
Cognitive Load	4.9 (1.4)	6.1 (1.6)	$t(178)= -5.1$, $p<0.001$
Learning Satisfaction	4.1 (0.7)	3.2 (1.1)	$t(178)= 6.8$, $p<0.001$

Note:

Resource Coherence/Satisfaction: 1 = Very poor, 5 = Excellent

Cognitive Load: 1 = Low load, 10 = High load (reverse-scored)

Analysis of resource coherence ratings revealed that students using the KGDI system perceived significantly stronger integration among learning resources. The experimental group's mean rating of 4.0 (SD=0.7) on the 5-point scale contrasted with the control group's average of 3.2 (SD=0.9). This 0.8-point difference ($t(178)=7.2$, $p<0.001$) confirms the knowledge graph's success in creating a unified learning environment as perceived by users.

Cognitive load measurements also showed significant differences. KGDI participants reported lower effort levels (M=4.9, SD=1.4 on the 10-point scale), whereas their counterparts using conventional approaches described a higher cognitive burden (M=6.1, SD=1.6). Objective behavioral data further supported the finding: time-tracking revealed that control group students spent 32% more time switching between platforms than KGDI students.

While less pronounced than other dimensions, satisfaction outcomes still showed educationally meaningful differences. KGDI users' average rating of 4.1 (SD=0.7) on satisfaction items like "This learning approach was engaging" significantly exceeded the control group's 3.2 (SD=1.1). The 0.9-point difference in satisfaction scores was both statistically and educationally significant ($t(178)=6.8$, $p<0.001$), suggesting that learners genuinely preferred the new model.

The consistent positive outcomes observed across all three dimensions collectively demonstrate the efficacy of the knowledge graph system. Beyond improving objective learning outcomes, the intervention significantly enhanced learners' subjective experiences. Specifically, it reduced the cognitive demands caused by fragmented resources while simultaneously increasing the perceived coherence among learning materials and raising overall satisfaction levels.

Instructional Time Efficiency

The instructor's time logs provided evidence of workflow efficiency improvements under the KGDI system. Over the 16-week intervention period, systematic tracking of three key teaching tasks revealed substantial time reductions compared to traditional methods, as presented in Table3.

Table 3. *Instructor time efficiency comparison*

Task Category	KGDI Group (SD)	Control Group (SD)	Time Difference (95% CI)	Time Reduction	Statistical Significance
Grading	3.5 (0.9)	5.4 (1.2)	-1.9 ± 0.3 [-2.2, -1.6]	↓35.2%	$t(15) = -9.7$, $p<0.001$, $d=6.3$
Resource Coordination	2.1 (0.6)	3.8 (1.1)	-1.7 ± 0.4 [-2.0, -1.4]	↓44.7%	$t(15) = -7.8$, $p<0.001$, $d=4.3$
Personalized Feedback	2.3 (0.7)	3.5 (0.9)	-1.2 ± 0.3 [-1.5, -0.9]	↓34.3%	$t(15) = -6.2$, $p<0.001$, $d=4.0$
Total	7.9 (1.5)	12.7 (2.3)	-4.8 ± 0.6 [-5.4, -4.2]	↓37.8%	$t(15) = -12.1$, $p<0.001$, $d=8.0$

Notes:

Time units = hours/week (M ± SD)

n = 16 paired observations (same instructor over 16 weeks)

Time Difference = KGDI group time - Control group time

d = Cohen's effect size for paired samples

Analysis of grading tasks showed that automated grading reduced weekly grading time by 35.2%—from 5.4 hours (SD=1.2) to 3.5 hours (SD=0.9) ($t(15)=-9.7$, $p<0.001$, $d=6.3$). This time reduction was achieved through two key processes: first, Automated Scoring Technology (AST) based analysis automates the grading of most objective questions; second, AI identifies 20% of problematic cases for focused review, allowing instructors to only check these cases instead of all answers. The knowledge graph's resource organization similarly decreased preparation time, cutting coordination time by 44.7% from 3.8 hours (SD=1.1) to 2.1 hours (SD=0.6) weekly ($t(15)=-7.8$, $p<0.001$, $d=4.3$) via LOOM metadata auto-tagging that eliminated manual cross-platform sorting. Individualized feedback delivery also became more efficient, with weekly time investment decreasing from 3.5 hours (SD=0.9) to 2.3 hours (SD=0.7) ($t(15)= -6.2$, $p<0.001$, $d=4.0$), which was achieved by using GPT-4 generated templates designed for common error patterns.

Overall, KGDI reduced total weekly workload by 37.8% from 12.7 hours (SD=2.3) to 7.9 hours (SD=1.5) ($t(15)=-12.1$, $p<0.001$, $d=8.0$). The model succeeded in achieving these time savings while simultaneously improving learning outcomes and student experiences, showcasing its dual benefit of enhancing both educational quality and teaching efficiency.

Interview Findings

Analysis of interviews with 12 participants (six KGDI students and six control group students) and the instructor revealed three key patterns. These findings, drawn from 310 minutes of recordings, demonstrated strong inter-coder agreement ($\kappa=0.82$).

Integrated Resources Enhancing Learning Efficiency

KGDI students consistently reported that the knowledge graph's unified interface improved study efficiency. One student noted: "Before, I wasted time searching through different platforms. Now the system connects everything automatically." The instructor observed this shift firsthand: "Students stopped asking where to find materials and could spend more time on learning itself." This very change aligns with survey findings showing a reduced cognitive load, demonstrating how integrated access minimizes mental effort.

Personalized Learning Supporting Diverse Learner Needs

The system's adaptive features supported learners at different levels. A struggling student explained: "When I had trouble with phonetics, the system gave me simpler examples first." Higher-achieving students also benefited: "The syntax challenges pushed me just enough." The instructor confirmed: "I could identify who needed help before class and adjust accordingly." This explains the KGDI group's strong academic performance.

Automated Workflows Boosting Teaching Efficiency

The instructor described significant workflow improvements, particularly highlighting two key changes. First, automatic resource organization eliminated manual gathering time: "I used to spend hours collecting materials from different platforms. Now the system organizes them for me." Second, automated error detection enhanced feedback efficiency: "The system highlights common mistakes, so my feedback targets actual misunderstandings." These first-hand accounts correspond with measured time reductions in both resource coordination and grading tasks.

Meanwhile, control group students reported developing compensatory strategies: "I kept separate notes to integrate resources from different systems", further confirming the cognitive load differences identified in survey data.

The interview findings collectively demonstrate that the KGDI model enhances learning and teaching through three core mechanisms: integrated resource access, personalized content sequencing, and intelligent task automation. These qualitative insights complement the quantitative results, revealing how the system's design directly contributes to improved outcomes, reduced cognitive load, and enhanced efficiency.

Discussion

The empirical outcomes of this study confirm the effectiveness of the KGDI model in addressing resource fragmentation in blended learning. Building on evidence from our findings, three key dimensions are further explored: theoretical advancements in blended learning frameworks, practical innovations, and implementation considerations.

Theoretical Advancements in Blended Learning

This research extends cognitive load theory (Sweller, 2010) by demonstrating how knowledge graphs reduce unnecessary cognitive effort. The significant decrease in cognitive load among KGDI learners (mean difference=1.2 points, $p<0.001$) is directly linked to the system's automated resource integration, which eliminates manual search efforts that would otherwise distract from core learning tasks. These findings confirm that intelligent knowledge organization enhances working memory capacity, as evidenced by the 10.4-percentage-point advantage in essay tasks requiring theoretical synthesis (82.3% vs. 71.9%, $p<0.001$)—a result consistent with Mayer's (2020) cognitive theory of multimedia learning, which emphasizes coherent information presentation as a key determinant of learning efficiency.

The KGDI approach applies constructivist principles through adaptive scaffolding, embodying Vygotsky's zone of proximal development (Chaiklin, 2003) in digital form. By adapting materials to individual skill levels, the system creates personalized learning pathways that enable self-directed knowledge building across platforms—bridging traditional constructivism with modern adaptive learning systems. Unlike static blended models, KGDI's dynamic structure supports both structured knowledge acquisition and creative conceptual

connections, resolving a longstanding tension in educational technology theory.

Crucially, the study redefines educational technology frameworks by demonstrating KGDI's dual strength: structural integrity for core concepts and semantic flexibility for interdisciplinary connections. Unlike conventional systems that serve as passive content repositories (Balyshev et al., 2024), KGDI creates conceptual frameworks where linguistic elements become dynamic connection points, solving identified coherence problems in blended learning design while preserving the adaptability needed for complex knowledge domains.

Practical Innovations in Blended Learning

The KGDI model demonstrates significant improvements in addressing long-standing challenges of blended learning implementation. By automating resource management and progress tracking, the system achieves a 37.8% reduction in instructor workload ($p<0.001$), allowing educators to focus on higher-value teaching activities. These efficiency gains are supported by recent research by Dörrenbächer-Ulrich et al. (2024), which found that intelligent resource integration can reduce administrative burdens by 30-60%. Importantly, these time savings do not come at the expense of teaching quality, as evidenced by the KGDI group's superior academic performance across all assessment types.

Beyond efficiency gains, the model's adaptive sequencing capability effectively addresses personalization gaps in traditional blended environments. The system's dynamic adjustment of content difficulty—providing simplified examples for struggling students while offering advanced challenges for high achievers—creates what Wong et al. (2019) term "precision scaffolding". This is particularly valuable in mixed-ability classrooms, where conventional blended learning can inadvertently widen achievement gaps. KGDI successfully narrowed these gaps, supporting UNESCO's (2024) call for the equitable integration of technology in education.

Such improvements position KGDI as a meaningful advance over current blended learning approaches. By simultaneously enhancing instructional efficiency and enabling personalized learning, the system addresses what were once seen as competing priorities in educational technology design. Combined with earlier theoretical advancements, knowledge graph-driven integration emerges as a promising direction for the future of blended learning.

Implementation Considerations in Blended Learning

While the KGDI model demonstrates clear benefits, its implementation varies across different academic disciplines. The system works well for structured fields (e.g., linguistics with clear concept relationships) but less effectively in unstructured areas (e.g., humanities). As Guo et al. (2025) found, building knowledge frameworks for literature or philosophy takes much more expert time than for linguistics. Their flexible concepts and interpretive nature make automated connection mapping more challenging, which aligns with the OECD's (2023) finding that technology works differently across subjects, suggesting that KGDI may require subject-specific adjustments.

The technical requirements pose another challenge. KGDI's real-time data analysis and cross-platform connectivity demand cloud computing resources that may be beyond the reach of smaller institutions. Data from EDUCAUSE (2024) shows that only 41% of regional universities in developing areas have sufficient server capacity for such systems—findings that support UNESCO's (2023) concerns that educational technology could widen existing gaps between well-resourced and less-resourced schools if not implemented carefully.

Additionally, while our 16-week trial confirmed short-term effectiveness, a long-term evaluation is necessary to assess the sustained impact, particularly regarding knowledge retention and institutional adaptation, which is a standard practice for evaluating educational innovations.

Despite these considerations, the accumulated evidence supports KGDI as a theoretically grounded and practically viable advancement in blended learning. By addressing both the structural and semantic needs of modern education, while acknowledging the discipline-specific realities of implementation, the model offers a promising path forward for technology-enhanced learning.

Conclusion

The primary aim of this study was to develop and evaluate a Knowledge Graph-Driven Integration (KGDI) model to address resource fragmentation in blended learning environments. Using a quasi-experimental design with 180 undergraduate students, we systematically assessed its effectiveness across multiple dimensions.

Results demonstrate that the KGDI model significantly improves learning outcomes, with experimental group students achieving a mean final examination score of 82.1 compared to 76.2 in the control group ($p<0.001$). The integrated approach substantially reduced students' cognitive load (mean difference = 1.2 points) while significantly enhancing perceived resource coherence (mean difference = 0.8 points). For instructors, the system reduced weekly workload by 37.8% through automated resource coordination and progress tracking.

These findings advance both theory and practice in three key aspects. First, they provide empirical evidence that knowledge graph-based organization effectively reduces cognitive overload in technology-enhanced learning. Second, they demonstrate how intelligent systems can strike a balance between instructional quality and efficiency. Third, they offer practical implementation insights for scaling adaptive learning technologies in higher education.

However, the study also reveals essential limitations that suggest directions for future research. The model performs better in structured disciplines, such as linguistics, than in open-ended subjects. Additionally, the 16-week timeline limited assessment of long-term retention, and computational demands may pose barriers for institutions with limited technical infrastructure. Future studies should extend evaluation timelines and develop low-resource knowledge graph construction methods to promote broader adoption.

In conclusion, the KGDI framework presents a promising solution for solving resource fragmentation in blended learning. Intelligently connecting multi-platform resources through knowledge graphs creates more coherent learning experiences while reducing unnecessary cognitive burdens. As educational institutions increasingly adopt blended learning models, this research provides both a theoretical framework and practical solutions for optimizing technology integration in teaching and learning processes.

Funding: This work was supported by the 2024 General Research Project of Zhejiang Provincial Department of Education (Grant No. Y202455247); the 2024 Commissioned Research Project of Zhejiang Yuexiu University Research Base (Grant No. 2024WGYYWH01); and the 2025

Zhejiang Provincial Education Science Planning Project (Grant No. 2025SCG399).
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 check the grammar. After using this tool, the author reviewed and edited the content as needed and took full responsibility for the publication's content.

About the Author:

Lingling Lou is an associate professor at the School of English Studies at Zhejiang Yuexiu University. She has published over ten high-quality academic articles in peer-reviewed journals. Her research interests include foreign language teaching and intercultural communication.
ORCID: https://orcid.org/0009-0008-5190-0654

References

Abirami, A. M., Pudumalar, S., & Pandeeswari, S. T. (2023). Designing an effective teaching learning environment for an undergraduate software engineering course. *International Journal of Adult Education and Technology*, *14*(1), 1–21. https://doi.org/10.4018/IJAET.321655

Al Hajri, M., Bani Orabah, S., Domalaon, D., & Bijani, H. (2025). Evaluating the quality of blended learning in higher education: Teachers' and students' perceptions in an Omani EFL context. *Arab World English Journal*, Special Issue on CALL, (11), 20-42. https://doi.org/10.24093/awej/call11.2

Alharbi, W., & Basiouni, A. (2025). The impact of blended learning on EFL student satisfaction and retention: A customer satisfaction model approach. *Arab World English Journal*, *16*(1), 179–206. https://doi.org/10.24093/awej/vol16no1.11

Aljuaid, H. (2024). The impact of artificial intelligence tools on academic writing instruction in higher education: A systematic review. *Arab World English Journal* (Special Issue on ChatGPT), 26-55. https://doi.org/10.24093/awej/ChatGPT.2

Alsalihi, H. D. A. (2025). The Effect of a Blended Approach in Enhancing Paraphrasing and Summarization Skills among EFL Female Students in Academic Writing. *Arab World English Journal*, Special Issue on CALL (11), 177-191. https://doi.org/10.24093/awej/call11.11

Balyshev, P., Darinskaia, L., Molodtsova, G., & Oskina, A. (2024). Organizing students' online teamwork for sustainable development. *International Journal of Online Pedagogy and Course Design*, *14*(1), 1–16. https://doi.org/10.4018/IJOPCD.353300

Bernard, R. M., Borokhovski, E., Schmid, R. F., Tamim, R. M., & Abrami, P. C. (2014). A meta-analysis of blended learning and technology use in higher education: From the general to the applied. *Journal of Computing in Higher Education*, *26*(1), 87–122. https://doi.org/10.1007/s12528-013-9077-3

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

Chaiklin, S. (2003). The zone of proximal development in Vygotsky's theory of learning and school instruction. *Investigative Ophthalmology & Visual Science*, *58*(3), 1411.

Denbel, D. G. (2023). Challenges of Teaching and Learning Mathematics Courses on Online

Platforms. *International Journal of Online Pedagogy and Course Design*, *13*(1), 1–15. https://doi.org/10.4018/IJOPCD.321155

Dörrenbächer-Ulrich, L., Sparfeldt, J. R., & Perels, F. (2024). Knowing how to learn: Development and validation of the strategy knowledge test for self-regulated learning (SKT-SRL) for college students. *Metacognition and Learning*, *19*, 1–45. https://doi.org/10.1007/s11409-024-09379-w

EDUCAUSE. (2024). *2024 EDUCAUSE Horizon Report: Teaching and Learning Edition*. EDUCAUSE.

Fan, J. (2024). A big data and neural networks-driven approach to designing a student management system. *Soft Computing*, *28*(2), 1255–1276. https://doi.org/10.1007/s00500-023-09524-8

Fedorenko, S., & Kravchenko, T. (2023). Multimodal resources and students' motivation in English for specific purposes. *Arab World English Journal*, *14*(1), 59–70. https://doi.org/10.24093/awej/vol14no1.4

Garrison, D. R., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *The Internet and Higher Education*, *7*(2), 95–105. https://doi.org/10.1016/j.iheduc.2004.02.001

Graham, C. R. (2013). Emerging practice and research in blended learning. In M. G. Moore (Ed.), *Handbook of distance education* (3rd ed., pp. 333–350). Routledge.

Guo, Z., Guan, H., He, C., Xu, Y., & Liu, R. (2025). Knowledge graph-enhanced interleaved multi-head attention knowledge tracing. *International Journal of Data Warehousing and Mining*, *21*(1), 1–20. https://doi.org/10.4018/IJDWM.377619

Huang, R., Spector, J. M., & Yang, J. (2021). *Educational technology: A primer for the 21st century* (2nd ed.). Springer.

Ji, S., Pan, S., Cambria, E., Marttinen, P., & Philip, S. Y. (2021). A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE Transactions on Neural Networks and Learning Systems*, *33*(2), 494–514. https://doi.org/10.1109/TNNLS.2021.3070843

Mayer, R. E. (2020). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 57 – 72). Cambridge University Press. https://doi.org/10.1017/9781108894333.008

Mir, N. A., Fathima, A., & Suppala, S. (2023). A new paradigm for blended learning: Leveraging inverted pedagogy and digital communication to foster effective learning. *International Journal of Adult Education and Technology*, *14*(1), 1–10. https://doi.org/10.4018/IJAET.321654

Ni, N., & He, Z. (2024). The application of blended learning in animation majors at colleges and universities in the era of "Internet +" *International Journal of Web-Based Learning and Teaching Technologies*, *19*(1), 1–12. https://doi.org/10.4018/IJWLTT.336835

OECD. (2023). *OECD digital education outlook 2023: Towards an effective digital education ecosystem*. OECD Publishing. https://doi.org/10.1787/c74f03de-en

Shaya, N., AbuKhait, R., Madani, R., & Ahmed, V. (2025). Conceptualizing blended learning models as a sustainable and inclusive educational approach. *International Journal of Sustainability in Higher Education*, *26*(9), 90–111. https://doi.org/10.1108/IJSHE-03-2024-0167

Skulmowski, A., & Xu, K. (2022). Understanding cognitive load in digital and online learning: A new perspective. *Educational Psychology Review*, *34*(1), 171-196.

https://doi.org/10.1007/s10648-021-09624-7
Su, C., Jiang, Q., Han, Y., Wang, T., & He, Q. (2025). Knowledge graph-driven decision support for manufacturing process: A graph neural network-based knowledge reasoning approach. *Advanced Engineering Informatics*, *64*, 103098. https://doi.org/10.1016/j.aei.2024.103098
Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, *22*(2), 123–138. https://doi.org/10.1007/s10648-010-9128-5
Tabassum, Z. A., & Mohd Saad, M. R. (2025). A decadal examination of community of inquiry and blended learning in EFL/ESL development: A systematic review. *Arab World English Journal*, *16*(1), 179–206. https://doi.org/10.24093/awej/vol16no1.11
Tambunan, H., Silitonga, M., & Sidabutar, U. B. (2021). Blended learning in teaching technical teacher candidates with various types of learning styles. *International Journal of Mobile and Blended Learning*, *13*(3), 1-13. https://doi.org/10.4018/IJMBL.2021070104
UNESCO. (2023). *Global education monitoring report 2023: Technology in education: A tool on whose terms?* UNESCO.
UNESCO. (2024). *Global education monitoring report 2024: Education leadership*. UNESCO.
Wang, H., & Li, A. (2021). A systematic approach for an English education model based on the neural network algorithm. *Journal of Intelligent & Fuzzy Systems*, *40*(2), 3455-3466. https://doi.org/10.3233/JIFS-189383
Wang, K., Chai, C. S., Liang, J. C., & Sang, G. (2024). Exploring teachers' behavioural intentions to design artificial intelligence-assisted learning in Chinese K–12 education. *Technology, Pedagogy and Education*, *33*(5), 629–645. https://doi.org/10.1080/1475939X.2024.2369241
Wong, W., Chao, T., Chang, C., & Chen, K. (2019). Online scaffolding for data modeling in low-cost physical labs. *International Journal of Distance Education Technologies*, *17*(4), 1–20. https://doi.org/10.4018/IJDET.2019100101
Xiang, Y., Luo, Y., Ning, Z., Liu, Y., Yang, Z., Li, Z., et al. (2025). Construction method of multimodal knowledge graph for safety management in hydropower underground engineering. *Journal of Tsinghua University (Science and Technology)*, *65*(3), 433–445. https://doi.org/10.16511/j.cnki.qhdxxb.2025.26.014
Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education - where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp. 227-243
DOI: https://dx.doi.org/10.24093/awej/vol16no3.13

Exploring Students' Perception Towards the Use of Grammarly to Improve Grammatical Accuracy in Essay Writing

Nur Hallyda Mohd Sham
Department of Language and Humanities, Faculty of Educational Studies
University Putra Malaysia, Selangor, Malaysia
Corresponding Author: nur_hallyda@yahoo.com

Nooreen Noordin
Department of Language and Humanities, Faculty of Educational Studies
University Putra Malaysia, Selangor, Malaysia
Email: **nooreen@upm.edu.my**

Norhakimah Khaiessa Ahmad
Department of Language and Humanities, Faculty of Educational Studies
University Putra Malaysia, Selangor, Malaysia
Email: norhakimah@upm.edu.my

Received: 06/17/2025 Accepted: 08/19/2025 Published: 09/20/2025

Abstract
The teaching of grammar poses challenges for both educators and learners, often resulting in frustration and disengagement. Thus, it is impertinent to find ways to make grammar lessons effective. This study investigates the effectiveness of using the Grammarly application as a tool for providing corrective feedback to improve students' grammatical accuracy in essay writing. Specifically, it aimed to explore students' perceptions of the Grammarly application's use in improving their writing skills. The study also aims to highlight the significance of integrating digital tools into language instruction to support autonomous learning and motivation. A qualitative approach was adopted, involving semi-structured interviews with ten 19-year-old students from a public higher education institution in Selangor, Malaysia. The data were thematically analysed using four key themes: reasons, challenges, motivation, and method. Findings suggest that students perceive Grammarly as a helpful tool that helps them construct clearer sentences, increases their awareness of grammatical errors, and boosts their confidence and motivation to write. These insights offer practical implications for educators considering digital feedback tools in their grammar instruction. It is recommended that future studies investigate Grammarly's effectiveness in diverse contexts and with larger populations to understand its broader educational impact better.

Keywords: Automated corrective feedback, Grammarly, perception, essay writing, grammatical accuracy

Cite as: Mohd Sham, N.H., Noordin, N., & Ahmad, N.K. (2025). Exploring Students' Perception Towards the Use of Grammarly to Improve Grammatical Accuracy in Essay Writing. *Arab World English Journal*. *16*(3): 227-243 https://dx.doi.org/10.24093/awej/vol16no3.13

Introduction

Teaching grammar can be challenging for both educators and learners, necessitating the use of effective instructional strategies to support learning. Student writing frequently exhibits issues with grammar, punctuation, and spelling (Abdul Rahman et al, 2022). In Malaysia, ESL students encounter difficulties with English grammar due to limited instructional time, which hinders their understanding of grammatical concepts (Ngo et al., 2021). Historically, research on grammar pedagogy has focused on three teaching methodologies: focus on form, focus on meaning, and a combination of focus on form and meaning. The focus-on-forms method, which prioritises explicit grammar instruction, has been criticised for its reliance on teacher-led approaches and disregard for student needs (Özlem, 2021). This critique has catalysed a movement toward a focus on meaning, which eschews explicit instructional techniques, such as direct error correction. Identifying and correcting errors can be daunting for instructors, particularly in larger classes. Nevertheless, if errors remain unaddressed, students may internalise incorrect grammatical patterns that persist in their interlanguage (Albelihi, 2024). Therefore, an essential component of grammar instruction is the approach to providing feedback and correcting grammatical errors.

In recent years, there has been a growing interest in leveraging technology to enhance grammar teaching and improve learners' language accuracy (Feng Wu, 2022; Tu, 2022; Siti Aisyiyah, 2024). In the Malaysian context, several experimental studies have investigated the effectiveness of various corrective tools and software in language learning. A study by Tan (2020) emphasizes the need for Malaysian higher education institutions to integrate educational technological tools to enhance assessment and feedback practices. The current level of integration is deemed insufficient, prompting a call for a revamped framework to utilize technology more effectively in feedback processes. Research by Ganapathy et al (2020) focuses on written corrective feedback (WCF) in Malaysian secondary schools, particularly in ESL contexts. The study reveals that both teachers and students recognize the importance of WCF in improving writing skills. However, the feedback is often unfocused and indirect, with a preference for more comprehensive error marking. Hamid et al. (2021) explore the use of Google Docs for providing written electronic corrective feedback in Arabic language courses at a Malaysian university. Their findings indicate a preference for indirect feedback with metalinguistic comments among lecturers, while students favour direct feedback due to time constraints and a lack of confidence in self-correction. Thus, with the increasing number of technological tools available online, it is crucial to identify those that meet students' expectations and needs. Most importantly, the overall impact of technology on learner achievement should be evaluated based on its perceived effectiveness, relevant content, adequate practice, immediate feedback, engaging visual aids, ease of system use, and flexibility (Dahbi, 2023). Therefore, this study aims to explore students' perceptions of using ACF for grammar feedback, specifically through the Grammarly application. By investigating the attitudes, beliefs, and experiences of language learners, this research aims to determine the effectiveness of such tools in enhancing their writing skills and whether they feel more confident in applying grammatical rules independently. Accordingly, the following research questions are formulated:

1. What technological tools do post-secondary students use to improve their writing and grammatical accuracy?
2. What are post-secondary students' perceptions towards the use of the Grammarly application in improving their writing skills and grammatical accuracy?

Literature Review

Overview of Grammar in Language Learning

In hindsight, grammar instruction has been a fundamental component of language education, with various approaches evolving, including the grammar-translation method and communicative language teaching (Slater & Li, 2025). In contrast, more modern methods promote contextual learning, where grammar is taught through real-life examples and practical application, allowing students to understand the subtleties of language in a meaningful way. Currently, grammar teaching has shifted from a focus on writing to a broader understanding of language rules, including syntax and morphology (Slater & Li, 2025).

"It is noteworthy that a vast majority of the students associate grammar with rules and structures, and some of them find the term 'grammar' nerve-wracking" (Ameen, 2023, p. 144). Teaching grammar in Malaysian schools and universities presents a multifaceted challenge, influenced by pedagogical, linguistic, and socio-cultural factors. As a Muslim majority country, "the main reasons include the perceived complexity of English grammar rules and the difficulty in finding similarities between English grammar and any other language they are familiar with (i.e., Malay and Arabic)" (Rasyiqah et al, 2025, p. 9). As a result, educators face difficulties in choosing effective instructional methods, addressing linguistic differences between Malay and English, and overcoming students' attitudes towards language learning. These challenges stem from time constraints and the need for tailored instructional strategies to meet the diverse needs of learners. Challenges in grammar learning can also arise from varying student proficiency levels, differing learning styles, and the need to balance traditional methods with innovative practices. "Things are further complicated by the fact that many learners in foreign language contexts start learning an additional language after the critical period, their out-of-class exposure to this language is typically limited and somewhat by default they are provided with explicit instruction which focuses on rules that are likely to be remembered even when they achieve high levels of TL proficiency." (Pawlak, 2023, p. 2).

Pedagogical Issues in ESL Grammar Learning

In view of different types of teaching methods, explicit instruction can be time-consuming, while implicit methods may not provide the necessary clarity for all students (Rahman et al., 2024). Some studies believed that comprehension of grammar takes place when it is taught incidentally – focus on form (Kang & Lee, 2020; Amogne & Simegn, 2023), while others managed to prove the success of teaching grammar explicitly, which focuses on forms (Xodabande & Boroughani, 2023; Azizpour & Alavinia, 2021). Modern grammar teaching often integrates both explicit and implicit instructional strategies, recognizing the benefits of each approach (Normawati & Nugrahaeni, 2024). This blend of methodologies not only enhances student engagement but also fosters a deeper appreciation for the intricacies of language, ultimately preparing learners to navigate diverse communicative contexts with confidence and competence. Aside from that, many teachers lack sufficient training in grammar instruction, which affects their confidence and effectiveness in teaching complex grammatical rules, a problem that proves particularly significant in a secondary school setting (Munir & Ahmad, 2023; Chandran et al., 2022). Some teachers also resort to using their mother tongue to alleviate discomfort and confusion among students when teaching grammar (Natoo, 2024).

The Importance of Corrective Feedback

One of the methods to consistently improve the language is to be aware of one's errors and progress, which can be achieved through corrective feedback. The effectiveness of corrective feedback is influenced by various factors, including the type of feedback, the context in which it is delivered, and the learners' characteristics. Research has shown that different forms of corrective feedback, including written, oral, and metalinguistic feedback, can significantly impact learners' acquisition of grammar and their writing performance. (Iqbal & Saad, 2023; Maulida & Arianti, 2024; Liu & Hwang, 2023). A study by Tsao (2025) proves the importance of corrective feedback on students' writing ability, with 60% of the participants preferring teacher WCF, specifically metalinguistic explanations. Due to this, Corrective Feedback (CF) has always been used in writing classes to highlight students' mistakes and progress, in the hope that they can learn from them and avoid making the same mistakes in the future. This is empirical, as teachers must continually adjust different types of CF based on the situation and students' varying levels of proficiency. Effective CF not only helps students identify their errors but also fosters a growth mindset, encouraging them to view mistakes as valuable learning opportunities rather than setbacks.

Use of Grammar Checker to Improve Grammatical Accuracy

The use of a grammar checker or any form of automated writing evaluation (AWE) has become increasingly popular among both educators and students. Jamshed (2024) examines the impact of ChatGPT on the writing skills of English language learners. He has seen a reduction in errors across multiple grammatical categories compared to students who received traditional teacher feedback. Another study by Sun & Asmawi (2023) investigated the use of WeChat to teach Business English Writing (BEW). The approach managed to yield improvements in terms of layout, lexical resources, organisation, coherence, and grammar due to their exposure to model business letters, including lexical chunks, formulaic sequences, and structured organisation. These tools provide instant feedback and suggestions for improvement, allowing learners to enhance their writing skills in real-time. Another more commonly used AWE tool is Grammarly, a natural language processing machine that provides instant grammatical feedback to improve students' writing. Grammarly is also one of the most used AWE tools among EFL learners, according to a study by Alshehri (2025), with 73.3% of its respondents having experience using the application to improve their writing. Several studies also demonstrate its effectiveness in reducing errors, particularly in subject-verb agreement and verb tense (Vu, 2024; Gavina & Ibay-Pamo, 2024). The tool also highlights improvements in linguistic accuracy and writing performance, providing reliable and valid feedback that contributes to enhanced grammatical accuracy (Astutik et al., 2024).

Moreover, Elmotri (2025) investigated the explicitness of Grammarly's corrective feedback and students' perceptions when using the application. The students believe that the Grammarly application is less complicated, easier to use, and user-friendly, with a less crowded writing interface. Pedagogically, AWE tools have effectively enhanced teachers' efficiency by reducing grading time and allowing for more individualised feedback (Shadiev & Feng, 2023; Tan et al., 2022). The tools' ability to provide detailed feedback reports helps students to understand better and address their grammatical shortcomings (Su et al., 2024). This trend reflects a growing recognition of the importance of technology in education, as it empowers students to take ownership of their learning and develop greater self-awareness regarding their language use. As a result, integrating grammar checkers into the curriculum not only aids in developing technical

writing skills but also promotes critical thinking as students learn to analyse and revise their work more effectively.

Students' Perception and Preference on the Type of Corrective Feedback

Identifying students' perceptions and preferences when utilising specific teaching methods is crucial to ensure their effectiveness. Past research has shown that the use of AWE among ESL learners is well-received, with many identifying it as a helpful tool for correcting errors related to grammar, punctuation, and sentence structure. This feature is often highlighted as the most helpful aspect of the tool (Resiana et al., 2024; Dewi, 2023). The Grammarly tool has explicitly been linked to a reduction in grammatical errors and an overall improvement in writing skills (Vu, 2024; Suryanto et al., 2024), as well as increased motivation and confidence (Dong & Shi, 2021; Ranalli, 2021). Elmotri et al (2025) think that "systems like Grammarly, which uses colour coding for different errors, were particularly favoured, emphasising the need for intuitive interfaces in educational technology" (p. 395). In terms of Malaysian students, they have expressed positive perceptions towards the implementation of AWE, viewing it as a beneficial tool for improving writing skills. This positive attitude is crucial for the successful adoption of AWE in educational settings (Luqmanul et al., 2023). Moreover, the use of AWE in ESL classrooms has implications for teaching methodologies, suggesting a shift towards more technology-integrated instruction that can provide timely and relevant feedback to students. Despite this, there has been little study on Malaysian students' perceptions when using AWE, especially at the diploma level, where writing skills are a priority.

Method

The study employed a qualitative case approach, utilising a semi-structured interview. Creswell (2013) emphasized that "identifying the type of interview is practical and will net the most useful information to answer research questions" (p. 148). In this study, the semi-structured interview was conducted as a group with all participants responding to the same set of questions. After conducting interviews, researchers meticulously transcribe and analyse the data, identifying patterns, themes, and insights that emerge from participants' narratives. This iterative process of data analysis enables researchers to uncover rich and nuanced understandings of the case, which form the basis for developing theories, practical implications, and recommendations for future research or action.

Participants

10 Malaysian post-secondary students (three female and seven male), aged 19 years, enrolled in a diploma program at a government-funded institution in Selangor, Malaysia, were selected through purposive sampling. Prior to the interview session, these students underwent four sessions of writing class, during which they were exposed to the use of an automated-corrective feedback tool, namely Grammarly.

Instruments

Interview questions were adapted from the Attitudes towards Computer-Assisted Language Learning (A-CALL) questionnaire (Jalali & Ardebili, 2013), based on the Technology Acceptance Model (Davis, 1989), and were reviewed and piloted. The interviews consisted of nine questions,

structured into four blocks: the first block collected demographic data, the second block addressed research question one, and the third and final blocks focused on answering research question two.

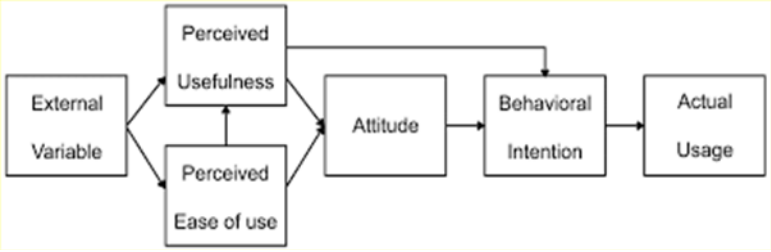


Figure 1. Technology Acceptance Model (Davis, 1989)

The data from the interview was analysed based on three steps: Transcription, Familiarisation and Coding. Firstly, the audio recordings were transcribed verbatim to ensure accuracy and sufficiency. Next, the transcriptions were read a few times to become familiar with any general themes. The data from the structured questions focused more on the demographic information, as well as the technological tools they have used or are familiar with. The remaining questions were designed to elicit open-ended and diverse responses. Eventually, both sets of data were coded and interpreted by identifying common keywords and details. Patterns and trends related to the study's focus were highlighted and annotated. These patterns were transferred into a table, which was categorised based on keywords, i.e., reasons, challenges, motivation, and method of learning. Lastly, those connections and variations were identified and justified.

Procedures

The interview session lasted about 40-50 minutes and was conducted in a face-to-face mode. The interview questions were asked in English. However, the students were allowed to converse in English and Bahasa (if necessary) to create a comfortable sharing session. The interview was audio-recorded and subsequently transcribed verbatim into a Word document for coding with careful and thorough consideration during transcription. The transcript was proofread, and any Bahasa words were translated into English. Audio files were safely stored and downloaded onto a computer.

Results

Experience in using ICT tools to learn English

To investigate the students' perception of using Grammarly tools, they were first asked about their general experience with online tools for learning the English language. Based on the interview, all students use and are familiar with the online dictionary to search for meanings. They also rely on the Google search engine to obtain information. Two respondents mentioned that they sometimes use Google Translate to find the meanings of words from their mother tongue language to English.

Sometimes I type a word I do not know the meaning of into the Google search engine and read the meanings that come up (Respondent one).

I will type the word that I need to know. Then I add the word 'synonym' next to it so that I can get more examples of vocabulary (Respondent two).

When asked about the term' grammar checker application,' most respondents are not familiar with the term but agree that there are specific applications that can help them check their grammar, such as Quillbot and Google Docs. Out of the 10 respondents, only two were familiar with the Grammarly application prior to the study being conducted.

Attitude and Experience in Learning English

The findings, which explored students' attitudes and experiences in learning English, were categorised into four key areas: reasons, challenges, motivation, and methods. Based on the interview, the researcher concluded that all the students have a positive attitude and experience when learning English, and the main reason is that they were exposed to the language from kindergarten and developed a love for it as they grew. One student (respondent two) mentioned that there was less pressure to learn English because it is something she has been learning for a long time.

What I like about English is that I have been learning it since I was in kindergarten, and my parents have taught me English from a young age. So it feels natural to learn the language. (Respondent four)

The students also identified challenges they face when learning English. It can be concluded that most of the difficulties occurred when they had to comprehend English grammar. Some of the concerns include when and how to apply grammatical structure when writing, and they experience confusion when using prepositions and tenses. Several students expressed particular concern when they realised their grammar mistakes. One student said they always found it difficult to answer error identification questions when they were in school.

My biggest challenge was when I realised I had been making the same basic grammatical mistakes. I only realised it when my teachers corrected it. (Respondent two)

My problem is with my lack of vocabulary. It affects my essay writing performance. (Respondent four)

Although I am already 19 years old, I still have problems understanding prepositions such as on, it, at. (Respondent five)

My problem is in pronunciation. I always have problems when trying to speak English. (Respondent eight)

My problem when learning English is the tenses. I think there are too many tenses and I get confused easily. (Respondent seven)

The students were also asked about their motivation to keep on learning English. In summary, the responses can be divided into intrinsic and extrinsic motivations. Three respondents think that they are motivated to learn English because they can use the language when travelling overseas and conversing with foreigners. Other respondents think that he/she is motivated to learn English because his/her siblings are fluent and want to follow their footsteps. One respondent is eager to improve their English because they want to get into law school. Other respondents are motivated to learn English because they feel insecure around their peers who speak English better. These findings indicate that all students generally have healthy and positive drives to learn the language, and they are in an environment where English is valued.

I want to learn English because I want to travel. If we go anywhere, people basically use English. So I want to learn more so that I can talk to strangers or foreigners without causing misunderstandings. (Respondent two)

What motivates me to learn English is my surroundings. All my close friends and family speak good English. So I want to be like them. (Respondent seven)
I have a sibling who scored well on her Malaysian University English Test (MUET) band. I want to be like her. (Respondent eight)
What motivates me to learn English is that I want to pursue a career in law as the next step. It's really important to be fluent in English. (Respondent nine)

Besides that, the students were also asked about their learning methods when it comes to English, and the responses were divided into two aspects: formal and informal. It can be concluded that all students underwent both types of learning methods. Informally, they learn English by watching movies and videos on social media, watching cartoons and using mobile applications.
I usually learn by watching videos on social media such as Tiktok. When local Tiktokers use English in their content, I will try to find errors. I learn from my mistakes and correct them myself (Respondent one).
I usually learn by doing exercises such as Fill-in-the-Blank (Respondent four).
I like to watch cartoons like Paw Patrol. I learn from the cartoons through their English conversation. If I don't understand a word, I will search it on Google Translate (Respondent three).
I enjoy doing word searches in my free time to learn more vocabulary (Respondent six).

Other than that, they received formal English lessons at school, which involved exercises and memorising words and texts. Based on these findings, it can be concluded that the students underwent a comprehensive learning experience that combined knowledge exposure, sources of motivation, and challenges. These continuous and active participations in their learning progress have proven that they have a general positive attribute to learning English.

Perception on the Use of Grammarly

To retain L2 learners' interest and boost their confidence in using automated corrective feedback, their preferences and perceptions should play a crucial role in the design and development cycle. In other words, to assess the effectiveness of the Grammarly application, it is necessary to study the students' perceptions of their experience using it. In this aspect, the researcher analysed the responses and divided them into two categories, i.e., negative and positive perceptions. In summary, it can be concluded that most students have a positive perception of the Grammarly application. The positive response focuses on awareness of grammatical mistakes, improved sentence structure, and a better level of motivation. One respondent reported that they like that Grammarly helps them correct sentences from informal to formal. On a similar view, one respondent shared that the application provides useful suggestions for writing clearer sentences.
"Grammarly is not only correct but suggests possible alternative sentences such as when I typed "I want to buy a book", the application suggests if I should write 'I want to buy books" instead" (Respondent five),

The suggestion makes the students realize that the two sentences illustrate two different meanings depending on the overall paragraph and contexts. These findings suggest that the feedback given makes the students rethink what they wrote and what type of noun they are going with which eventually teaches them to engage in self-evaluation as they compose or revise their essays. As the students use the tool more consistently, the feedback loop encourages them to take ownership of their writing process, ultimately fostering greater independence and self-direction in their learning journey (Wilson & MacArthur, 2024). Another respondent shared that the application helps him/her become more aware of the mistakes explicitly.

After I use Grammarly, I now understand the difference between the word 'other' and 'another' (Respondent four)
Grammarly helps me realize the correct way of using the word 'everyone' as in it should be followed by a singular verb instead of the plural verb." (*Respondent three)*
I realised that I have been making the same mistake when writing "its means" instead of "it means" since I started using the application. Now, that is the first thing I check every time I am done writing." (Respondent four)

This resonates with a study by Halim et al (2022) that showed Grammarly enhances students' self-directed learning in L2 writing by making them aware that Grammarly helped their writing process. Evidently, it is demonstrated through the students' initiative and motivation to learn as they experience needs and interests in a particular subject.

Respondents four and six also pointed out that Grammarly has helped them realise the differences between American English and British English. Since the Grammarly application is developed by an American company, the application is designed to flag British spelling as an error. This confusion has allowed the teacher to clarify to the students that other words are spelt differently depending on where they are used. The explanation has led to more discussion between the teacher and the students regarding its usage and appropriateness. This finding has proven that the application can spark more interactive and contextualised discussions about a grammar topic. This finding is in line with the previous studies that opt for Grammarly software as the intervention which noted that students perceived the software as a useful tool to use and a significant software to understand grammar rules (Dewi, 2023; Nurhidayah & Irawati, 2024)

Another respondent (Respondent three) thought that they feel more motivated to write because the application assigns marks every time they paste their essay, and seeing their marks improve has made them more driven to write.
Before using Grammarly, I never double-checked my essay after I finished writing. After using Grammarly several times, I notice that the software assigns marks every time I paste my essay. So, I feel bad every time I get a low mark. Now, I proofread my essay before pasting it into the software so that I can see improvements in my mark. (Respondent three)

This quote suggests that the Grammarly application encourages self-regulated and autonomous behaviour in users when it comes to evaluating their writing progress (Su, 2024). On top of that, the marking system feature provided by the application further motivates the students to aim better every time they try to complete their writing task. This finding further supports the notions from previous studies that automated writing evaluation tools motivate students more effectively, as they simplify the writing process and make it more engaging (Komara & Putra, 2024; Raihan et al., 2024). Some respondents also shared that they have started using Grammarly for their hobbies. This aligns with one of the items in the Technology Acceptance Model (TAM), specifically 'Behavioural Intent', where users exhibit a given behaviour, which then becomes a predictor of actual behaviour. The respondent has started to use the Grammarly application not just for writing assignment purposes, but also as a meaningful activity in their daily life.
"I like to post something on social media, and now I check my texts on Grammarly before posting on social media" (Respondent two).
"I like to write reviews of movies or series that I watch, and I will post them somewhere everyone can read. Now I use Grammarly to make sure my review is (grammatically) perfect" (Respondent four).

However, the application also brought out some negative perceptions from the students. They believe that although the application helps correct their sentence structure, some of the feedback also alters the overall meaning of the sentence.

"The application wants to correct our sentence structure, but then it changes what we want to say (in the essay)" (Respondent nine).

This finding aligns with several previous studies, which have reported that some students express negative perceptions due to misleading feedback and non-contextual suggestions (Faisal, 2023; Suryanto, 2024; Indiah Pane et al., 2025). Another respondent thinks that the free version of the application can only detect basic mistakes, and the price to subscribe to the premium version is too expensive. Several previous studies echo the same sentiments (Ummah & Bisriya, 2022; Dewi, 2023).

Preferred Types of Feedback

The correct type of feedback will help students to improve their comprehension. In this study, students were asked about their preferences for types of feedback when it comes to improving their writing and grammar. Although they believe that the Grammarly application is helping them sufficiently, the majority of respondents prefer a balance between teacher feedback and automated feedback. Respondent five pointed out that:

"*I feel grateful every time I receive feedback from my teacher but I understand that teachers can be busy. So the automated feedback is beneficial every time my teachers are not available to give feedback" (Respondent 10).*

Another respondent believes that the Grammarly application should be used in conjunction with teachers' assistance, as it provides a human touch and more personalised feedback that the tool may not be able to offer.

I appreciate receiving feedback from my teacher, but I understand that teachers are busy and not always available. So when my teacher is not available, I will use automated feedback to learn on my own. (Respondent 10)

I am a shy person. I will try to get my feedback from the online application first. If I still don't understand, I will refer to my teacher. (Respondent nine)

I prefer teacher feedback because I would like to listen to teachers' comment about my mistake face-to-face. I also like to ask follow-up questions. (Respondent five)

I think both types of feedback (teacher and automated) are good. But I prefer teacher feedback because I feel it is more personal. (Respondent six)

Based on these responses, it can be concluded that students still primarily rely on teachers' feedback and prefer to use automated feedback as a complement. This finding aligns with previous studies, which suggest that the Automated Writing Evaluation tool alone is insufficient. A study by Mat (2024) concludes that the traditional feedback method is still favoured by many. Some students also express distrust towards AWE-generated feedback and show a preference for human raters over automated scores (Astutik et al, 2024).

Discussion

The qualitative data were obtained through semi-structured interviews with ten students who underwent four weeks of writing lessons with Grammarly as their grammar checking assistant. The data were transcribed and analysed by identifying and categorising common themes.

Analysing interview data involves several steps: coding techniques, grouping similar types of information, and relating different ideas and themes to one another.

The study's first research question focused on students' experience in using ICT tools in general. The findings indicated that the students were familiar with using online dictionaries to find the meanings of words. They also used the Google search engine to find information and relied on Google Translate to get a quick translation of certain English phrases. When asked about the term "grammar checker application," most of them were unfamiliar with the term, but they agreed that there are specific online applications that can help them check their grammar.

The findings were further focused on the students' attitudes and experiences in learning English, with an emphasis on four key areas: reasons, challenges, motivation, and methods. It can be concluded that all respondents had a generally healthy and positive experience learning English, and all students identified the grammar aspect as the part they found most challenging. In terms of motivation, the findings are categorized into intrinsic and extrinsic motivations. Three out of ten students are intrinsically motivated when it comes to learning the English language. In contrast, the rest of the students believe that English is important for their career path and a competitive social environment. In terms of learning methods, all students underwent both formal and informal language lessons. They learned English in school by doing exercises and memorizing texts. Informally, the methods vary, ranging from learning English through movies and television series to the use of mobile applications. These findings indicate that students have a positive attitude towards their English language learning experience, as evidenced by their active participation in the learning process.

Regarding the use of the Grammarly application, it can be concluded that most students show a positive perception of it, mainly due to three reasons: awareness of grammatical mistakes, improved sentence structure, and a better level of motivation. Two respondents shared that the Grammarly application helps them to improve their sentence structure. This led them to rethink and question the intended meaning of their writing. Additionally, two respondents noted that they became more aware of the grammatical mistakes they had made in their writing. This has led to a follow-up discussion between them and the teacher about the feedback provided by the Grammarly application. In addition to that, two additional respondents shared that the Grammarly application boosts their motivation to write more and check for language errors before posting any writing on their social media accounts.

However, some respondents thought that the Grammarly application does not fully assist them in their comprehension. Occasionally, the application generates corrections that do not match what the writer means. Additionally, other issues that occur include the feedback being too brief and simplistic due to the free version, and users believe that the price to subscribe to the premium version is too high.

Lastly, the respondents were asked about their preferred type of feedback. The majority of the respondents prefer a combination of teacher feedback and automated feedback. They feel that a balance of both types of feedback can further solidify the grammar input and, at the same time, make them understand better. Based on these responses, it can be concluded that students still primarily rely on teachers' feedback and prefer to use automated feedback as a complement.

Conclusion

The results of this study confirmed the effectiveness of the Grammarly tool and its benefits in improving students' grammatical accuracy and writing skills. Based on the findings gathered

from the interview, several implications have arisen. The importance of feedback in grammar learning is proven to be crucial in this study. The results illustrate that students' awareness of their language errors is heightened, in the sense that they are more careful when producing English texts in formal assignments or during their daily activities. Moreover, they learn to adapt the use of a grammar checker application to their learning styles and preferences by realising that the usage may motivate them more to write in English, or vice versa. Nevertheless, the application helps students to explore more methods to self-assess their work and to receive prompt feedback. Further research is needed to investigate students' acceptance, availability, and proficiency levels, as well as their impact on other language skills, such as speaking and reading. Additionally, since this study focused on only ten students, further research on a larger population is needed to verify its effectiveness more thoroughly. Additionally, the current study only involved post-secondary students from a single institution. Further study is needed on students from diverse backgrounds to assess the feasibility and accuracy of this approach. Other ACF tools can also be explored to study their features and effectiveness. Lastly, the current study employed a small-group, closed-interview approach. Future research should focus on more empirical studies involving a larger dataset to gain a more in-depth and comprehensive perspective.

Acknowledgement
Not applicable
Funding: This research is not funded.
Acknowledgements: 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 utilised Scispace and ChatGPT to assist with literature review writing, and Grammarly to verify grammatical corrections. After using these tools, the authors reviewed and edited the content as necessary, taking full responsibility for the publication's content.

About the Authors:

Nur Hallyda Mohd Sham is currently pursuing her Master's Degree in TESL at the Faculty of Educational Studies, Universiti Putra Malaysia. She works as an English language teacher at a secondary school in Kuala Lumpur, Malaysia. Her research interests are teaching English as a second language and Educational Technology. https://orcid.org/0009-0002-8624-3304

Nooreen Noordin holds a doctoral degree from Universiti Putra Malaysia in the field of Teaching English as a Second Language. She is a senior lecturer at the Department of Language and Humanities Education, Faculty of Educational Studies, Universiti Putra Malaysia. Her research interests centers on language learning technology, content-based instruction and learning styles. https://orcid.org/0000-0002-4970-2682

Norhakimah Khaiessa Ahmad is a senior lecturer at Universiti Putra Malaysia, specializing in Second Language Teacher Education (SLTE). Her research centers on teacher and assessment identity within diverse instructional contexts. Her work advances scholarly discourse in English language teaching through publications and international conferences. https://orcid.org/0000-0002-0788-4553

References

Abdul Rahman, N. A., Zulkornain, L. H., & Hamzah, N. H. (2022). Exploring Artificial Intelligence using Automated Writing Evaluation for Writing Skills. *Environment-Behaviour Proceedings Journal*. 7(SI9), 547–553. https://doi.org/10.21834/ebpj.v7iSI9.4304ResearchGate+3

Albelihi, H. H. M., & Al-Ahdal, A. A. M. H. (2024). Overcoming error fossilization in Academic writing: Strategies for Saudi EFL learners to move beyond the plateau. *Asian-Pacific Journal of Second and Foreign Language Education, 9*(1). https://doi.org/10.1186/s40862-024-00303-y

Alshehri, M. (2025). Integrating Automated Writing Evaluation into Saudi EFL Students' Writing Practice. *Arab World English Journal,* 16(1):70-87. https://dx.doi.org/10.24093/awej/vol16no1.5

Ameen, R. M. (2023). Students' Attitudes towards Learning English Grammar: The Case of the English Department Students at the University of Raparin. *Arab World English Journal*, 14 (2) 138-145. DOI: https://dx.doi.org/10.24093/awej/vol14no2.10

Amogne, D., & Simegn, B. (2023). Students' motivation to learn grammar: The contribution of focus on form instruction. *Journal of Educational and Social Research, 13*(3), 96–106. https://doi.org/10.36941/jesr-2023-0060

Astutik, I., Widiati, U., Ratri, D. P., Jonathans, P. M., Nurkamilah, N., Devanti, Y. M., & Harfal, Z. (2024). Transformative Practices: Integrating Automated Writing Evaluation in Higher Education Writing Classrooms - A Systematic Review. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, 6(3), 423–441. https://doi.org/10.23917/ijolae.v6i3.23675

Azizpour, S., & Alavinia, P. (2021). The impact of focus on form and focus on forms instruction on grammar acquisition of the subjunctive by Iranian advanced EFL learners. *Teaching English Language*, 15(1), 215–249. https://doi.org/10.22132/TEL.2021.134368

Chandran, V. N., Albakri, I. S. M. A., Shukor, S. S. binti, Ismail, N., Tahir, M. H. M., Mokhtar, M., & Zulkepli, N. (2022). Malaysian English language novice teachers' challenges and support during initial years of teaching. *Studies in English Language and Education, 9*(2), 443–461. https://doi.org/10.24815/siele.v9i2.22974

Creswell, J. W. (2013). Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Sage Publications.

Dahbi, M. (2023). Integrating an Intelligent Language Tutoring System in Teaching English Grammar. *Arab World English Journal*, 14 (4) 189-196. DOI: https://dx.doi.org/10.24093/awej/vol14no4.11

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340

Dong, Y., & Shi, L. (2021). Using Grammarly to support students' source-based writing practices. Assessing Writing, 50, 100564. https://doi.org/10.1016/J.ASW.2021.100564

Dewi, U. K. (2023). Grammarly as Automated Writing Evaluation: Its Effectiveness from EFL Students' Perceptions. *Lingua Cultura, 16*(2), 155–161. https://doi.org/10.21512/lc.v16i2.8315

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

Faisal, F. (2023). Utilizing Grammarly in an Academic Writing Process: Higher-Education Students' Perceived Views. *JELTL (Journal of English Language Teaching and Linguistics)*, *8*(1), 23. https://doi.org/10.21462/jeltl.v8i1.1006

Ganapathy, M., Tan, D. A. L., & Phan, J. (2020). Impact of Written Corrective Feedback on Malaysian ESL Secondary Students' Writing Performance. *3L: The Southeast Asian Journal of English Language Studies*, 26(3), 139–153. https://doi.org/10.17576/3L-2020-2603-11

Gavina, M. A., & Ibay-Pamo, L. M. (2024). Effect of Automated Writing Evaluation in Higher Education Academic Writing Performance. *EDULANGUE*, 6(2), 118–137. https://doi.org/10.20414/edulangue.v6i2.8300

Halim, A., Sharina, . S., & Zur, S. (2022). Grammarly as a Tool to Enhance Students' Self-Directed Learning. *KnE Social Sciences*, *7*(8), 5–13. https://doi.org/10.18502/kss.v7i8.10719

Hamid, M. A. A., Sahrir, M. S., & Razali, K. (2021). A preference analysis and justification of Arabic written corrective feedback among instructors and undergraduates. *Indonesian Journal of Applied Linguistics*, *10*(3). https://doi.org/10.17509/IJAL.V10I3.31756

Indiah Pane, I. I., Alvindi, A., Dimas, D., & Sadri, I. (2025). Exploring Students' Perceptions of Grammarly as a Tool for Enhancing Grammar Accuracy in Writing. *Vision:Jurnal Bahasa Dan Pendidikan*, *20*(2), 90. https://doi.org/10.30829/vis.v20i2.4192

Iqbal, A., & Saad, M. A. (2023). The effect of focused and comprehensive written corrective feedback on writing accuracy: A comparative study. *AJELP: Asian Journal of English Language and Pedagogy*, 11(1), 120–133. https://doi.org/10.37134/ajelp.vol11.1.9.2023

Jalali, S., & Ardebili, M. (2013).The perceptions of Iranian EFL learners toward computer-assisted language learning. Paper presented at the International Conference on Current Trends in ELT, 20-22 May, Urmia, Iran. Retrieved from https://www.melta.org.my/journals/MAJER/downloads/majer10_01_04.pdf

Jamshed, M., Ahmed, A.S.M.M., Sarfaraj, M., Warda, W.U. (2024). The Impact of ChatGPT on English Language Learners' Writing Skills: An Assessment of AI Feedback on Mobile. *International Journal of Interactive Mobile Technologies (iJIM)*, 18(19), pp. 18–36. https://doi.org/10.3991/ijim.v18i19.50361

Kang, E. Y., & Lee, J. H. (2020). Effects of focus on form instruction through listening in blended learning on the development of grammar and listening skills. *Computer Assisted Language Learning*, *33*(8), 891–912. https://doi.org/10.15738/kjell.20..202011.662

Komara, C., & Putra, A. R. (2024). Students' perception of using grammarly website to support their english writing skill. *Klasikal: Journal of Education, Language Teaching and Science*, 265-272. https://doi.org/10.52208/klasikal.v6i2.1145

Liu, Y., & Hwang, G.-J. (2023). Effects of metalinguistic corrective feedback on novice EFL students' digital game-based grammar learning performances, perceptions, and behavioural patterns. *British Journal of Educational Technology*, 54(2), 456–473. https://doi.org/10.1111/bjet.13400

Luqmanul Hakim Zulkornain*, Azman Che Mat, and Nurul Ajleaa Abdul Rahman, "Does Automated Writing Evaluation (AWE): Improve Students' Writing? Focus on Technical Aspects and Readability," *International Journal of Information and Education*

Technology, vol. 13, no. 10, pp. 1656-1662, 2023. https://doi.org/10.18178/ijiet.2023.13.10.1974

Mat, A. C. (2024). Automated Writing Evaluation: Users' Perception and Expectations. *International Journal of Information and Education Technology*, *14*(2), 183–192. https://doi.org/10.18178/ijiet.2024.14.2.2039

Maulida, W. J., & Arianti, T. (2024). Teachers' use of oral corrective feedback during grammar classes at SMPN 189. *Foremost Journal*, 5(2), 104–111. https://doi.org/10.33592/foremost.v5i2.3693

Munir, N., & Ahmad, M. B. (2023). Challenges in teaching grammar to secondary school English learners in District Vehari. *Global Language Review, 8*(1), 31–43. http://dx.doi.org/10.31703/glr.2023(VIII-II).04

Natoo, B. (2024). *English language – challenges in teaching grammar* (pp. 112–119). https://doi.org/10.58532/v3balt4p8ch2

Ngo, K. L., Mohd Nor, J., Rani, M., & Abdullah, S. (2021). A Preliminary Study on English Grammar Proficiency Knowledge of Diploma Students. *International Journal of Humanities, Philosophy and Language*, 4 (15), 13-20. Retrieved from https://gaexcellence.com/ijhpl/article/view/544

Normawati, A., & Nugrahaeni, D. A. (2024). Grammar teaching and learning in English language class: Students' view. *English Education: Journal of English Teaching and Research, 9*(1), 41–48. https://doi.org/10.29407/jetar.v9i1.21536

Nurhidayah, A., & Irawati, T. (2024). Grammarly Through the Lens of Student Perception. *Jurnal Bahasa, Sastra Dan Studi Amerika*, *30*(1), 11. https://doi.org/10.20961/jbssa.v30i1.87858

Özlem F. (2021): Current Issue in Grammar Teaching, *The Criterion: An International Journal in English* Vol. 12, Issue-III, June 2021 ISSN: 0976-8165, 223-239. Retrieved from https://www.researchgate.net/publication/353039191_Current_Issues_in_Grammar_Teaching

Pawlak, M. (2023). Grammar Learning Strategies: Towards a Pedagogical Intervention. *Language Teaching Research Quarterly*. Vol 39, 174-191. https://doi.org/10.32038/ltrq.2024.39.12

Raihan, R., Febriani, B., & Irianti, L. (2023). EFL Students' Perception of Using Grammarly in Undergraduate Paper Writing. *International Journal of Innovation and Education Research (IJIER)*, 26-36. https://doi.org/10.33369/ijier.v2i1.29089

Ranalli, J., & Yamashita, T. (2022). Automated written corrective feedback: Error-correction performance and timing of delivery. *Language Learning & Technology, 26*(1). 1-25 Retrieved from https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/14e06dd5-5eab-49ac-9186-8a91e380fe4f/content

Rasyiqah Batrisya M.Z., et al. (2025). Addressing English Grammar Learning Challenges Among Malaysian Islamic Studies Students. *Malaysian Journal of Social Sciences and Humanities (MJSSH), 10*(1), 1-12. https://doi.org/10.47405/mjssh.v10i1.3199

Resiana, A. T., Zamzam, A., Putera, L. J., Amrullah, A., & Arfah, H. (2024). Effectiveness of Grammarly Application on the Students' Argumentative Writing Progress. *Journal of*

English Education Forum, *4*(3), 153–159. https://doi.org/10.29303/jeef.v4i3.722

Siti Aisyiyah., Novawan, A., Dewangga, V., & Bunarkaheni, S. (2024). Teaching grammar by using technologies: Unlocking language pedagogical potential. *Journal of English in Academic and Professional Communication*, *10*(1), 36–45. https://doi.org/10.25047/jeapco.v10i1.4615

Slater, T., & Li, J. (2025). Grammar in Language Teaching and Education. *The Encyclopedia of Applied Linguistics*, 1–9. https://doi.org/10.1002/9781405198431.wbeal0479.pub2

Shadiev, R., & Feng, Y. (2023). Using automated corrective feedback tools in language learning: A review study. *Interactive Learning Environments*, 1–29. https://doi.org/10.1080/10494820.2022.2153145

Su, Y. (2024). A Correlation Study of Automated Writing Evaluation System (Grammarly) and Chinese EFL Learners' Writing Self-Efficacy in Their Self-Regulated Learning. *Journal of Electrical Systems*. 1874-1895. https://doi.org/10.52783/jes.3904

Sun, L., & Asmawi, A. (2023). The effect of WeChat-based instruction on Chinese EFL Undergraduates' business English writing performance. *International Journal of Instruction*, 16(1), 43- 60. https://doi.org/10.29333/iji.2023.1613a

Suryanto, S., Habiburrahim, H., Akmal., S., Zainuddin, Z., Safrul, M.S., & Hanani, F. (2024). Scrutinizing the Impacts of Grammarly Application on Students' Writing Performance and Perception. *Jurnal Ilmiah Peuradeun*, *12*(2), 465-490. https://doi.org/10.26811/peuradeun.v12i2.1235

Tan, S., Cho, Y. W., & Xu, W. (2022). Exploring the effects of automated written corrective feedback, computer-mediated peer feedback and their combination mode on EFL learner's writing performance. *Interactive Learning Environments*, 1–11. https://doi.org/10.1080/10494820.2022.2066137

Tsao, J. J. (2025). EFL Students' Perceptions of Reading Portfolios and Teacher Feedback on Reflective Writing. *Arab World English Journal*, 16(1): 3-26. https://dx.doi.org/10.24093/awej/vol16no1.1

Tu, T. H. P. (2022). The effects of using education technology tools on learning grammar in secondary schools. *International Journal of Language Instruction*, 1(1), 41-52. https://doi.org/10.54855/ijli.22115

Ummah, L. K., & Bisriyah, M. (2022). EFL students' perception of Grammarly's feedback and how they deal with the inaccuracy. *JEES (Journal of English Educators Society)*, *7*(2). 163-172. https://doi.org/10.21070/jees.v7i2.1687

Vu, P. T. (2024). Using Grammarly in enhancing students' grammatical accuracy in English writing: A case at a private college in Hanoi. *Proceedings of the AsiaCALL International Conference*, 6, 81–96. https://doi.org/10.54855/paic.2466

Wilson, J., & MacArthur, C. A. (2024). Exploring the Role of Automated Writing Evaluation as a Formative Assessment Tool Supporting Self-Regulated Learning in Writing (pp. 197–220). *Informa*. https://doi.org/10.4324/9781003397618-14

Wu, F., Huang, S., Cheng, L. (2022). Analyzing the Application of Multimedia Technology Assisted English Grammar Teaching in Colleges. *Hindawi Scientific Programming.* 1-11. https://doi.org/10.1155/2022/4422754

Xodabande, I., & Boroughani, T. (2023). Mobile-assisted focus on forms in English for academic purposes instruction: Investigating the impacts on learning academic words.

Frontiers in Psychology, *14*, 1071555. 1-8.
https://doi.org/10.3389/fpsyg.2023.1071555Fr

Appendices

Appendix A: Interview Questions

A. Interviewee Background

1. Can you state your name and age?
2. How long have you been studying at this institution?

B. Attitude and Experience in learning English language

1. What do you like about English language?
2. What are some of the challenges you face when learning English language?
3. What motivates you to keep improving your English language proficiency?
4. How do you usually learn English grammar?

C. Experience in using ICT tools

1. In terms of studying, what are the ICT tools or applications that you usually use? (word processing, grammar checker, search engine)

D. Perception of the use of Grammarly

1. How does Grammarly help in improving your writing?
2. Does Grammarly help you to notice the errors more easily in your writing?
3. Does your understanding of the grammar structure become clearer after using Grammarly?
4. Do you feel more motivated to complete your writing tasks after receiving feedback from Grammarly?
5. What are your thoughts on the feedback generated by Grammarly?

E. Preferred types of feedback

1. What is your preferred type of language feedback? Why?

F. Demographics

Post Interview Comments and/or Observations:

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp. 244-258
DOI: https://dx.doi.org/10.24093/awej/vol16no3.14

Preserving Authorial Voice in Academic Texts in the Age of Generative AI: A Thematic Literature Review

Jokha Al Hosni
Centre for Preparatory Studies, Sultan Qaboos University, Muscat, Oman
Email: alhosnij@squ.edu.om

Received:05/21/2025 Accepted:08/10/2025 Published:09/20/2025

Abstract

Recent advancements in generative artificial intelligence (AI) tools have increasingly blurred the boundary between student-authored and AI-generated academic texts, raising urgent concerns about the erosion of academic integrity, creativity, and authenticity in student writing. While the existing body of literature has primarily focused on identifying stylistic differences to detect AI-generated content, this review adopts a distinct and underexplored perspective. This thematic review aims to examine how the use of AI writing tools, particularly in second language contexts, may influence students' *authorial voice* and *linguistic fingerprint*. The primary research questions guiding this study are: What are the key stylistic and linguistic differences between AI-generated and human-authored academic texts, and how do these differences affect the authorial voice and linguistic fingerprint in academic writing? This review concludes by highlighting key pedagogical implications for fostering authentic expression and preserving individual identity in academic writing. Drawing on 18 scholarly papers published between 2023 and 2025, this thematic literature review synthesizes up-to-date empirical findings across three core dimensions: (1) stylistic and linguistic features, (2) authorial voice and linguistic fingerprint, and (3) pedagogical implications for writing instruction. The evidence reveals that although AI-generated texts tend to be grammatically polished and structurally coherent, they often lack the emotional nuance, subtlety, rhetorical richness, and distinct linguistic markers that characterize human-authored writing, which may threaten the linguistic fingerprints of individual writers. This review highlights the urgent need for pedagogical strategies that not only integrate AI tools effectively into academic writing instruction but also actively protect students' individuality, authorial voice, and creative expression. This is particularly relevant in English as a second language (ESL) university settings, where the majority of the synthesized studies were conducted.

Keywords: Academic writing, academic integrity, authorial voice, ChatGPT, generative artificial intelligence, linguistic fingerprint, student identity, stylistic analysis

How to cite: Al Hosni, J. (2025). Preserving Authorial Voice in Academic Texts in the Age of Generative AI: A Thematic Literature Review. *Arab World English Journal, 16*(3):244-258. https://dx.doi.org/10.24093/awej/vol16no3.14

Introduction

The rapid evolution of generative artificial intelligence (AI), particularly AI chatbots like ChatGPT, Copilot, Gemini, and Claude, has significantly transformed the landscape of education and language teaching. A recent global survey conducted by the Digital Education Council among universities and corporations, which surveyed 3839 undergraduate and postgraduate students in different fields of study across 16 countries, found that most students are using AI tools. The survey indicated that 86% of students have already integrated AI into their academic work, with 54% using it weekly. ChatGPT was found to be the most commonly used chatbot (Digital Education Council, 2024).

Over the past three years, there has been a significant increase in research on the use of AI chatbots in English language learning, highlighting their potential benefits for language learners. AI-powered chatbots have proven effective in enhancing various language skills, including reading (Songsiengchai, 2023), vocabulary (Özçelik & Ekşi, 2024; Songsiengchai, 2023; Zhang & Huang, 2024), grammar (Harunasari, 2023; Songsiengchai, 2023), speaking (Du & Daniel, 2024; Duong & Suppasetseree, 2024; Khasawneh , 2023; Rusmiyanto et al., 2023; Songsiengchai, 2023) and writing (Baskara, 2023; Harunasari, 2023; Hawanti & Zubaydulloevna, 2023; Kwon et al., 2023; Silitonga et al., 2023; Yan, 2023; Yuan et al., 2024).

The use of AI chatbots as writing assistants has become increasingly popular due to their remarkable ability to generate ideas, themes and perspectives (Kasneci et al., 2023; Taecharungroj, 2023), reduce the time and effort needed to produce well-structured and coherent texts (Lund et al., 2023; Yan, 2023), and assist with proofreading by identifying and correcting spelling, grammar and syntax errors, thereby enhancing the overall quality of student writing (Imran & Almusharraf, 2023). However, there have also been growing concerns regarding the academic integrity and authenticity of the academic texts produced by students. Thus, many studies have tried to find ways to detect AI-written texts to uphold academic standards and maintain trust in student-authored work.

As reflected in the findings section of this paper, the current body of literature has extensively explored the stylistic differences between AI-generated and human-generated texts, primarily to develop criteria for detecting AI-produced content in student submissions. However, authorial voice and linguistic fingerprint have remained underexplored, with only a few recent studies beginning to examine how the use of AI chatbots may influence a writer's ability to express stance and personal style. This gap is particularly concerning in second language learning, where expressing personal academic voice is already a challenge. Therefore, this paper is an attempt to address this gap by shedding light on the potential loss of authorial voice in academic writing. The review attempts to answer the following research questions:

1. What are the key stylistic and linguistic differences between AI-generated and human-authored academic texts?
2. To what extent does the use of generative AI tools (e.g., ChatGPT) affect the authorial voice and linguistic fingerprint of authors in academic writing?
3. What are the pedagogical implications of integrating generative AI tools into academic writing instruction to safeguard students' authorial identity and linguistic fingerprint, particularly in university ESL classes?

Literature review

Since authorial voice and linguistic fingerprint are key concepts in this review, it is important to understand how these abstract concepts have been defined and operationalized in academic writing. These terms refer to the unique stylistic and rhetorical choices that reflect a writer's identity, stance, and engagement with the subject matter. An author's voice is expressed through vocabulary, sentence structure, tone, and perspective, which signals confidence and ownership in their writing. Similarly, a linguistic fingerprint encompasses the consistent patterns in language use, such as word choice, grammar, punctuation, and spelling, which uniquely characterize an individual's writing style. While linguistic fingerprinting is widely used in fields like forensic linguistics to establish authorship (Herbold et al., 2023), its significance in academic writing, especially among language learners, lies in preserving authenticity and individuality. Herbold et al. (2023) found that student writing often retains distinctive stylistic traits, suggesting that learners' voices and identities can be recognized through their written texts. In language learning contexts, developing a personal voice is not only a marker of proficiency but also a reflection of learner identity. Theoretical perspectives from scholars, such as Hyland (2018), Ivanič (1998), and Matsuda and Tardy (2007), further emphasized that authorship in academic writing extends beyond content delivery. It involves constructing a rhetorical presence shaped by social, cultural, and disciplinary norms (Biber, 2006; Charmaz & Mitchell, 1996; Tardy, 2012). Thus, voice is multifaceted and dynamic. It is shaped by the expectations of different genres and academic communities. Cultivating a unique voice plays a critical role in both academic success and personal development, particularly for second language writers.

Tannen (1993) and Ivanič (1994) offered methodological guidance that proposed that elements, such as word choice, sentence structure, hedges and boosters, emphatic expressions, and the use of personal pronouns, can serve as tangible markers to reflect the writer's stance and presence. Therefore, the review of the relevant studies in this paper analyzed a range of stylistic and linguistic features, including lexical diversity, lexical density, vocabulary richness, sentence length, and the use of linguistic devices, such as hedges and boosters.

Methodology

To provide a comprehensive and focused review, this study synthesized findings from 18 research papers that primarily focused on the comparative analysis of stylistic and linguistic features in AI-generated and human-authored academic texts of undergraduate and postgraduate students, including those that specifically highlighted aspects of authorial voice. A total of 33 papers were initially retrieved; however, they were filtered using the inclusion criteria listed in the table below.

Table 1. *Search criteria*

Category	Details
Keywords	"AI-generated text", "ChatGPT", "human-generated text", "stylistic features", "authorial voice", "linguistic fingerprint", "undergraduate/graduate students", "academic writing", "stylometric", "ESL/EFL"
Inclusion Criteria	● Qualitative, quantitative, and mixed-method comparative studies and systematic and tech reviews ● Published between 2023 and 2025 ● Focused on academic writing produced by undergraduate and graduate students ● Sourced from peer-reviewed journals, conference proceedings, and dissertations

ISSN: 2229-9327

This review adopts a thematic approach to synthesize the selected literature, which categorizes findings into three key themes: linguistic and stylistic features, authorial voice and linguistic fingerprint, and pedagogical implications. The structure reflects a deliberate progression from surface-level textual characteristics to deeper and more abstract dimensions of writing. The decision to begin with linguistic and stylistic features is based on their observable and quantifiable nature, which forms the foundation for identifying contrasts between AI-generated and human-authored texts. These surface markers, such as syntax, cohesion, lexical choices, and text rhythm, serve as a gateway to examining more nuanced aspects like authorial voice and emotional presence. By establishing these baseline differences first, the review then transitions to explore how such features impact the expression of individuality, cognitive engagement, and emotional resonance in academic writing.

Findings

Linguistic and Stylistic Features Comparison

This section outlines the key linguistic and stylistic differences between AI-generated and human-authored academic texts as identified in the synthesized literature.

Zindela (2023) investigated the linguistic differences between 20 AI-generated and 20 human-written essays by first-year Tswana-speaking English students at a South African university. The study found that AI-generated essays using ChatGPT-3.5 featured more complex syntactic structures, such as longer sentences and varied clause types, and they demonstrated higher lexical diversity by incorporating a wider range of low-frequency vocabulary. In comparison, human-written essays included a greater proportion of content words, indicating a denser delivery of information. Additionally, they made more frequent use of function words, which reflects natural language usage patterns.

AlAfnan and MohdZuki (2023) did not conduct a human–AI comparison; instead, they focused on a genre-based analysis of ChatGPT-4-generated responses across three genres: case studies, business correspondence, and academic writing. Their analysis focused on various linguistic and stylistic aspects, including sentence length, paragraph structure, tense, voice, mood, pronoun usage, word choice, keyword density, lexical density, lexical diversity, and reading ease. The study revealed that ChatGPT-4-generated academic writing typically consists of paragraphs containing three to four sentences, each averaging 16 to 19 words. The sentences are predominantly written in the declarative mood, primarily using the active voice. However, agentless passives appear occasionally, with a range of present tenses. The use of keywords and lexical density is high, but the lexical diversity is low, which results in an overall reading ease of moderate difficulty. This is further complicated by the presence of undefined abbreviations. These findings complement those of Zindela (2023) by reinforcing the notion that AI texts are technically dense, which may reduce the organic and context-sensitive cohesion typically found in human language.

Mizumoto et al. (2024) conducted a study involving 140 first-year Japanese university students enrolled in an EFL (English as a Foreign Language) program to examine whether essays written by Japanese EFL learners could be distinguished from those generated by ChatGPT. The results revealed that AI-generated essays typically displayed greater lexical diversity and syntactic complexity, giving them a more polished and formal appearance. In contrast, student-written essays contained more modals, epistemic markers, and discourse markers compared to ChatGPT as they attempted to express their voice. For example, more hedges and boosters were used in

human-written essays, which indicates the writer's stance and engagement with the reader, a finding that was also echoed by Johansson (2023) and Amirjalili et al. (2024). Hedges express uncertainty or caution, while boosters convey confidence or emphasis. The study also showed that student essays exhibited a higher frequency and larger range of reader pronouns, reflecting a more personalized and context-specific writing style.

Wang et al. (2024) conducted a large-scale analysis of 4,895 journal entries written by 789 university students enrolled in a mental-wellbeing course and compared them with 225 parallel essays generated by ChatGPT-4. The findings revealed that the AI-produced texts were consistently more difficult to read than the student-authored entries. Specifically, GPT-4 essays utilized a broader vocabulary with longer words, but constructed shorter sentences. Conversely, human essays featured longer sentences with a more limited vocabulary. These results reinforce the previous observations, but they contradict both AlAfnan & Mohdzuki's and Zindela's results regarding sentence length, which could be due to differences in genre, task design, or prompt specificity that influenced how ChatGPT structured its responses compared to students' natural writing.

To provide a more multidimensional perspective, Biber's (1988) multidimensional framework served as the analytical foundation for Sardinha's (2024), Ali's (2024), and Markey et al.'s (2024) studies. Biber's Multidimensional Framework is a corpus-based analytical model that identifies co-occurring linguistic features to describe variation across registers along key dimensions, which are 1. involvement vs. information density, 2. narrative vs. non-narrative focus, 3. explicitness, 4. persuasion, 5. abstraction, and 6. real-time elaboration. By using Biber's framework, these studies offered nuanced insights into the discourse characteristics of human and machine-produced writing.

Sardinha (2024) provided a systematic comparative analysis of AI-generated (using ChatGPT 3.5) and human-authored texts across multiple registers, including academic prose. Focusing on academic writing, the study examined introductions extracted from published research articles from two disciplines, Applied Linguistics and Chemistry, and evaluated how AI-produced texts align with human-written counterparts along five linguistic dimensions. The results revealed that, among the various registers analyzed, AI-generated academic texts exhibited the closest resemblance to human-authored academic writing. Nevertheless, notable discrepancies were observed. On Dimension 3 (explicitness), AI academic texts in both disciplines were found to be substantially less explicit compared to human-authored introductions. Additionally, the study highlighted that AI-generated academic texts tended to underutilize abstraction markers (Dimension 5) and exhibit reduced rhetorical functions, such as persuasive language (Dimension 4), when compared to human-written academic prose.

These findings were reinforced by Ali (2024), who applied the same analytical framework to compare the introduction sections of 20 dissertations written by Saudi postgraduate students with 20 corresponding texts generated by ChatGPT. The study examined more than 150 linguistic features, categorizing them according to Biber's Multidimensional Framework. The analysis revealed that AI-generated introductions were more informational (Dimension 1) and more explicit (Dimension 3), but less narrative (Dimension 2) and less overtly argumentative (Dimension 4) than human-written texts. A notable difference also emerged for Dimension 5, where human-authored introductions displayed greater abstraction, while ChatGPT's outputs tended toward more concrete and non-abstract discourse. Although both sets of texts featured a wide range of

lexico-grammatical patterns, the human-authored versions demonstrated deeper rhetorical sophistication and stronger alignment with disciplinary conventions.

Markey et al. (2024) investigated how ChatGPT-3.5-generated academic introductions compare to human writing by analyzing 200 texts from first-year undergraduates, 174 high-quality texts from advanced students, 159 texts from published academics, and 200 ChatGPT-3.5-generated texts. Using Biber's multidimensional framework, they projected all texts onto Dimension 1 and Dimension 4. They found that ChatGPT's introductions were significantly more informationally dense (Dimension 1) but consistently exhibited a sedimented, rigid structure and produced prose that felt fluffy and empty, which lacked concrete examples and rhetorical depth. While its argumentation features (Dimension 4) superficially resemble human-written proposals, the texts remained formulaic and dialogically closed.

While Sardinha (2024) and Ali (2024) both highlighted that AI-generated academic texts tend to underutilize abstraction markers and, thus, produce more concrete discourse (Dimension 5), their findings diverge slightly on explicitness (Dimension 3). Sardinha reports that AI texts are less explicit than human-authored introductions, whereas Ali finds that AI texts appear more explicit, but still lack narrative depth and rhetorical complexity. This could be explained by variations in the genres selected, the academic levels of the human comparison groups, and differing interpretations of what constitutes explicitness and abstraction across studies. Overall, the collective findings of these three studies demonstrate that ChatGPT-generated academic introductions can mimic the surface-level informational structure and formal conventions of human-written texts. However, they consistently fall short in rhetorical depth, abstraction, and persuasive richness. This may result in prose that is often formulaic and rigid.

Finally, in a recent study, Durak et al. (2025) investigated stylistic distinctions using a sample of 128 education faculty students. The results indicated that AI-written texts differ from human-written texts in terms of syntactic use of sentences, singular words, conjugations, and stop words. The findings revealed that human-generated articles exhibited higher average singular word usage and longer sentence lengths compared to AI-generated articles, which suggests a more complex and nuanced language structure in human writing. Vocabulary usage in AI-generated texts was also more repetitive, whereas human-written texts showcased a richer and more varied lexicon. In terms of sentence structure, human-written articles displayed more complex and nuanced constructions compared to the formulaic patterns typical of AI-generated texts.

Building on these findings, the following section synthesizes insights from studies that specifically examine how such linguistic and stylistic differences influence the expression of authorial voice in academic writing.

Impact on Authorial Voice and Linguistic Fingerprint

In a tech review on the use of ChatGPT in second language writing, Barrot (2023) addressed key constructs, such as language form, style, coherence, cohesion, and writing conventions. The review noted a rigid template-like structure in ChatGPT's outputs, with similar organizational patterns applied across different topics, a recurring finding that is highlighted in the previous section. The researcher believed that ChatGPT, as an AI chatbot, "may mimic a high degree of human intelligence and excel in many aspects of writing. However, this tool cannot totally replace humans because it lacks the emotional depth and life experiences that contribute to an individual's voice, identity, and distinctiveness in writing" (p.4). This finding is echoed in Oritsegbemi's (2023) review of empirical studies on AI's impact on the emotional dimensions of

human communication, in which she similarly concluded that AI cannot fully reproduce the complexity and subjectivity of human intelligence and emotion due to its lack of innate feelings and understanding of abstract concepts.

As one of the earliest empirical studies in this area, Johansson (2023) conducted a comparative analysis of argumentative essays produced by undergraduate English literature students and those generated by ChatGPT. The study explored dimensions, such as authorial presence, assertiveness, and self-identification. The findings revealed that the AI-generated essays were contextually appropriate and coherent, but they lacked factual precision and the nuanced authorial traits found in human writing. Human-authored texts exhibited greater specificity, rhetorical variety, and engagement with sources. Moreover, the effective employment of hedges and boosters contributes to a stronger sense of academic voice and identity. In contrast, the AI-generated texts demonstrated repetitive phrasing and a generalized and less personalized tone.

Similarly, Fedoriv et al. (2023) analyzed student-submitted academic essays from a university Learning Management System (LMS) to distinguish human and AI (ChatGPT-3.5) discourse. The study's findings revealed that human-authored texts naturally integrate discourse markers, cultural context, and authentic emotional depth. These elements demonstrate clear signs of critical thinking, planning, and socio-cultural awareness. In contrast, AI-generated texts are coherent and linguistically polished, but they often overuse ornate language, repeat phrases, and display stylistic mismatches. Unlike human writing, AI writing lacks nuanced cultural understanding and tends to produce template-like structures, with limited evidence of true creativity or deep reasoning.

From an evaluative perspective, Mahyoob et al. (2023) proposed a framework for assessing ChatGPT's effectiveness in academic writing across various genres. A critical analysis was conducted on selected ChatGPT-generated texts using a framework built around six core principles: Relatedness, Adequacy, Limitation, Authenticity, Cognition, and Redundancy. The findings revealed that, despite ChatGPT's impressive language capabilities, several issues persist in its academic outputs. These include content repetition, illogical reasoning, fabricated references, factual inaccuracies, hallucinations, and limited pragmatic interpretation. Therefore, the researchers concluded that ChatGPT has limited capacities to deliver academically sound, authentic, and reliable content.

Echoing similar patterns, Mizumoto et al. (2024) noted that human-generated academic texts tend to include more hedging and boosting devices, reflecting a nuanced expression of stance and engagement. AI-generated texts, though fluent and grammatically sound, lacked these subtle features, resulting in a detached and neutral tone.

In a perception-based study aimed at mapping the impact of ChatGPT on academic writing, Malik et al. (2023) surveyed 245 undergraduate students from 25 tertiary Indonesian institutions to explore their attitudes toward AI tools in academic essay writing. The results revealed generally positive perceptions. Participants recognized the value of AI-powered tools for grammar correction, plagiarism detection, language translation, and essay outline generation. However, some participants expressed concerns about the potential negative effects of AI on creativity, critical thinking, and ethical writing practices.

Al-Zubaidi et al. (2024) complemented this finding by revealing concerns from a broader academic community about the lack of emotional depth and rhetorical nuance in AI-generated texts. The researchers surveyed 180 undergraduate students, postgraduate students, and faculty members at Moroccan universities. The findings identified the lack of human touch as one of the

main limitations of using ChatGPT, alongside its inability to understand sarcasm and irony. This absence of emotional nuance and contextual sensitivity often results in texts that may fail to capture the depth, subtlety, and rhetorical richness typically found in human-authored academic writing.

In a complementary finding, Wang et al. (2024) highlighted several shortcomings of AI-generated writing, including its general and repetitive language, robotic tone, lack of emotional nuance, absence of personal voice, and limited critical thinking and creativity. Participants in the study suggested that AI-generated writing could be improved by incorporating personal stories, connections to the topic, feelings, and thoughts, and by removing repetitive language.

Adding a detailed case comparison, Amirjalili et al. (2024) conducted a comparative analysis of authorship and voice by examining two essays: one written by a second-year English literature student and the other generated by ChatGPT, using a refined version of Helms-Park and Stapleton's Voice Intensity Rating Scale. The analysis focused on assertiveness, self-identification, and authorial presence. The findings reported that the human-authored essay skillfully blended active and passive structures, employed rhetorical questions, robustly engaged the reader, and demonstrated precise sourcing, which contributed to a nuanced and distinctive voice. In contrast, the AI text maintained coherence and occasional first-person usage. However, it relied heavily on active voice, lacked hedges and boosters, and frequently repeated formulaic phrases. This compromised lexical diversity, factual accuracy, and contextual depth, which are more prevalent in the human-authored essay. These findings underscore the persistent gap in capturing authentic scholarly voice in AI-generated writing.

An additional perspective was introduced by Amirjalili et al. (2024), who linked the erosion of authorial voice to students' experiences of *cognitive dissonance* (CD) when engaging with generative AI tools. The researchers coined the term "AIgiarism" to describe a nuanced form of plagiarism rooted in the internal conflict between upholding academic integrity and relying on the efficiency of AI-generated content. This cognitive dissonance, defined as the discomfort caused by holding contradictory beliefs, may lead students to suppress or compromise their authentic writing style in favor of AI-generated text. Over time, this tension can result in a detachment from personal voice. This may reinforce concerns about the diminishing presence of individuality and identity in student writing shaped by AI.

Similarly, by employing a voice analytic rubric, Nañola et al. (2025) extended this discussion by examining differences in authorial voice across 12 undergraduate student-written essays and 12 AI-generated texts using Lehman and Sułkowski's Voice Analytic Rubric. Their findings revealed that student essays predominantly exhibited a Collective (C) voice, characteristic of novice academic writers aligning with disciplinary norms, while AI-generated texts often employed an Individual (I) voice, associated with a more detached and expert-like stance. Notably, the study indicated that the predictable and highly polished nature of AI writing can potentially conceal students' developing academic voice.

Durak et al. (2025) added to these observations by noting that AI-generated texts occasionally lacked the depth and critical insight characteristic of human authorship, as human writers display greater adaptability in tone and style to suit different educational topics.

Jiang and Hyland (2025) further illustrated these concerns by comparing 145 student-written essays with 145 ChatGPT-generated essays to examine the presence of engagement markers, such as rhetorical questions, personal commentary, and direct reader appeals. Their findings revealed that student-authored essays featured significantly more engagement markers, resulting in a more dynamic and persuasive tone. Human-written essays stood out for their

rhetorical impact. For instance, they frequently used questions to provoke thought, offered personal reflections, and addressed the reader directly. These techniques enhance engagement. In contrast, AI-generated essays, though coherent and grammatically accurate, lack personal voice, interactional features, and rhetorical nuance, which makes them feel mechanical and less compelling. It is believed that these limitations were attributed to ChatGPT's reliance on statistical patterns rather than authentic communicative intent.

Pedagogical Implications for Second Language Teaching and Learning

This section synthesizes the pedagogical implications to help preserve students' authorial voice, especially for language teaching and learning in higher education, as most of the findings are gathered from research focusing on ESL learners in higher education contexts.

The main pedagogical implication emerging from recent studies is that writing instruction must deliberately safeguard and cultivate students' authentic authorial voice in an era where AI tools like ChatGPT are increasingly integrated into academic work. Teachers should adopt an integrated yet critical approach by using AI outputs as supportive tools while ensuring students remain the primary authors of their texts. Ali (2024) highlighted how the overlap between ChatGPT's language patterns and those of human authors challenges traditional notions of authorship and creativity, which underscores the need to examine AI's role in shaping academic voice critically. Similarly, Malik et al. (2023) advocated for balanced AI integration that complements rather than replaces human creativity and critical thinking. Fedoriv et al. (2023) cautioned that overreliance on AI deprives, especially non-native speakers, of opportunities for genuine expression and cognitive engagement. To counter this, Barrot (2023) suggested that second language teachers should help students cultivate a distinctive voice by weaving personal experiences and stories into their writing, an element difficult for AI to replicate. Nañola et al. (2025) reinforced this by proposing that teachers embed authorial voice into instruction and assessment by expanding assessment rubrics with context-specific features. They also suggested analyzing AI's stylistic patterns to distinguish human writing from AI-generated text.

To develop students' awareness of rhetorical choices and stylistic variation, Markey et al. (2024) and Mizumoto et al. (2024) recommended that instructors use AI-generated texts to teach genre conventions, cohesion, and the impact of linguistic choices on rhetorical effect. This can help students recognize and avoid rigid syntax, abstract overgeneralizations, and a lack of exemplification. Jiang and Hyland (2025) similarly emphasized explicit instruction in rhetorical strategies to address gaps in engagement markers. They suggested that comparing AI and student drafts can sharpen students' ability to personalize and refine their writing. Amirjalili et al. (2024) called for targeted activities to build nuance, individuality, and lexical diversity while fostering students' technological literacy to use AI responsibly.

Discussion

Across these synthesized comparative studies, a consistent pattern emerges highlighting notable linguistic and stylistic differences between AI-generated and human-written academic texts. AI-produced academic prose often exhibits higher syntactic complexity and greater lexical diversity on the surface. However, it tends to lack more profound rhetorical cohesion and discipline-specific nuances. Studies by Zindela (2023), Mizumoto et al. (2024), and Wang et al. (2024) show that AI texts typically rely on denser information packing and a broader vocabulary, yet human-authored texts demonstrate more authentic use of hedges, boosters, modals, and reader

engagement markers, which reflect a more substantial authorial presence and contextual awareness. Research employing Biber's multidimensional framework (Ali, 2024; Markey et al., 2024; Sardinha, 2024) further confirmed that AI-generated academic writing scores higher in information density but falls short in explicit referencing, narrative depth, abstraction, and persuasive language. These stylistic gaps result in AI texts that often feel rigid, formulaic, and superficially polished. Additionally, they lack the rhetorical richness, emotional depth, cohesive argumentation, and disciplinary alignment that are found in human-generated academic writing.

The reviewed studies on authorial voice and linguistic fingerprint have identified key characteristics that distinguish the academic voice of AI-generated academic texts and human-generated ones. The table below summarizes these differences.

Table 2. *Key differences in authorial voice between AI-generated and human-generated texts*

Academic text	Key characteristics
AI-generated	• cannot fully replicate the emotional depth, complexity, subjectivity, and life experience • lacks innate feelings and understanding of abstract concepts, with limited pragmatic interpretation • overuses ornate language • has a detached and neutral tone • has a robotic tone
Human-generated	• exhibits greater specificity and high engagement with sources • shows cultural context and authentic emotional depth • shows clear signs of critical thinking, planning, and socio-cultural awareness • exhibits subtlety and rhetorical richness • has a dynamic and persuasive tone

The reviewed studies also emphasized the need to implement pedagogical strategies that intentionally preserve authorial voice, which includes intentional curriculum and assessment design, reflective use of AI, and sustained pedagogical focus on originality, rhetorical awareness, and creative expression. This can ensure that the distinctive human elements of academic writing are not lost in an AI-enhanced learning environment.

Conclusion and Future Directions

This review sheds light on a critical yet underexplored dimension of AI in education: how generative AI tools affect students' authorial voice and linguistic fingerprint, which are key markers of identity, ownership, and rhetorical presence in academic writing. Unlike prior work, which focused on detecting AI-generated texts, this appears to be the first literature review to synthesize recent studies specifically through the lens of voice and stylistic individuality of academic texts, particularly in higher education contexts.

Findings revealed that while AI-generated texts are grammatically sound and lexically diverse, they often lack the emotional depth, rhetorical nuance, and personal engagement found in student-authored writing, resulting in a rigid and formulaic tone and structure. This has significant pedagogical implications. Educators must move beyond emphasizing only formal correctness and instead adopt teaching approaches that help students cultivate a distinctive writing voice. Effective strategies include encouraging comparative analysis of AI- and human-generated texts, embedding opportunities for personal narratives and reflective writing within the curriculum, and promoting ethical AI use that preserves individuality and originality. Most compellingly, it is suggested that writing assessment rubrics should explicitly value and reward the integration of personal stories,

critical reflection, authentic voice, and a purposeful use of discursive features. This can ensure that students' unique rhetorical identities remain central to academic writing in the age of Generative AI.

Despite its contributions, this review has some limitations that should be acknowledged. First, the reviewed research focused solely on higher education; thus, its insights may not apply to other educational levels. Secondly, as the field is rapidly evolving, some findings may quickly become outdated due to the continuous advancement of Generative AI tools and their increasing resemblance to human-generated training data. These limitations point to the need for ongoing, context-sensitive research to deepen insights into how best to preserve students' authorial voice and rhetorical individuality in AI-supported learning environments.

Future research should examine how long-term exposure to AI tools affects voice development, explore voice expression across different genres, and investigate student perceptions of authorship in AI-assisted writing. This review lays the groundwork for a much-needed pedagogical shift, one that places the human writer back at the center of academic authorship.

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 author used ChatGPT in order to enhance organization and readability. After using ChatGPT, the author reviewed and edited the content as needed and takes full responsibility for the content of the publication.

About the Author:

Jokha Al Hosni is a senior instructor at the Centre for Preparatory Studies at Sultan Qaboos University, Oman. She holds an MS in Curriculum and Instruction from Purdue University, USA. She has been teaching English foundation and ESP courses for 15 years. Her research interests include curriculum design, language assessment, social and emotional learning, and competency-based learning.
ORCID ID: https://orcid.org/0000-0002-2148-4572

References

AlAfnan, M. A., & MohdZuki, S. F. (2023). Do Artificial Intelligence Chatbots Have a Writing Style? An Investigation into the Stylistic Features of ChatGPT-4. *Journal of Artificial Intelligence and Technology*, *3*(3), 85–94. https://doi.org/10.37965/jait.2023.0267

Ali, A. (2024). Locating the Intersection of Generative Artificial Intelligence and Human English Writing Skills: A Comparative Study. *Arab World English Journal (AWEJ)* Special Issue on ChatGPT, April 2024: 112-123. DOI: https://dx.doi.org/10.24093/awej/ChatGPT.7

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

Amirjalili, F., Neysani, M., & Nikbakht, A. (2024). Exploring the boundaries of authorship: A comparative analysis of AI-generated text and human academic writing in English literature. *Frontiers in Education, 9*,1347421. https://www.frontiersin.org/journals/education/articles/10.3389/feduc.2024.1347421/full

Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing, 57*, 100745. https://doi.org/10.1016/j.asw.2023.100745

Baskara, F. R. (2023). Integrating ChatGPT into EFL writing instruction: Benefits and challenges. *International Journal of Education and Learning, 5(*1), 44-55. https://pubs2.ascee.org/index.php/ijele/article/view/858/0

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. John Benjamins Publishing.

Biber, D. (1988). *Variation across Speech and Writing*. Cambridge: Cambridge University Press.

Charmaz, K., & Mitchell, R. G. (1996). The Myth of Silent Authorship: Self, Substance, and Style in Ethnographic Writing. *Symbolic Interaction, 19*(4), 285–302. https://doi.org/10.1525/si.1996.19.4.285

Digital Education Council. (2024). *How students use AI: The evolving relationship between AI and higher education* [Blog post]. Digital Education Council. https://www.digitaleducationcouncil.com/post/how-students-use-ai-the-evolving-relationship-between-ai-and-higher-education

Due, J., & Daniel, B. (2024). A Systematic Review of AI-Powered Chatbots in EFL Speaking Practice: Transforming Language Education. *Computers and Education: Artificial Intelligence*,6, 100230. DOI: https://doi.org/10.1016/j.caeai.2024.100230

Duong, T., & Suppasetseree, S. (2024). The Effects of an Artificial Intelligence Voice Chatbot on Improving Vietnamese Undergraduate Students' English Speaking Skills. International *Journal of Learning, Teaching and Educational Research,23*(3), 293-321. https://www.ijlter.org/index.php/ijlter/article/view/9787

Durak, H., Eğin, F., & Onan, A. (2025). A Comparison of Human-Written Versus AI-Generated Text in Discussions at Educational Settings: Investigating Features for ChatGPT, Gemini, and BingAI. *European Journal of Education, 60*(1), e70014. https://onlinelibrary.wiley.com/doi/full/10.1111/ejed.70014

Fedoriv, Y., Shuhai, A., & Pirozhenko, I. (2023). Linguo-cognitive markers in human vs. AI text attribution: A case study of narrative and descriptive discourse. *Актуальні питання гуманітарних наук, 66*(3), 130–142. https://doi.org/10.24919/2308-4863/66-3-20

Harunasari, S. Y. (2023). Examining the effectiveness of AI-integrated approach in EFL writing: A case of ChatGPT. *International Journal of Progressive Sciences and Technology (IJPSAT), 39*(2), 357-368. https://ijpsat.org/index.php/ijpsat/article/view/5516

Hawanti, S., & Zubaydulloevna, K. M. (2023). AI chatbot-based learning: alleviating students' anxiety in English writing classroom. *Bulletin of Social Informatics Theory and Application,7*(2), 182-192. DOI:10.31763/businta.v7i2.659

Herbold, S., Hautli-Janisz, A., Heuer, U., Kikteva, Z., & Trautsch, A. (2023). A large-scale comparison of human-written versus ChatGPT-generated essays. *Scientific reports, 13*(1), 18617. https://www.nature.com/articles/s41598-023-45644-9

Hyland, K. (2018). *Metadiscourse: Exploring interaction in writing* (2nd ed.). Bloomsbury Academic.

Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology, 15*(4), ep464. https://www.cedtech.net/article/analyzing-the-role-of-chatgpt-as-a-writing-assistant-at-higher-education-level-a-systematic-review-13605

Ivanič, R. (1998). *Writing and identity: The discoursal construction of identity in academic writing*. John Benjamins Publishing.

Ivanič, R. (1994). I is for interpersonal: Discoursal construction of writer identity and the teaching of writing. *Linguistics and Education, 6*(1), 3–15. https://www.sciencedirect.com/science/article/abs/pii/0898589894900183

Jiang, F. K., & Hyland, K. (2025). Does ChatGPT write like a student? Engagement markers in argumentative essays. *Written Communication*. Advance online publication. https://doi.org/10.1177/07410883251328311

Johansson, I. R. (2023). A Tale of Two Texts, a Robot, and Authorship: a comparison between a human-written and a ChatGPT-generated text [BA dissertation, Malmö University]. https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1767661&dswid=-6519

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, *103*, 102274. https://www.sciencedirect.com/science/article/pii/S1041608023000195

Khasawneh, M. A. S. (2023). Integrating Ai-Based Virtual Conversation Partners in Enhancing Speaking Skills In Foreign Languages: Insights From University Students. *Journal of Southwest Jiaotong University*, *58*(5). https://jsju.org/index.php/journal/article/view/1861

Kwon, S. K., Shin, D., & Lee, Y. (2023). The application of chatbot as an L2 writing practice tool. *Language Learning & Technology*, *27*(1), 1–19. https://doi.org/10125/73541

Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S., & Wang, Z. (2023). ChatGPT and a new academic reality: Artificial Intelligence- written research papers and the ethics of the large language models in scholarly publishing. *Journal of the Association for Information Science and Technology*, *74*(5), 570-581. https://asistdl.onlinelibrary.wiley.com/doi/10.1002/asi.24750

Mahyoob, M., Al-Garaady, J., & Alblwi, A. (2023). A proposed framework for human-like language processing of ChatGPT in academic writing. *International Journal of Emerging Technologies in Learning (iJET)*, *18*(14). https://online-journals.org/index.php/i-jet/article/view/41725

Malik, A. R., Pratiwi, Y., Andajani, K., Numertayasa, I. W., Suharti, S., & Darwis, A. (2023). Exploring artificial intelligence in academic essay: higher education student's perspective. *International Journal of Educational Research Open*, *5*, 100296. https://www.sciencedirect.com/science/article/pii/S2666374023000717

Markey, B., Brown, D. W., Laudenbach, M., & Kohler, A. (2024). Dense and disconnected: Analyzing the sedimented style of ChatGPT-generated text at scale. *Written Communication*, *41*(4), 571-600. https://journals.sagepub.com/doi/10.1177/07410883241263528

Matsuda, P. K., & Tardy, C. M. (2007). Voice in academic writing: The rhetorical construction of author identity in blind manuscript review. *English for Specific Purposes*, *26*(2), 235–249. https://doi.org/10.1016/j.esp.2006.10.001

Mizumoto, A., Yasuda, S., & Tamura, Y. (2024). Identifying ChatGPT-generated texts in EFL students' writing: Through comparative analysis of linguistic fingerprints. *Applied Corpus Linguistics*, *4*(3), 100106. https://www.sciencedirect.com/science/article/pii/S2666799124000236

Nañola, E. L., Arroyo, R. L., Hermosura, N. J. T., Ragil, M., Sabanal, J. N. U., & Mendoza, H. B. (2025). Recognizing the artificial: A comparative voice analysis of AI-generated and L2 undergraduate student-authored academic essays. *System, 130,* 103611. https://www.sciencedirect.com/science/article/pii/S0346251X25000211

Oritsegbemi, O. (2023). Human Intelligence versus AI: Implications for Emotional Aspects of Human Communication. Journal of Advanced Research in Social Sciences, 6(2), 76-85. https://pdfs.semanticscholar.org/38be/7053382f39b7dd65a77cf3077a65b8a48875.pdf

Özçelik, N., & Ekşi, G. (2024). Cultivating writing skills: the role of ChatGPT as a learning assistant—a case study. *Smart Learning Environments, 11*(1), 10. https://slejournal.springeropen.com/articles/10.1186/s40561-024-00296-8

Rusmiyanto, R., Huriati, N., Fitriani, N., Tyas, N. K., Rofi'i, A., & Sari, M. N. (2023). The role of artificial intelligence (AI) in developing English language learner's communication skills. *Journal on Education, 6*(1), 750-757. https://jonedu.org/index.php/joe/article/view/2990

Sardinha, T. B. (2024). AI-generated vs human-authored texts: A multidimensional comparison. *Applied Corpus Linguistics, 4*(1), 100083. https://www.sciencedirect.com/science/article/pii/S2666799123000436

Silitonga, L. M., Hawanti, S., Aziez, F., Furqon, M., Zain, D. S. M., Anjarani, S., & Wu, T. T. (2023). The impact of AI chatbot-based learning on students' motivation in English writing classroom. In: Huang, YM., Rocha, T. (Eds.), *Innovative Technologies and Learning* (pp.542-549). Cham: Springer Nature Switzerland. https://link.springer.com/chapter/10.1007/978-3-031-40113-8_53

Songsiengchai, S., Sereerat, B. O., & Watananimitgul, W. (2023). Leveraging Artificial Intelligence (AI): Chat GPT for Effective English Language Learning among Thai Students. *English Language Teaching, 16*(11), 1-68. https://ccsenet.org/journal/index.php/elt/article/view/0/49421

Taecharungroj, V. (2023). "What can ChatGPT do?" Analyzing early reactions to the innovative AI chatbot on Twitter. *Big Data and Cognitive Computing, 7*(1), 35. https://www.mdpi.com/2504-2289/7/1/35

Tannen, D. (1993). *Framing in Discourse*. Oxford University Press.

Tardy, C. M. (2012). Voice construction, assessment, and extra-textual identity. *Research in the Teaching of English, 47*(1), 64-99. https://publicationsncte.org/content/journals/10.58680/rte201220672

Wang, S., Cristianini, N., & Hood, B. M. (2024). Stylometric Comparison between ChatGPT and Human Essays. In *Workshop Proceedings of the 18th International AAAI Conference on Web and Social Media*. https://doi.org/10.36190/2024.32

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

Yuan, Y., Li, H., & Sawaengdist, A. (2024). The impact of ChatGPT on learners in English academic writing: opportunities and challenges in education. *Language Learning in Higher Education, 14*(1), 41-56. https://www.degruyterbrill.com/document/doi/10.1515/cercles-2023-0006/html?lang=en

Zhang, Z., & Huang, X. (2024). The impact of chatbots based on large language models on second language vocabulary acquisition. *Heliyon, 10*(3). https://www.sciencedirect.com/science/article/pii/S2405844024014014

Zindela, N. (2023). Comparing Measures of Syntactic and Lexical Complexity in Artificial Intelligence and L2 Human-Generated Argumentative Essays. *International Journal of Education and Development Using Information and Communication Technology*, *19*(3), 50-68. https://eric.ed.gov/?id=EJ1413411

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.259-269
DOI: https://dx.doi.org/10.24093/awej/vol16no3.15

The Trends in the Integration of AI Technologies in English Language Teaching: Systematic Literature Review

Wani Nurfahani Mohd Sapuan
Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia
Corresponding Author: p131169@siswa.ukm.edu.my

Nur Ainil Sulaiman
Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia
Email: nurainil@ukm.edu.my

Maslawati Mohamad
Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia
Email: maslawati@ukm.edu.my

Received: 06/24/2025 Accepted:08/16/2025 Published: 09/20/2025

Abstract
This systematic literature review (SLR) focuses on AI, which is generally a new topic in education, with limited reviews on its use and roles, especially in ELT practices. Three research questions guided this review, which are: 1) How have AI technologies in ELT been defined and conceptualised in recent years? 2) What role does AI play in supporting ELT in recent years? 3) What strategies can be implemented to ensure the responsible and ethical use of AI technologies in ELT? The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) coding scheme, which is an established systematic review framework, was utilised as a guideline in selecting the articles for the review. Three phases were involved: identification, screening, and inclusion. As a result, the review included twenty-four out of thirty-two papers from Scopus and Web of Science (WOS) for analysis. The findings reveal that AI has progressed from just an intelligent system to a practical technology in the ELT field. It plays an essential role in supporting ELT practices through available tools and applications for both educators and learners. Lastly, the findings indicate that it is crucial to use AI in a principled manner, develop systematic guidelines and ethical frameworks, and conduct further research on AI in ELT to ensure proper use of AI among educators. Overall, this comprehensive review may serve as a valuable reference for future research in understanding the effectiveness of AI technologies in ELT practices.

Keywords: Artificial Intelligence (AI), English Language Teaching (ELT), PRISMA coding scheme, Systematic Literature Review (SLR)

How to cite: Sapuan, W.N.M., Sulaiman, N. A., & Mohamad, M. (2025). The Trends in the Integration of AI Technologies in English Language Teaching: Systematic Literature Review. *Arab World English Journal, 16*(3):259-269. https://dx.doi.org/10.24093/awej/vol16no3.15

Introduction

The ELT landscape has transformed tremendously due to the pandemic. The pandemic has shifted traditional face-to-face learning to online learning. Then, it transformed learning into blended learning in the post-pandemic era. These changes have led to the integration of technology in ELT to ensure that the teaching and learning process is not affected by the pandemic. Furthermore, there is a notable trend towards student-centred approaches, highlighting task-based learning and communicative language teaching, which emphasise practical communication skills over traditional teaching methods (Valledor et al., 2023). This phenomenon proves how the pandemic led to the integration of technology in education at a global scale (Sharadgah & Sa'di, 2022). According to Huang et al. (2023), AI is currently the latest technology in the education field. The flexible AI tools and applications support the ever-changing landscape in the ELT field (Alhalangy & AbdAlgane, 2023). Thus, the AI integration in the ELT practices is relevant as Crompton et al. (2024) state that AI may positively change English education and enhance both teaching quality and learning experiences.

Previous systematic reviews have generally focused on AI integration in education (Crompton et al., 2024). Lack of recent reviews of AI integration into ELT despite its rapid development in education (Sharadgah & Sa'di, 2022). Hence, this review highlights the rise of AI technologies in the ELT field, so both educators and learners are aware of what they can harvest from this new technological development. The exposure impacts not only their teaching and learning experience but also their lives in general. According to Kucuk (2024), "The functionality and practicality of artificial intelligence, which is in every aspect of people's lives, are among the factors that make it the most preferred" (p. 101). This review aims to examine AI integration into ELT from 2020 to 2024, focusing on the three research questions: 1) How have AI technologies in ELT been defined and conceptualised in recent years? 2) What role does AI play in supporting ELT in recent years? 3) What strategies can be implemented to ensure the responsible and ethical use of AI technologies in ELT?

Literature Review

AI Technologies in ELT

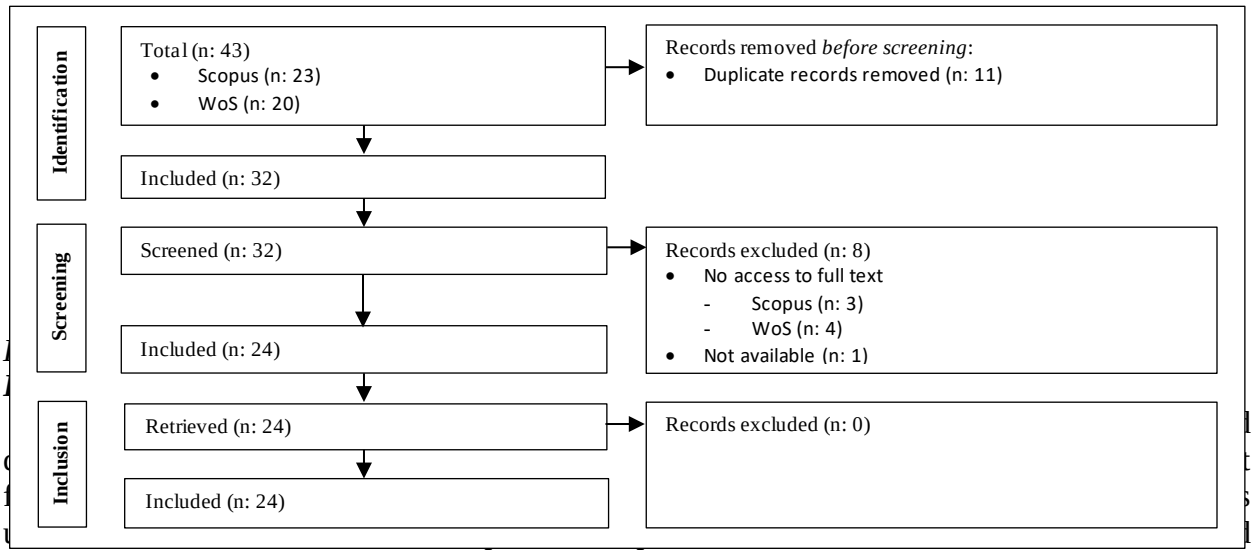
The advancement of information technology and educational concepts has shed some light on the ELT field (Yuan et al., 2023). Currently, AI is the latest technology in education (Huang et al., 2023). It has changed how students learn English by offering personalised learning paths, immediate feedback, and adaptive language activities tailored to students' levels (Binu, 2024). Furthermore, the emergence of AI language learning applications also provides interactive practice in real-time situations, reflecting more efficient ELT methods (Meniado, 2023). Al Mukhallafi (2020) highlighted that AI could improve the ELT field in the areas of instructional and assessment methods due to its personalised learning approach. Overall, the impact of AI technologies is evident in the ELT field (Sharadgah & Sa'di, 2022). The integration of AI technologies in education has removed traditional concepts and introduced new perspectives for future education (Kucuk, 2024). Their contribution is widely acknowledged in the education setting, particularly in the ELT field (Delgado et al., 2020). Therefore, English lecturers should incorporate AI technologies to teach the current generation effectively (Slimi et al., 2022).

AI Tools and Applications in ELT

AI tools have evolved and incorporated advanced data processes like data perception, analysis, and utilisation (Antonenko & Abramowitz, 2022). The evolution in AI-driven language tools leads to the transformation of traditional teaching methods by enabling personalised and adaptive instruction (Govindarajan & Christuraj, 2023). Among AI-based applications that have been used widely in the ELT field are AI-based writing tools, machine translation, generative AI, Naver Papago, Google Translate, and Grammarly (Lee et al., 2024b). Furthermore, there is ChatGPT, which is always accessible and can offer alternatives and solutions based on specific needs (Kucuk, 2024). In China, the Liulishuo application has been used for speech synthesis, pronunciation evaluation, speech recognition, and syntax error detection (Chang et al., 2022). Learners can improve their pronunciation issues after using the application for two months. Owoc et al. (2021) highlighted that Nuance's speech recognition software can assist educators to prepare assignments, organise tasks, and manage communications through notifications or emails. They also mentioned Coggni, a conversational AI software that uses virtual assistants for providing immediate feedback and personalised tutoring, aligns with students' needs. Both applications are practical as language learning aids for students and educators. Thus, staying updated with the advancement of technology in education is necessary, consistent with the growing number of AI tools and applications (Kucuk, 2024).

Methodology

The study used the PRISMA coding scheme, an established set of rules for performing systematic reviews (Shadiev & Yang, 2020). The rules ensure all the articles meet the criteria for inclusion and are free from the requirements for removal. The article selection process, following the PRISMA coding scheme, was divided into three phases: identification, screening, and inclusion (refer to Figure 1). The Mendeley software assisted the selection process.



with the keywords. Unfortunately, 11 duplicates have been excluded from the article list. Table 1 shows the identification of articles during this phase.

Table 1. *Article identification*

Abstracting and Indexing Databases.	Total Search	Duplicate	Total
Scopus	23	11	32
World of Science (WOS)	20		

Phase 2. Screening

After the identification process, there was a screening process to check if the articles met the standards for inclusion. As a result, the process removed eight articles due to no access to the full text (Scopus three; WOS four), and one article was unavailable (Refer to Table 2). The other 24 articles were accepted as they were relevant to both AI and ELT areas. For the resources list, refer to Appendix A.

Table 2. *Screening*

Search Engine	**Total Search**	**No access to the full text**	**Not available**	**Total**
Scopus	23	3	0	24
World of Science (WOS)	9	4	1	

Phase 3. Inclusion

Following the screening process, further reviews of the articles were conducted to find related research that fits the study's scope. This systematic literature review followed a clear set of guidelines that includes articles and conference papers that: 1) are relevant to the topic of this research and published in English, 2) revolve around both local and international settings, 3) specifically address the application of AI technology in the ELT field, 4) were published over five years, ranging from 2020 to 2024.

Results

How AI Technologies in ELT Have Been Defined and Conceptualised in Recent Years?

Based on the analysis, AI technologies have evolved significantly in recent years, encompassing changes in their definitions and concepts. According to Hockly (2023), AI refers to "a system's ability to correctly interpret external data, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation" (p. 2). Sharadgah and Sa'di (2022), who cited Bin and Mandal (2019), highlighted that AI is a discipline that integrates some aspects of computer science, cybernetics, linguistics, and others. Kusuma et al. (2024) cited Antonenko and Abramowitz (2022), who stated that the development of AI technologies includes advanced data processes such as data perception, analysis, and utilisation. Binu (2024) described AI as a technology application that involves natural language processing and machine learning for delivering personalised instruction and feedback. Delgado et al. (2020) defined AI as a system designed to exhibit human-like intelligent behaviours.

The conceptualisation of AI technologies in the ELT field has significantly progressed due to its potential to change the educational landscape. For instance, Crompton et al. (2024) noted the acknowledgement of AI at the tertiary level. It demonstrates that AI has caught the attention in academic contexts. This observation aligns with Tang and Foley (2022), who highlight the significant need for expertise in both AI and English due to technological development and global reach. Then, Sharadgah and Sa'di (2022) stated that AI has expanded its influence across sectors, including education. The expansion highlights its progression in the education field. Next, Yuan et al. (2023) pointed out that AI and 5G networks have supported innovations in college English

classes recently. This progress fosters new approaches to language teaching. Furthermore, Rahimi and Sevilla-Pavón (2024) highlighted that AI has encouraged innovations in educational practices, especially in programming and teacher training areas. Finally, Alhalangy and AbdAlgane (2023) stated that AI-assisted technologies have the potential to make education more flexible and comprehensive, while also promoting the development of relevant skills needed in today's world.

What Role Does AI Play in Supporting ELT in Recent Years?

The findings reveal that the roles of AI technologies in the ELT field in recent years have primarily focused on supporting both English language teaching and learning. AI can improve language skills and maintain student engagement in education (Alhalangy & AbdAlgane, 2023). Moreover, AI-driven technologies save educators' time in preparing teaching materials such as worksheets, exercises, and activities in language classrooms (Govindarajan & Christuraj, 2023). In addition, Delgado et al. (2020) emphasised the inclusivity of AI-based applications in education, which could cater to learners' different needs and ensure accessible and equal educational opportunities to all language learners. Alhalangy and AbdAlgane (2023) highlighted that the flexibility of the AI tools in education to support the language teaching experience was agreed upon by the language experts.

AI technologies have been developed for their roles in supporting language learning in various ways (Valledor et al., 2023), including supporting language comprehension among learners (Al-khresheh, 2024; Mabuan, 2024; Özçelik & Ekşi, 2024) due to their flexibility (Alhalangy & AbdAlgane, 2023; Delgado et al., 2020), and capacities to assist learners' language development (Mabuan, 2024). These technologies facilitate error checking (Lee et al., 2024b; Özçelik & Ekşi, 2024) and language translation (Govindarajan & Christuraj, 2023; Mabuan, 2024), which makes these technologies preferable as a language learning tool (Alhalangy & AbdAlgane, 2023; Özçelik & Ekşi, 2024). Furthermore, they assist learners in writing tasks, provide corrective feedback, perform grammar checks, and promote self-learning (Binu, 2024), alongside assisting with homework completion (Govindarajan & Christuraj, 2023; Özçelik & Ekşi, 2024). AI technologies also provide language practices across all skills, including grammar (Binu, 2024; Kucuk, 2024), reading (Binu, 2024; Govindarajan & Christuraj, 2023), writing (Alhalangy & AbdAlgane, 2023; Binu, 2024; Mabuan, 2024; Özçelik & Ekşi, 2024), listening (Govindarajan & Christuraj, 2023), and speaking (Alhalangy & AbdAlgane, 2023; Ironsi, 2023). Moreover, they offer personalised learning to learners (Al-khresheh, 2024; Kusuma et al., 2024; Özçelik & Ekşi, 2024; Yuan et al., 2023). Feature such as two-way communication provided by AI applications offers conversational tools and interactive learning activities to learners (Hung et al., 2023; Kucuk, 2024). AI chatbot activities increased learners' confidence to communicate in English in an authentic setting (Lee et al., 2024a).

What Strategies Can Be Implemented to Ensure the Responsible and Ethical Use of AI Technologies in ELT?

The first strategy found in past literature is using AI technologies in a reasonable manner, full of principles, especially in the post-pandemic era, where these technologies are still developing and becoming mainstream to the public (Hockly, 2023). Kohnke et al. (2023a) highlighted the same idea, recommending that English educators use AI technologies in a principled manner to enhance language learning. The second strategy is to develop guidelines and ethical frameworks that ensure data privacy and transparency throughout the process (Mabuan, 2024). Kohnke et al.

(2023a) suggested that all institutions to develop appropriate guidelines for using AI tools in teaching and assessment practices. Additionally, Kohnke et al. (2023b) shared several tips for English educators to stay updated with AI technologies. The suggestions include attending professional training, treating AI as a supporting tool in teaching and learning, establishing guidelines and policies for data privacy and ethical use, and upgrading access and infrastructure.

Discussion

The analysis reveals that AI in ELT has made significant progress by combining technological advancements with pedagogical innovations in education. This progression demonstrates that AI is no longer a theoretical construct, but it is now a practical application in teaching and learning. This development aligns with Sharadgah and Sa'di (2022), who note that AI technologies have gained their place in ELT. Furthermore, they are also acknowledged in education widely (Delgado et al., 2020).

Moreover, the analysis shows that AI technologies play a crucial role in supporting ELT across all skills with their available tools and applications. Educators could save time in preparing materials and delivering their teaching practice. Learners can experience personalised and interactive language learning. This would enhance students' engagement in the classroom. The analysis parallels that of Valledor et al. (2023), as they state that AI technologies support language learning. Thus, it explains why it is a must to stay updated with the technology development in education, as highlighted by Kucuk (2024).

Additionally, the analysis highlights that educators should use AI technologies properly. Thereby, they should attend training to get themselves well prepared with all the required skills to seamlessly integrate AI into their language classroom. It is relevant for educators to get professional help in paving the way to learn how to use AI technologies effectively in their teaching practices. This recommendation is relevant, as Huang et al. (2023) note that AI technologies are still new in the education field. This effort is also consistent with Slimi et al. (2022), who state that teaching approaches should meet learners' needs and accommodate the current technological development. Other than that, the analysis also indicates that developing guidelines and frameworks is vital in ensuring the ethical use of AI among educators. Since education is a highly regulated industry, it is relevant for these documents to be in place as a source of reference and guidance for educators.

Conclusion

In conclusion, past literature on AI integration in the ELT field reveals significant insights across three key research questions. The first research question focused on how AI technologies in ELT have been defined and conceptualised in recent years. Recent literature shows significant developments in the conceptualisation and use of AI technologies in the ELT field. AI has progressed and transformed traditional teaching methods into more effective ELT practices. The second research question addressed the roles of AI technologies in supporting ELT practices. Previous literature indicates that the attributes of AI technologies in the ELT field mainly emphasise their abilities to support many aspects of ELT and language learning. The final research question centred on suggestions for educators to use AI technologies in their ELT practices. Past studies suggested that English educators use AI in a principled manner, develop appropriate guidelines and ethical frameworks, and stay current with AI developments.

Future studies may expand the scope of the literature review by incorporating more recent and diverse sources related to AI integration in the ELT field. It could enhance the depth of the analysis. Although this study offers a comprehensive overview of AI in ELT, future research may include various perspectives on its ongoing development, challenges, and impacts in educational contexts. Such findings would provide additional insights into how AI technologies are applied and perceived globally. This expansion would also help identify any new trends or gaps in the present research. Furthermore, gathering empirical data from surveys, interviews, or case studies involving educators and students could provide real-world insights into the use of AI tools in the ELT settings. Further research would reveal specific issues experienced by users. In addition, evaluating the long-term impact of integrating AI into the ELT field could provide a better understanding of the effectiveness of AI-based technologies. Although the present study already highlights the use of AI technologies in the ELT field, a longitudinal study assesses how AI tools affect language learning outcomes over time. It could help improve AI-driven educational technologies and guide their future 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 the article, the author used ChatGPT and Grammarly to improve readability and language. The author is responsible for the content of the publication.

About the Authors:

Wani Nurfahani Mohd Sapuan is a PhD candidate at the Faculty of Education (TESL), UKM. She can be contacted via email at p131169@siswa.ukm.edu.my. Her research interests revolve around English Language Teaching (ELT) and educational technology areas. ORCID: https://orcid.org/0009-0006-9055-5457

Dr. Nur Ainil Sulaiman is a lecturer at the Faculty of Education, UKM. She can be contacted via email at nurainil@ukm.edu.my. Her research interests are Second Language Acquisition (SLA), vocabulary learning, and Sociolinguistics. ORCID: https://orcid.org/0000-0001-6212-7494

Dr. Maslawati Mohamad is a senior lecturer at the Faculty of Education, UKM. She can be contacted via email at maslawati@ukm.edu.my. Her research areas include Teaching English as a Second Language, and innovations in teaching and learning. ORCID: https://orcid.org/0000-0002-8979-58950000-0002-8979-5895

References

Al-khresheh, M. H. (2024). Bridging Technology and Pedagogy from a Global Lens: Teachers' Perspectives on Integrating ChatGPT in English Language Teaching. *Computers and Education: Artificial Intelligence, 6*(2024), 1-12. https://doi.org/10.1016/j.caeai.2024.100218

Al Mukhallafi, T. R. (2020). Using Artificial Intelligence for Developing English Language Teaching/Learning: An Analytical Study from University Students' Perspective.

International Journal of English Linguistics, 10(6), 40-53. https://doi.org/10.5539/ijel.v10n6p40

Alhalangy, A. G. I., & AbdAlgane, M. (2023). Exploring the Impact of AI on the EFL Context: A Case Study of Saudi Universities. *Journal of Intercultural Communication*, *23*(2), 41–49. https://doi.org/10.36923/jicc.v23i2.125

Antonenko, P., & Abramowitz, B. (2022). In-service Teachers' (Mis)conceptions of Artificial Intelligence in K-12 Science Education. *Journal of Research on Technology in Education*, *55*(1), 64-78. https://doi.org/10.1080/15391523.2022.2119450

Binu, P. M. (2024). Affordances and Challenges of Integrating Artificial Intelligence into English Language Education: A Critical Analysis. *English Scholars Beyond Borders*, *10*(1), 34-51. https://www.englishscholarsbeyondborders.org/wp-content/uploads/2024/06/Binu-1-1.pdf

Chang, Y., Lee, S., Wong, S. F., & Jeong, S.-P. (2022). AI-Powered Learning Application Use and Gratification: An Integrative Model. *Information Technology & People (2022), 35*(7), 2115-2139. https://doi.org/10.1108/ITP-09-2020-0632

Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English Language Teaching: Affordances and Challenges. *British Journal of Educational Technology, 55*(6), 2503-2529. https://doi.org/10.1111/bjet.13460

Delgado, H. O. K., Fay, A. de A., Sebastiany, M. J., & Silva, A. D. C. (2020). Artificial Intelligence Adaptive Learning Tools: The Teaching of English in Focus. *BELT Journal*, *11*(2). https://doi.org/10.15448/2178-3640.2020.2.38749

Govindarajan, R., & Christuraj, G. (2023). Opportunities and Challenges of Using ChatGPT in the ELT Scenario of Utas, Nizwa, Oman. *Journal of Teaching English for Specific and Academic Purposes*, *11*(3), 593-605. https://doi.org/10.22190/JTESAP230529046G

Hockly, N. (2023). Artificial Intelligence in English Language Teaching: The Good, the Bad and the Ugly. *RELC Journal, 54*(2), 445–451. https://doi.org/10.1177/00336882231168504

Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, Research Issues and Applications of Artificial Intelligence in Language Education. *Technology & Society, 26*(1), 112-131. https://doi.org/10.30191/ETS.202301_26(1).0009

Hung, C.-Y., Lin, Y.-T., Yu, S.-J., & Sun, J. C.-Y. (2023). Effects of AR- and VR-based Wearables in Teaching English: The Application of an ARCS Model-Based Learning Design to Improve Elementary School Students' Learning Motivation and Performance. *Journal of Computer Assisted Learning, 39*(5), 1510–1527. https://doi.org/10.1111/jcal.12814

Ironsi, C. S. (2023). Investigating the Use of Virtual Reality to Improve Speaking Skills: Insights from Students and Teachers. *Smart Learning Environments*, *10*(53). https://doi.org/10.1186/s40561-023-00272-8

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023a). ChatGPT for Language Teaching and Learning. *RELC Journal, 54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023b). Exploring Generative Artificial Intelligence Preparedness among University Language Instructors: A Case Study. *Computers and Education: Artificial Intelligence, 5*, 1-8. https://doi.org/10.1016/j.caeai.2023.100156

Kucuk, T. (2024). ChatGPT Integrated Grammar Teaching and Learning in EFL Classes: A Study on Tishk International University Students in Erbil, Iraq. *Arab World English*

Journal Special Issue on ChatGPT, 100–111. https://dx.doi.org/10.24093/awej/ChatGPT.6

Kusuma, I. P. I., Roni, M., Dewi, K. S., & Mahendrayana, G. (2024). Revealing the Potential of ChatGPT for English Language Teaching: EFL Preservice Teachers' Teaching Practicum Experience. *Studies in English Language and Education*, *11*(2), 650–670. https://doi.org/10.24815/siele.v11i2.34748

Lee, S., Jeon, J., & Choe, H. (2024a). Enhancing Pre-Service Teachers' Global Englishes Awareness with Technology: A Focus on AI Chatbots in Metaverse Environments. *TESOL Quarterly*, *59*(1), 49-74. https://doi.org/10.1002/tesq.3300

Lee, Y.-J., Davis, R. O., & Lee, S. O. (2024b). University Students' Perceptions of Artificial Intelligence-Based Tools for English Writing Courses. *Online Journal of Communication and Media Technologies*, *14*(1), 1-11. https://doi.org/10.30935/ojcmt/14195

Mabuan, R. A. (2024). ChatGPT and ELT: Exploring Teachers' Voices. *International Journal of Technology in Education*, *7*(1), 128–153. https://doi.org/10.46328/ijte.523

Meniado, J. (2023). The Impact of ChatGPT on English Language Teaching, Learning, and Assessment: A Rapid Review of Literature. *Arab World English Journal*, *14*(4), 3–18. https://dx.doi.org/10.24093/awej/vol14no4.1

Owoc, M. L., Sawicka, A. & Weichbroth, P. (2021). Artificial Intelligence Technologies in Education: Benefits, Challenges and Strategies of Implementation. In M. L. Owoc, A. Sawicka, & P. Weichbroth (Eds.), *Artificial Intelligence for Knowledge Management (AI4KM 2019)* (pp. 37–58). Springer. https://dx.doi.org/10.1007/978-3-030-85001-2_4

Özçelik, N. P., & Ekşi, G. Y. (2024). Cultivating Writing Skills: The Role of ChatGPT as a Learning Assistant-A Case Study. *Smart Learning Environments*, *11*(10), 1-18. https://doi.org/10.1186/s40561-024-00296-8

Rahimi, A. R., & Sevilla-Pavón, A. (2024). The Role of ChatGPT Readiness in Shaping Language Teachers' Language Teaching Innovation and Meeting Accountability: A Bisymmetric Approach. *Computers and Education: Artificial Intelligence*, *7*, 1-19. https://doi.org/10.1016/j.caeai.2024.100258

Shadiev, R., & Yang, M. (2020). Review of Studies on Technology-Enhanced Language Learning and Teaching. *Sustainability 2020*, *12*(2), 1-22. https://doi.org/10.3390/su12020524

Sharadgah, T. A., & Sa'di, R. A. (2022). A Systematic Review of Research on the Use of Artificial Intelligence in English Language Teaching and Learning (2015-2021): What are the Current Effects? *Journal of Information Technology Education: Research, 21,* 337–377. https://doi.org/10.28945/4999

Slimi, Z., Al Alawai, F., Al Alyani, H., Al Abri, S., Al-Farsi, F. A., & Al Balushi, K. (2022). Writing Issues in ESL and their Potential Solutions: Case Study IMCO's Foundation Students. *Journal of Educational and Social Research*, *12*(6), 81. https://doi.org/10.36941/jesr-2022-0146

Tang, J., & Foley, J. (2022). A Case Study on the Effectiveness of Applying Content and Language Integrated Learning in an Artificial Intelligence English Reading Course. *Arab World English Journal*, *13*(3), 236–253. https://doi.org/10.24093/awej/vol13no3.15

Valledor, A., Olmedo, A., Hellín, C. J., Tayebi, A., Otón-Tortosa, S., & Gómez, J. (2023). The Eclectic Approach in English Language Teaching Applications: A Qualitative Synthesis of the Literature. *Sustainability 2023, 15*(15), 1-31. https://doi.org/10.3390/su151511978

Yuan, J., Liu, Y., Han, X., Li, A., & Zhao, L. (2023). Educational Metaverse: An Exploration and Practice of VR Wisdom Teaching Model in Chinese Open University English Course. *Interactive Technology and Smart Education*, *20*(3), 403–421. https://doi.org/10.1108/ITSE-10-2022-0140

Appendix A
Resources List

Bil	Authors	Year	Title	Source
1	Delgado, H. O. K., Fay, A. de A., Sebastiany, M. J., & Silva, A. D. C.	2020	Artificial Intelligence Adaptive Learning Tools: The Teaching of English in Focus	SCOPUS
2.	Sharadgah, T. A., & Sa'di, R. A.	2022	A Systematic Review of Research on the Use of Artificial Intelligence in English Language Teaching and Learning (2015-2021): What are the Current Effects?	WoS SCOPUS
3.	Tang, J., & Foley, J.	2022	A Case Study on the Effectiveness of Applying Content and Language Integrated Learning in an Artificial Intelligence English Reading Course	WoS
4.	Slimi, Z., Al Alawai, F., Al Alyani, H., Al Abri, S., Al-Farsi, F. A., & Al Balushi, K.	2022	Writing Issues in ESL and their Potential Solutions: Case Study IMCO's Foundation Students	SCOPUS
5.	Alhalangy, A. G. I., & AbdAlgane, M.	2023	Exploring the Impact of AI on The EFL Context: A Case Study of Saudi Universities.	SCOPUS
6.	Ironsi, C. S.	2023	Investigating the Use of Virtual Reality to Improve Speaking Skills: Insights from Students and Teachers	WoS SCOPUS
7.	Yuan, J., Liu, Y., Han, X., Li, A., & Zhao, L.	2023	Educational Metaverse: An Exploration and Practice of VR Wisdom Teaching Model in Chinese Open University English Course	WoS SCOPUS
8.	Hung, C.-Y., Lin, Y.-T., Yu, S.-J., & Sun, J. C.-Y.	2023	Effects of AR- and VR-Based Wearables in Teaching English: The Application of an ARCS Model-Based Learning Design to Improve Elementary School Students' Learning Motivation and Performance	WoS
9.	Valledor, A., Olmedo, A., Hellín, C. J., Tayebi, A., Otón-Tortosa, S., & Gómez, J.	2023	The Eclectic Approach in English Language Teaching Applications: A Qualitative Synthesis of the Literature	SCOPUS
10.	Kohnke, L., Moorhouse, B. L., & Zou, D	2023	Exploring Generative Artificial Intelligence Preparedness among University Language Instructors: A Case Study	SCOPUS
11.	Hockly, N.	2023	Artificial Intelligence in English Language Teaching: The Good, the Bad and the Ugly	WoS SCOPUS
12.	Meniado, J.	2023	The Impact of ChatGPT on English Language Teaching, Learning, and Assessment: A Rapid Review of Literature	WoS

13.	Kohnke, L., Moorhouse, B. L., & Zou, D.	2023	ChatGPT for Language Teaching and Learning	WoS SCOPUS
14.	Govindarajan, R., & Christuraj, G.	2023	Opportunities and Challenges of Using ChatGPT in the ELT Scenario of Utas, Nizwa, Oman	WoS SCOPUS
15.	Crompton, H., Edmett, A., Ichaporia, N., & Burke, D.	2024	AI and English Language Teaching: Affordances and Challenges	WoS SCOPUS
16.	Lee, Y.-J., Davis, R. O., & Lee, S. O.	2024	University Students' Perceptions of Artificial Intelligence-Based Tools for English Writing Courses	WoS SCOPUS
17.	Mabuan, R. A.	2024	ChatGPT and ELT: Exploring Teachers' Voices	WoS
18.	Özçelik, N. P., & Ekşi, G. Y.	2024	Cultivating Writing Skills: The Role of ChatGPT as a Learning Assistant – A Case Study	WoS
19.	Lee, S., Jeon, J., & Choe, H.	2024	Enhancing Pre-Service Teachers' Global Englishes Awareness with Technology: A Focus on AI Chatbots in Metaverse Environments	WoS SCOPUS
20.	Kucuk, T.	2024	ChatGPT Integrated Grammar Teaching and Learning in EFL Classes: A Study on Tishk International University Students in Erbil, Iraq	WoS
21.	Kusuma, I. P. I., Roni, M., Dewi, K. S., & Mahendrayana, G.	2024	Revealing the Potential of ChatGPT for English language Teaching: EFL Preservice Teachers' Teaching Practicum Experience	SCOPUS
22.	Al-khresheh, M. H.	2024	Bridging Technology and Pedagogy from a Global Lens: Teachers' Perspectives on Integrating ChatGPT in English Language Teaching	SCOPUS
23.	Binu, P. M.	2024	Affordances and Challenges of Integrating Artificial Intelligence into English Language Education: A Critical Analysis	SCOPUS
24.	Rahimi, A. R., & Sevilla-Pavón, A.	2024	The Role of ChatGPT Readiness in Shaping Language Teachers' Language Teaching Innovation and Meeting Accountability: A Bisymmetric Approach	SCOPUS

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.270-284
DOI: https://dx.doi.org/10.24093/awej/vol16no3.16

Realizing ESL-oriented Education in the Singaporean Setting: A Scoping Review

Phan Ngoc Tuong Vy
School of Foreign Languages, Can Tho University, Vietnam
Email: vyp2025008@gstudent.ctu.edu.vn

Trinh Quoc Lap
School of Foreign Languages, Can Tho University, Vietnam
Email: tqlap@ctu.edu.vn

Le Cong Tuan
School of Foreign Languages, Can Tho University, Vietnam
Email: lctuan@ctu.edu.vn

Dang Vu Hoai Nhan
School of Foreign Languages, Can Tho University, Vietnam
Email: dvhnhan@ctu.edu.vn

Le Xuan Mai
School of Foreign Languages, Can Tho University, Vietnam
Corresponding Author: lxmai@ctu.edu.vn

Received:06/09/2025 Accepted: 08/06/2025 Published: 09/20/2025

Abstract

As Vietnam contemplates a shift from teaching English as a foreign language to teaching English as a second language, it is essential to examine second language education in linguistically comparable contexts. The current study synthesizes twelve review and empirical studies of English language instruction in Singapore (2019–2025) and uses thematic analysis to identify four interrelated fields: learner knowledge, learner awareness, skill development, and teacher professional development. The study aims to elicit implications that are relevant to Vietnam's changing English language reforms. It addresses the question: What strategies from Singapore's experience can inform Vietnam's transition to second language instruction? Findings have accentuated the importance of culturally responsive teaching, multimodal and socio-phonetic pedagogy, foundational skill alignment across educational stages, and reflective teacher development guided by national language policy. Based on the evidence-based synthesis of insights, the paper hopes to offer practical implications for Vietnam about localized curricula, inclusive teacher training, or policy support for sociolinguistic diversity in the early phase of second language reforms tailored to Vietnam's context.

Keywords: ESL education, awareness, knowledge, multimodal pedagogy, skills, sociolinguistic diversity, teacher professional development

How to cite: Vy, P.N.T., Lap, T.Q., Tuan, L. C., Nhan, D. V. H., & Mai, L. X. (2025). Realizing ESL-oriented Education in the Singaporean Setting: A Scoping Review. *Arab World English Journal, 16*(3): 270-284. https://dx.doi.org/10.24093/awej/vol16no3.16

Introduction

Vietnam is entering a pivotal phase in its national language education reform which is marked by a growing effort to reposition English from a Foreign Language (EFL) to a Second Language (ESL) (Trinh et al., 2025). This shift is more and more pronounced in cities, internationalized studies, and tertiary education, where English is not merely used as a classroom language but also for work and social purposes. To be in step with such changes, English language education must transcend traditional models and embark on a socially placed, context-responsive stance that is attuned to learners' multilingual lives. The 2020–2025 National Foreign Language Project (NFLP 2025) has laid important groundwork for such transformation. Massive projects include developing a six-level CEFR-based English competence framework, building capacity of over 85,000 English teachers, and integrating English into content-based learning at various levels of schools (Nguyen & Duong, 2021). The project has also experimented with English as a Medium of Instruction (EMI) for fostering digital innovation in English Language Teaching (ELT) and attaining international partnerships (Bui, 2022). These partnership efforts once again show Vietnam's desire to have a more efficient and inclusive ELT system. However, practical models of how English can be made into a second language are still underdeveloped, particularly in curriculum design, teacher training, and learner motivation across the country. In envisioning next-phase reforms, Singapore emerges as a particularly instructive reference point. As a post-colonial, multilingual society like Vietnam, Singapore has successfully institutionalized English as a national working language through long-term, policy-driven reforms. Its ESL system is built on four interconnected domains: awareness of the societal and identity functions of English, knowledge of language structures and variation, skills for disciplinary and social communication, and sustained professional development (PD) to support globally competent pedagogy (Jones et al., 2025). As a result, this review aims to examine how Singapore's strategies in English language policy and practice can inform Vietnam's ongoing transition to ESL. By analyzing recent empirical and systematic studies published between 2019 and 2025, the review seeks to identify insights that may support Vietnam in building a more coherent and contextually grounded model for English as a second language through the research question:

What strategies from Singapore's experience can inform Vietnam's transition to second language instruction?

Literature Review

In an effort to facilitate Vietnam's seamless shift from English as a Foreign Language (EFL) to English as a Second Language (ESL), one must understand the basic differences of the two orientations as pedagogies. Sato and Storch (2020) clarified that EFL contexts typically position English as a subject taught within classroom walls, which have the tendency to stay isolated from learners' daily communication needs or societal uses. EFL students may not have as straightforward access to natural English use outside of formal instruction, limiting opportunities for deep learning. ESL environments, on the other hand, envision English as a socially situated language that permeates society, media, schools, and workplaces. Yoo and Kim (2023) go on to say that within ESL settings, English is not just the target of classroom instruction but a medium of everyday interaction and academic discourse as well. As countries around the world have shifted from treating English as a foreign language to positioning it as a second language, international studies have identified a number of recurring challenges. Another common issue is the difficulty happening for Indian related stakeholders in aligning national language education policies with the

multilingual realities of everyday life, particularly in post-colonial or linguistically diverse societies (Singh, 2025). Envisioning the shift from EFL to ESL, Bahanshal (2023) found that one major challenge in the Saudi Arabian context lies in designing curricula that promote the use of English in real-world settings beyond the classroom, while still acknowledging and valuing learners' first languages. Studies from Taiwan and Malaysia highlight shifting demands on teachers. While Taiwanese educators are increasingly expected to facilitate socially embedded, function-oriented language learning (Reynolds et al., 2022), non-specialist ESL teachers in Malaysia often struggle to confidently deliver communicative instruction, particularly in speaking (Zulkepli et al., 2024). Regarding developing nations like Nigeria, there are persistent disparities of access chances between urban and rural areas regarding English-language resources, technological infrastructure, and consistent instructional quality (Sabe & Luka, 2020).

The transformation of ELT calls for sustained, context-sensitive education overhauls. With this in mind, Singapore presents a compelling model since despite its colonial past and linguistic diversity, it has positioned English as a national working language through strategic planning, curriculum reform, teacher development, and public engagement. Innovations include the use of metalanguage for socio-phonetic instruction (Gu & Chen, 2019) and the integration of multimodal classroom resources (Weninger et al., 2020), both of which may suit Vietnamese learners given their transparent orthography and syllable-timed rhythm. However, few studies have attempted to systematically extract implications from Singapore's experience to inform Vietnam's current shift toward English as a second language of instruction, especially in terms of institutional practices and pedagogy. While digital literacies are coming to be more widely acknowledged, their application in ESL is not yet entirely worked out in both contexts. Against this backdrop, the present review draws on recent empirical studies (2019–2025) to examine the strategies from Singapore's experience that can inform Vietnam's transition to second language instruction.

Methodology

Data sources

As emphasized by Zainuddin et al. (2024), the purpose of conducting a systematic review is to reap synthesized evidence from previous research that can serve as a meaningful point of departure in informing pedagogical and policy practices in similar contexts. In accordance with this charge, the current study sets out to identify and explore the most significant strategies from Singapore's second language teaching experience that are transferable to Vietnam's recent language education reforms. Sources were accessed through a systematic retrieval of scholarly reliable databases and specific ASEAN-targeted repositories in education, using Boolean search strings for scope and relevance (Cooper et al., 2019). To strengthen the quality and contextual specificity of the review, there are three inclusion criteria applied: (1) Studies must explicitly examine ESL implementation, pedagogical design, teacher education, or policy reform within the Singaporean education system; (2) Publications must appear in Scopus-indexed journals, or in peer-reviewed academic conference proceedings affiliated with Singaporean higher education institutions; (3) Selected publications have sufficiently detailed abstracts to allow for reliable extraction of main thematic and methodological information.

After selection, this review draws on twelve peer-reviewed articles and academic studies published between 2019 and 2025 that critically examine ESL education in the Singaporean context. The selected literature includes a mix of empirical studies, qualitative, quantitative, as well as systematic reviews, allowing for a well-rounded understanding of micro-level instructional

practices and macro-level policy developments (see the summarized table in the Appendix section).

Data analysis

The current study deployed a thematic analysis method to examine the selected studies. Following the qualitative analysis framework of Naeem et al. (2023), the analysis proceeded through six phases: familiarization with the qualitative data, initial code generation, accumulation of related themes, review of the found themes, title putting the themes, and presentation of data. Twelve studies were coded inductively to allow salient categories to emerge from the data, while also being informed by prior theoretical work on second language acquisition, bilingual education, and language policy. To ensure the trustworthiness of the findings, two independent coders analyzed the full dataset. Independent initial codes were created and later contrasted with one another to check for consistency and inter-coder reliability. Conflicts were settled and agreed upon through repeated discussion, leading to an improved coding scheme, which enhanced the credibility and dependability of the thematic findings (Korstjens & Moser, 2018). The coded data were then grouped into four analytical domains: (1) learner awareness, (2) learner knowledge, (3) skills development, and (4) teacher PD. The thematic findings appear to build the basis for identifying relevant practical implications that could inform Vietnam's ESL transition journey.

Findings

This section synthesizes twelve papers on ESL education in Singapore, grouping them into four interrelated themes that offer actionable insights for Vietnam's shift from EFL to ESL. Each theme reflects an area of ESL education in learners' awareness, knowledge, skills, and teacher capacity development.

ESL learners' awareness development

This theme deals with the construction of linguistic awareness, sociocultural alignment, and bilingual identity among ESL learners in a multilingual setting. Singapore-based work has shown how language policy and pedagogical practice may shape learner identity and enable culture-situated ESL instruction. A Liew (2020) examined the pedagogical application of Singlish – Colloquial Singapore English (CSE) – in Standard English writing instruction in Secondary 3 classes. A quasi-experimental, mixed-methods study was employed to compare outcomes among experimental groups adopting CSE-informed instruction versus control groups that were provided with traditional immersion teaching. The findings indicated that incorporation of students' colloquial linguistic repertoire in instructional tasks enhanced motivation and raised metalinguistic awareness without compromising proficiency gains. Ethnographic fieldwork also illustrated the ways in which students actively participated in metalinguistic reflection on their own use of language and negotiated variations between informal and formal English registers. The current study presented strong local evidence for the advantages of mainstreaming students' linguistic experiences in ESL classes to promote identity-affirming learning processes.

Building on the theme of identity and bilingual awareness, O'Brien et al. (2021) made a synthesis of national-level studies conducted under the Education Research Funding Programme (ERFP) in Singapore. Their review positioned bilingualism as both a pedagogical goal and a broader sociocultural project which is embedded in Singaporean residents' national identity. They emphasized that the development of ESL competence is supported through an eco-system of

societal, familial, and institutional structures that reinforce both English and Mother Tongue Languages. Singapore's bilingual policy underpins this awareness systematically through consistent investment in ERFP, targeted teacher training, and curricula that legitimates English and home languages across educational contexts. The report feeds into a discussion of how English in Singapore is used as a world language and a vehicle for national integration. Most notably, it advocated for ESL pedagogies that align with culturally grounded bilingual education policies and promote sociolinguistic reflection among both learners and educators. Most recently, Ong et al. (2023) examined the pedagogical functions of code-switching in ESL classrooms through observations and in-depth interviews. While previous studies emphasized identity and structural support, Ong et al. provided classroom-level evidence of how language awareness is enacted in real time. The study revealed that code-switching acted as a cognitive aid for comprehension and further as a bridge connecting classroom instruction to students' linguistic awareness and cultural backgrounds. Teachers used code-switching to clarify abstract concepts, while students employed it for peer support and expression of new knowledge. The authors warned against overreliance on code-switching, which could hinder full immersion, but ultimately supported its strategic use as a means to validate multilingual identities in the ESL classroom.

ESL learners' knowledge development

The second theme explores how ESL learners build critical literacy, academic discourse skills, and multimodal understanding through diverse media and instruction promoting higher-order thinking. With a focus on socio-phonetic variation, Gu and Chen (2019) explored the patterns in phonology of Singaporean children relative to their American and British counterparts using acoustic phonetic analysis. While some characteristics of non-rhoticity like British English and vowel fronting like American English were found, the study also identified unique pronunciation features in Singapore English, including reduced vowel height contrast and less defined tense-lax vowel distinctions. These findings pointed out that Singapore English is evolving with its own localized norms which are influenced by its multilingual context. The study underscored the importance of incorporating metalanguage and analytical tools into ESL instruction to help learners recognize and adapt to socio-phonetic diversity while maintaining intelligibility. This provides a useful parallel for Vietnamese teacher training, where more explicit instruction in linguistic metalanguage, such as how to teach and contrast segmental features like /θ/ and /t̪/, could enhance pronunciation instruction in EMI and ESL classrooms.

A study from Weninger et al. (2020), who examined how Singaporean students develop conceptual and conversational knowledge in ESL classrooms by constructing meaning across language, images, gestures, and other semiotic resources. Rather than focusing on language production skills, the study placed emphasis on how students build a knowledge of meaning-making processes by working with multimodal texts. This was done through developing an understanding of how meaning is conveyed in various modes and how these may be mixed to decipher sophisticated messages. Their findings support broader curricular trends that position multiliteracies and media literacy on a central platform of ESL knowledge spaces in order to enable students to navigate and analyze today's communication in socially and culturally appropriate language. The study advocated for pedagogies that treat students' media practices not as distractions, but as assets for building reflective, academically grounded literacy. The study advocated for pedagogies that treat students' media practices not as distractions, but as assets for

building reflective, academically grounded literacy, which is an approach that Vietnamese curricula could adopt to align media-integrated learning with formal academic outcomes.

Jones et al. (2025) combined 159 Singapore ELT studies of 2017-2023 via systematic review. Through the integration, the study identified several macro-trends related to ESL knowledge development, with the most significant being a growing emphasis on cultivating explicit knowledge of English grammar, vocabulary, and discourse structure. Amidst these trends, the convergence of critical literacy, translanguaging, and English as an academic language was linked with a shift in how one thinks about linguistic knowledge and how to teach it. Rather than learning about language rules as isolated abilities, the review referred to pedagogical approaches that allow learners to learn about the grammar structure, expand lexical repertoire, and navigate genres within a discipline. The authors emphasized that good ESL teaching should initiate the metalinguistic awareness to critically engage with the lexical texts of the academic discourse.

ESL learners' skills development

The third theme explores how ESL learners in Singapore acquire and consolidate a range of integrated language skills essential for reading, pronunciation, academic communication, and workplace readiness. Rather than treating these abilities in sections, the following research explains how ESL proficiency develops through interdependent layers affected by the language background of students, teaching environments, and lifelong learning routes. Gottardo et al. (2021) explored how core reading-related skills, in particular, phonological navigation, decoding, and inferences from text, transfer reciprocally between students' L1 and L2. Under the bilingual education system of Singapore, wherein a majority of students begin with firm L1 literacy, such cross-linguistic correlation facilitates easier acquisition of English word recognition, reading fluency, and oral intelligibility. Their studies showed that ESL acquisition should actively tap students' L1 capacities to scaffold major English literacy from the outset. Extending this platform into more applied settings, Zhang et al. (2022) studied vocational English courses in schools such as Singapore's Institute of Technical Education (ITE). Their study determined how fundamental ESL skills, especially reading, technical terms, and oral delivery, are customized for workplace communication across different sectors. This alignment of language training with occupational use ensures that students acquire communicative competence conducive to real work requirements, including task-related documentation and interactive teamwork.

In the tertiary settings, Bo et al. (2023) discover that English language proficiency significantly forecasted educational accomplishment in English-medium instruction (EMI) university settings once adjusted for students' previous academic achievements. Their study helped establish the importance of well-structured skill formation with grammar, syntax, reading for academics, and formalized writing. Grammar, syntax, academic reading, and formal writing were the major predictors of GPA and disciplinary participation, which suggests that fluency is not sufficient without on-going exposure to conventions of academic discourse (Bo et al., 2023). These skills enable students to sustain scholarly performance, then actively participate in disciplinary discourse and scholarly expectations at the post-secondary level. With a pragmatic classroom perspective thrown in, Ong et al. (2023) examined how code-switching operates as a strategic scaffolding tool for ESL learners. ESL teachers often applied code-switching to provide abstract content explanations, while students used code-switching for peer support and comprehension. While the study acknowledged its benefits in supporting oral processing and conceptual knowledge, it also cautioned against L1 overdependence. Ongoing ability development, especially

in verbal English, was noted to depend on increasingly mitigated use of L1 and escalating immersive exposure to English usage.

ESL teacher PD

The final theme addresses how teacher education and ongoing PD in Singapore have evolved to meet the complexities of ESL instruction in a dynamic, multilingual society. Across the reviewed studies, four interrelated dimensions of teacher development emerge, including (1) systemic alignment with national bilingual policies and sociocultural realities; (2) multimodal pedagogy that equips teachers to handle new literacy demands; (3) ideological reorientation toward pluralistic English use; (4) and reflective practice that empowers teachers as agents of pedagogical inquiry. Systemically, O'Brien et al. (2021) synthesized scholarship on national bilingual programs funded by the Education Research Funding Programme (ERFP). They called for the significance of institutional and community-based processes to build long-term teacher capacity. In particular, the study advocated for systemic congruence between language-in-education policy and classroom environments, calling for greater support for ESL teachers in mediating English and Mother Tongue instruction. By locating teachers as identity makers and cultural mediators, the research reaffirmed the importance of aligned policy, school culture, and teacher development tracks. Focusing on pedagogical innovation, Lim (2024) piloted the development of a pedagogic metalanguage to support secondary English teachers in teaching digital news texts. Through lesson observations, teacher interviews, and student focus groups, the study showed how multimodal literacy could be integrated into everyday ESL instruction. Teachers who engaged with this framework reported greater sensitivity to students' communicative repertoires and improved ability to design tasks that reflect real-world literacy practices. Lim's findings also brought up the importance of equipping teachers with the conceptual and assessment tools needed to meet curricular demands.

Moving from institutional to ideological considerations, Marlina (2024) proposed an English as an International Language (EIL)-oriented framework for teacher PD across Southeast Asia, with feedback drawn from Singaporean in-service teachers. Her program directly challenged monolingual and native-speaker biases by validating localized and pluralistic English usage in classrooms. Teachers reported significant shifts in belief systems and pedagogical confidence, particularly in incorporating diverse linguistic identities into lessons. This study invites rethinking teacher development as training for skills and ideological reorientation toward inclusive, internationally relevant English teaching with a set of recommendations for anti-native-speakerist PD in Vietnam where native-speaker norms sweep aside local practices. Marlina's model can inform the development of context-aware workshops that place the legitimacy of Vietnamese English at the forefront, feature successful models of non-native English-speaking teachers, and engage participants in reflective discussion and media critique challenging the native-speaker myth. In this manner, teacher education can become a site for rethinking linguistic identity and restoring pedagogical confidence, especially in regions where learners' multiple repertoires are subordinated.

Lastly, in Jones et al.'s (2025) meta-level points of view, teacher education emerged as one of five dominant research themes, particularly through the lens of teacher cognition, agency, and inquiry-based practice. Teacher education emerged as one of five dominant research themes, particularly through the lens of teacher cognition, agency, and inquiry-based practice. The review documented a professional culture that encourages critical self-reflection, practitioner research,

and context-driven pedagogical adaptation. Even within a small island nation, the study emphasized that instructional quality and innovation are best driven by teacher autonomy rather than by a one-size-fits-all model. This finding has considerable implications for Vietnam, where top-down models of PD often limit teacher innovation. Jones et al.'s emphasis on teacher research engagement suggests the need to promote practitioner-led inquiry within Vietnamese teacher education systems. PD should be reconceptualized to include small-scale classroom research and the cultivation of teacher research communities. In practice, this may take the form of local teacher research circles coordinated by pedagogical universities, periodic reflective teaching workshops, and Vietnamese-language publication platforms where teachers can share locally grounded innovations.

Discussion

To inform Vietnam's pathway to ESL, this section addresses the research question by discussing four key strategies related to ESL awareness, knowledge, skills, and teacher development.

Promoting learner awareness about ESL

Singapore's attention to sociolinguistic positioning and ESL learners' identity construction provides a critical model for Vietnam, where regional dialects and ethnic languages exist as a significant influence in students' lives. As seen in Liew (2020) and Ong et al. (2023), Singaporean classrooms strategically incorporate code-switching and local varieties to affirm learners' identities while enhancing English comprehension since these pedagogical choices validate students' linguistic backgrounds and foster motivation through cultural relevance. Vietnamese classrooms, especially in the Mekong Delta, Central Highlands, and northern border regions, can similarly benefit from inclusive language policies that permit controlled use of L1 to bridge learners' lived experiences and English instruction. On top of that, the prevailing notion of a "standardized" or "ideal" English accent should also be revisited. In the context of increasing globalization, it may be more constructive to move away from viewing any single English accent as inherently superior or more prestigious. In ELT, particularly in multilingual contexts like Vietnam, pronunciation goals might be more effectively guided by principles such as intelligibility, communicative clarity, and appropriateness to context rather than conformity to native-speaker norms. Therefore, deeply ingrained assumptions must be revisited about accent, thus encouraging greater tolerance of accent variation and reducing learners' linguistic insecurity.

Synthesizing metalinguistic awareness and shared linguistic habits can enhance an increased sense of belongingness and learning interest in English, especially among ethnolinguistically heterogeneous students. Metalinguistic reflection, or learners' ability to think about and discuss how language works, enables them to compare structures, meanings, and usage across English and their home languages. For instance, students may observe that both English and Vietnamese follow a Subject–Verb–Object word order, yet differ in verb conjugation, as English verbs change according to tense (e.g. "go" vs. "went") while Vietnamese relies on time markers like "đã" or "sẽ" instead. Encouraging such reflections helps learners make sense of English in relation to what they already know, deepens their understanding of language structure, and affirms the relevance of their linguistic backgrounds in the learning process.

Strengthening ESL knowledge through multimodality

Singapore's integration of multimodal literacy and grammatical awareness into ESL curricula has spotlighted the importance of broadening the definition of language knowledge beyond textbook-based instruction. The work of Weninger et al. (2020) demonstrates how Singaporean educators engage learners in critical interaction with digital and visual texts, preparing them for 21st-century communication. Meanwhile, Gu & Chen (2019) emphasize that knowledge of pronunciation and grammar must be taught with sensitivity to local socio-phonetic variation, particularly for learners whose first language differs structurally and phonetically from English. These developments reflect the theory of multimodality, which views communication as occurring through multiple modes, not just language, but also visual, auditory, and spatial forms (Kress & Van Leeuwen, 2020). From this perspective, ESL learning involves more than grammar and pronunciation; it requires the ability to interpret and produce meaning across diverse semiotic resources. Furthermore, the adoption of visual and digital texts in the ESL curriculum in Singapore mirrors the continuing process of multimodalization, with the objective of familiarizing the learners with multimodal communication in real-life situations.

In this aspect, Vietnam has a comparative edge over Singapore's largely Chinese-speaking population, in particular English literacy acquisition. Vietnamese orthography, like English, uses a Latin-based alphabetic writing system, whereas Chinese is character-based and logographic. This means Vietnamese learners are already familiar with letter-sound correspondences and linear spelling conventions, which can make the processes of decoding, spelling, and reading in English more intuitive. When supported by explicit instruction and exposure to multimodal texts, this familiarity with alphabetic print can help Vietnamese students access English vocabulary and grammar more efficiently. Building on this orthographic similarity, Vietnam's ESL curriculum could integrate Singapore's multimodal approaches while tailoring vocabulary and grammar instruction to capitalize on learners' existing literacy skills. The instruction of English in Vietnamese contexts has long been more inclined toward grammatical correctness and vocabulary memorization than toward functional discourse and multimodal meaning-making. To further support ESL, classroom content must become more diverse to include media texts, visual matter, and critical thinking activities in accordance with authentic communication. Such content not only supports scholarly skills but also engages learners using formats already familiar to them from daily life. In particular, the learning of vocabulary can be supplemented by the use of contextual multimodal approaches, such as brief videos, memes, infographics, or social media interactive elements, delivering the words in a fun and "trending" format. As students encounter new words through dense and dynamic contexts, they will be more likely to retain and use them meaningfully.

Consolidating ESL skills development across educational and vocational pathways

The efforts of Vietnam to increase the role of EMI and vocational English learning can be strongly supported by the holistic approach by Singapore in ESL skills acquisition across academic and technical disciplines. Investigations by Gottardo et al. (2021) and Zhang et al. (2022) showcase how systematic development in reading fluency, speaking intelligibility, and task-specific lexis are entertained by Singapore through a well-articulated, multi-level approach. This two-path system prepares students with academic skills, like report writing or presenting, as well as vocational communication skills necessary for job-related work, hence a seamless move from education to employment. In reality, teachers and learners in Vietnam continuously face mismatches between general English curricula and other specific language needs of sectors like tourism, customer

service, information technology, and manufacturing (Budiarta et al., 2020). Singapore's planned integration of domain-specific English in institutions such as the Institute of Technical Education (ITE) is a good example. Instructional design within ITE incorporates realistic communicative tasks, including delivering safety briefings, authoring service documentation, or dealing with long-term client interaction, as part of technical training programs. Vietnam's vocational ESL programs, on the other hand, remain extremely fragmented and underdeveloped (Tang & Nguyen, 2023). To offset this, Vietnam may work to develop inter-ministerial coordination, which is critical to ensuring ESL education is aligned with regional economic drivers and labor market demand. Collaborations are likely to initiate co-creation of workplace-relevant ESL syllabi for industry clusters of high demand.

Equally important is pedagogical recognition of learners' literacies in Vietnamese and minority ethnic languages. As Gottardo et al. (2021) persuasively argue, explicit instruction in transferring L1 strategies, for instance, reading comprehension, narrative proficiency, or oral sequencing, will significantly accelerate L2 attainment. By embedding Gottardo et al. 's findings within vocational ESL modules, Vietnam can facilitate more equitable lifelong access to language learning for linguistically diverse and under-resourced learners. Finally, rural access gaps require not only policy alignment but also delivery innovation. Vietnam can adopt gamified and blended learning models, such as proposed by PubMed Central studies on digital literacy and student motivation. These designs combine interactive content, local language scaffolding, and offline capability to make vocational ESL more sustainable and inclusive. By incorporating ESL into practical and realistic vocational work and employing blended tools, Vietnam can establish a context-responsive, skill-based ESL system to promote workforce competitiveness and social mobility.

Advancing teacher PD for ESL transition

Singapore's experience demonstrates that teacher development lies at the center of ESL reform but later evolves in tandem with shifts in learner needs, language use, and curriculum requirements. Just as ESL learners benefit from progressing through stages of awareness, knowledge, and skill mastery, teachers similarly need to go through a similar process of professional learning. Marlina (2024) and Lim (2024) assert that this begins with building consciousness, away from fixed models of language or native-speaker norms toward a more expansive model of English as a global and context-sensitive language. This foundational awareness enables teachers to recognize the diverse linguistic backgrounds of their learners and the sociocultural dynamics of the English language use in real-world settings. Building on awareness and mindset evolution, ESL teachers need to strengthen their pedagogical knowledge of multilingual strategies, multimodal literacy, and learner-centered approaches. These areas of knowledge equip teachers to engage with contemporary ESL instruction, where text, image, sound, and context often intersect. From this expanded understanding, teachers can then develop the instructional skills necessary to apply new methods, such as designing inclusive tasks, scaffolding multimodal materials, or adapting content for learners with different linguistic profiles. In Singapore, this developmental progression is supported by systemic initiatives that encourage teachers to act as critical inquirers, or the professionals who can reflect on their classroom practice, adapt curricula to local realities, and innovate in response to shifting learner demands (Jones et al., 2025). In Vietnam, specifically in ethnic minority settings such as Khmer pupils in the Mekong Delta, translanguaging classes are applicable. Piloting translanguaging approaches that combine

Khmer and English, replicating Singapore's home language use policy, would assist in consolidating identity and clarity. Bilingual pedagogies based on community, context-specific curriculum, and teacher training with ethnic minority linguistic and cultural resources can facilitate connections between policy intentions and practiced classroom realities.

In the field of evolving English language training in Vietnam, plenty of scope is available to extend these findings. PD programs can be designed to guide teachers through the stages of awareness, knowledge, and practice to foster technical competence, reflective capacity, and instructional adaptability. Especially in multilingual and regionally diverse settings, localized teacher training can play a transformative role by helping educators connect curricular goals with classroom realities, draw on learners' linguistic assets, and participate confidently in the country's transition toward a sustainable ESL framework. One culturally embedded strength in the Vietnamese teaching profession is the high value placed on professional responsibility and intellectual authority. Many Vietnamese teachers feel a strong obligation to "know in advance" what learners might ask or struggle with, an expression of care, preparedness, and respect for their role. Rather than viewing this belief as a constraint, PD programs can build on it to cultivate proactive, research-informed teaching practices. By positioning ongoing learning, reflection, and adaptation as extensions of teachers' responsibility to learners, PD programs can affirm existing values while introducing teachers to inquiry-based pedagogy, curriculum co-design, or technology integration.

Conclusion

This review examined how to learn from Singapore's ESL school system to inform Vietnam's continuing transition from EFL to ESL instruction. On the basis of thematic synthesis of recent research, four interrelated strategies were identified: developing students' sense of sociolinguistic diversity, constructing ESL capacity through multimodal participation, creating transferable capabilities across academic and vocational pathways, and progressing teacher PD based on inquiry, inclusive practice, and reflective practice. Combined, these dimensions reflect a socially situated and function-oriented ESL education strategy that directly mirrors the priority foundations of Vietnam's current reform stage. Although Singapore's ESL achievement rests on synchronized top-down policy and powerful institutional backing, Vietnam's reform context might better draw on more grassroots pedagogical ingenuity and greater teacher autonomy. There is a great need and opportunity for empowering regional teachers as curriculum designers and implementers with diversity in region and language by the nation. Coordination across sectors, especially between education and labor agencies, will also be essential to ensure ESL education aligns with vocational needs and job placement. To promote equity, especially in rural and resource-poor areas, the application of blended and gamified learning approaches that are centered on mobile access and student motivation is highly promising, regional and comparative analyses state.

At the same time, this review acknowledges several limitations, which are in its inability to bring in relevant research from adjacent fields, such as sustainability in language teaching, that may provide additional insights into long-term reform. Though research from 2019 to 2023 is well represented, that from 2024–2025 is still limited, suggesting a potential recency problem in capturing Singapore's most recent innovations and policy trends in ESL. Looking to the future, research needs to respond to the role of AI-driven ESL tools, including adaptive feedback systems and reinforcement learning, in contributing to personalized teaching and learning outcome

enhancement. Longitudinal studies could be a worthwhile consideration to examine sustainability of Vietnam's ESL and EMI reforms in the long run, particularly regarding teacher agency, student learning, and community involvement, to ensure a contextually appropriate and socially inclusive transition.

Funding: This research is not funded.

Acknowledgments: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Authenticity Statement: This manuscript is an original work.

Artificial Intelligence Statement: AI tools (Grammarly and QuillBot) were used only for language editing. The authors remain fully responsible for the content's accuracy and integrity.

About the Authors:

Phan Ngoc Tuong Vy is a Ph.D. student at the School of Foreign Languages, Can Tho University, Vietnam. Her research interests are teacher motivation and PD.
ORCID: https://orcid.org/0009-0004-1058-0069

Trinh Quoc Lap is a lecturer at the School of Foreign Languages, Can Tho University, Vietnam. His research focuses on curriculum development and sociocultural approaches in ELT.
ORCID: https://orcid.org/0000-0002-1256-8125

Le Cong Tuan is a lecturer at the School of Foreign Languages, Can Tho University, Vietnam. His interests include task-based teaching and psychoeducation.
ORCID: https://orcid.org/0000-0003-3055-6539

Dang Vu Hoai Nhan is a lecturer at the School of Foreign Languages, Can Tho University, Vietnam. He specializes in ESP, technology in language learning, and education policy.
ORCID: https://orcid.org/0009-0006-2443-8148

Le Xuan Mai is a lecturer at the School of Foreign Languages, Can Tho University, Vietnam. Her research focuses on teacher and learner development. She is the corresponding author.
ORCID: https://orcid.org/0000-0002-1300-3804

References

Bahanshal, D. (2023). From EFL to ESL in the Eye of Saudi Vision 2030. *Asian Journal of Education and Social Studies*. https://doi.org/10.9734/ajess/2023/v48i41079

Bo, W. V., Fu, M., & Lim, W. Y. (2023). Revisiting English language proficiency and its impact on the academic performance of domestic university students in Singapore. *Language Testing, 40*(1), 133-152. https://doi.org/10.1177/02655322211064629

Budiarta, I., Rasna, I., Padmadewi, N., & Artini, L. (2020). Applying content and language integrated learning (CLIL): Teachers' perceptions and challenges. *International Journal of Innovation, Creativity and Change, 14*(3), 78-94. https://www.ijicc.net/index.php/volume-14-2020/200-vol-14-iss-3

Bui, T. H. (2022). A theoretical evaluation of the English textbooks for grade 12 students under the National Foreign Language Project 2008 – 2020. *TESOL Communications*. https://doi.org/10.46451/tc.20220203

Cooper, H., Hedges, L. V., & Valentine, J. C. (2019). The handbook of research synthesis and meta-analysis. Russell Sage Foundation.

Gottardo, A., Chen, X., & Huo, M. R. Y. (2021). Understanding within-and cross-language relations among language, preliteracy skills, and word reading in bilingual learners: Evidence from the science of reading. *Reading Research Quarterly*, *56*, S371-S390. https://doi.org/10.1002/rrq.410

Gu, Y., & Chen, N. F. (2019). Large-scale acoustic characterization of Singaporean children's English pronunciation. *The Journal of the Acoustical Society of America*, *146*(4_Supplement), 2924-2924. https://doi.org/10.1121/1.5137153

Jones, S. A., Loh, C. E., & Sabnani, R. L. (2025). Research in ELT and learning in Singapore: 2017–2023. *Language teaching*, 1-23. https://doi.org/10.1017/S0261444825000114

Korstjens, I., & Moser, A. (2018). Series: Practical guidance to qualitative research. Part 4: Trustworthiness and publishing. *European Journal of General Practice*, *24*(1), 120-124. https://doi.org/10.1080/13814788.2017.1375092

Kress, G., & Van Leeuwen, T. (2020). *Reading images: The grammar of visual design*. Routledge. https://doi.org/10.4324/9781003099857

Liew, W. M. C. Y. (2020). Investigating the effectiveness of language awareness approaches to ELT in Singapore secondary schools (Report No. OER 21/15 LCY). *Office of Education Research, National Institute of Education, Singapore*. https://hdl.handle.net/10497/22688

Lim, F. V. (2024). Thinking and talking about digital news in the Singapore secondary English classroom: A pilot study. *English Teaching & Learning*, *48*(1), 97-116. https://doi.org/10.1007/s42321-022-00134-5

Marlina, R. (2024). "Who says Singlish cannot be used?": An EIL-aware teacher PD course in Singapore. *TESOL Journal*, *16*(1), e899. https://doi.org/10.1002/tesj.899

Naeem, M., Ozuem, W., Howell, K., & Ranfagni, S. (2023). A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. *International journal of qualitative methods*, *22*. https://doi.org/10.1177/16094069231205789

Nguyen, H. T. T., & Duong, D. (2021). National Foreign Language policy. In T. Neimann, J. T. Y. Lin, & J. Paulus (Eds.), *Stagnancy issues and change initiatives for global education in the digital age* (pp. 212–235). https://doi.org/10.4018/978-1-7998-4993-3.CH010

O'Brien, B. A., Sun, H., Chua, D. F., Ow, L., & Sun, B. (2021). Local evidence synthesis on language learning and bilingualism. *Office of Education Research, National Institute of Education (Singapore)*. https://hdl.handle.net/10497/23499

Ong, E. S., Koh, W. T., & Teo, A. (2023). Code-switching in Singaporean esl classrooms: how it affects classroom discourse and student understanding. *African Journal of Emerging Issues*, *5*(13), 10-24. https://ajoeijournals.org/sys/index.php/ajoei/article/view/472

Reynolds, B., Shieh, J.-J., Ding, C., & Van Ha, X. (2022). Sustained content language teaching: Insights from an ESL and EFL course. *International Journal of English Studies*. https://doi.org/10.6018/ijes.507161

Sabe, C. Z., & Luka, E. (2020). Impediments to learning the English language in Nigerian education: A philosophical discourse. *International Journal of Educational Research*, *7*,

97-107. https://consensus.app/papers/impediments-to-learning-the-english-language-in-nigerian-luka-sabe/4a9c46f8ef3d59cdaf794aed2129ff3a/

Sato, M., & Storch, N. (2020). Context matters: Learner beliefs and interactional behaviors in an EFL vs. ESL context. *Language Teaching Research*, *26*, 919-942. https://doi.org/10.1177/1362168820923582

Singh, A. K. (2025). From policy to practice: Challenges and pathway for advancing multilingual education through India's national education policy 2020. *Policy Futures in Education*, *23*, 746-765. https://doi.org/10.1177/14782103251320532

Tang, C., & Nguyen, H. T. (2023). An evaluation of the effectiveness of Vietnam's vocational training project for rural laborers. *Adult Education Quarterly*, *73*, 286-309. https://doi.org/10.1177/07417136231172180

Tang, H. (2020). Linguistic landscaping in Singapore: multilingualism or the dominance of English and its dual identity in the local linguistic ecology? *International Journal of Multilingualism*, *17*, 152-173. https://doi.org/10.1080/14790718.2018.1467422

Trinh, L. Q., Phan, V. N. T., & Ngo, N. H. H. (2025). Second-career EFL teachers' experiences of adaptability and sustainability in ELT. *Educational Process International Journal*, *14*(1). https://doi.org/10.22521/edupij.2025.14.46

Weninger, C., Choo, S. S., Hu, G., Williams, P., & Kan, K. H. Y. (2020). Media literacy in the teaching of English in Singapore (Report No. OER 50/12 CW). *Office of Education Research, National Institute of Education (Singapore)*. https://doi.org/10.32658/10497/22725

Yoo, H.-K., & Kim, H. (2023). A cross-modal analysis of lexical sophistication: EFL and ESL learners in written and spoken production. *International Review of Applied Linguistics in Language Teaching*, *0*. https://doi.org/10.1515/iral-2023-0204

Zainuddin, N. M., Mohamad, M., Zakaria, N. Y., Sulaiman, N. A., Jalaludin, N. A., & Omar, H. (2024). Trends and Benefits of Online Distance Learning in the English as a Second Language Context: A Systematic Literature Review. *Arab World English Journal*. https://dx.doi.org/10.24093/awej/call10.18

Zhang, W., Xiaoya, S., & Cheung, Y. L. (2022). Vocational English language education in multilingual context: the case of Singapore. *Research in Post-Compulsory Education*, *27*(3), 436-453. https://doi.org/10.1080/13596748.2022.2076056

Zulkepli, N., Ariff Albakri, I. S. M., Puteri Zarina, M. K., & Farish Ishak, M. F. (2024). Using the ADDIE Model to Develop an Online Professional Development Program for Non-Specialist ESL Teachers. *Arab World English Journal*. https://dx.doi.org/10.24093/awej/call10.19

Appendix

A summary table of selected studies (2019-2025)

Theme	Study	Methods/Materials	Participants	Key findings
Awareness Development	Liew (2020)	Mixed methods: Quasi-experiment + observations	Secondary 3 students (Express & Normal streams)	Integrating Singlish boosted motivation and metalinguistic awareness.
	O'Brien et al. (2021)	Policy synthesis of 18 ERFP-funded projects (2013–2017)	N/A	Bilingual awareness is supported through coordinated policy, family, and school efforts.
	Ong et al. (2023)	Mixed methods: Classroom observations + interviews	Observations from 10 ESL classrooms and 30 semi-structured	Code-switching aids comprehension and affirms multilingual identities.

			interviews with teachers and students	
Knowledge Development	Gu & Chen (2019)	Acoustic analysis of vowel pronunciation using Praat	414 children (192 Singaporean, 140 American, 82 British; ages 6–13).	Unique local phonetics highlight the need for metalanguage-focused instruction.
	Weninger et al. (2020)	Mixed-methods: survey + classroom observation + curriculum co-design	202 secondary English teachers (survey); 4 teachers and ~80 upper secondary students (classroom phase)	Multimodal texts foster critical, socially grounded literacy.
	Jones et al. (2025)	Systematic review of 159 ELT studies	N/A	Emphasized trends: translanguaging, critical literacy, and academic discourse focus.
Skills Development	Gottardo et al. (2021)	Narrative review of 151 studies on L1–L2 reading factors in bilinguals	N/A	L1 literacy skills predict L2 fluency and reading comprehension.
	Zhang et al. (2022)	A desk-based review of 53 sources (27 journal articles, 2 books, 2 book chapters, 1 thesis, 21 government or online documents)	N/A	Workplace English must be job-specific and contextually relevant.
	Bo et al. (2023)	Quantitative: Regression analysis	514 Singaporean undergraduate students	ESL proficiency significantly predicts GPA beyond prior achievement.
	Ong et al. (2023)	Mixed methods: Observations + interviews	Observations from 10 ESL classrooms and 30 semi-structured interviews with teachers and students	Code-switching scaffolds comprehension and peer interaction.
Teacher PD	O'Brien et al. (2021)	Policy synthesis of 18 ERFP-funded projects (2013–2017)	N/A	PD should align with bilingual policy and foster cultural mediation.
	Lim (2024)	Qualitative case study: observations, interviews, focus groups	Two English teachers and students in two Secondary 1 classes (1 Express and 1 Normal stream)	Metalanguage enhances multimodal instruction and feedback practices.
	Marlina (2024)	Qualitative case study: EIL-based PD course evaluation + teacher feedback	In-service English teachers in Singapore	EIL-based PD challenges native-speakerism and empowers teachers.
	Jones et al. (2025)	Systematic review of 159 ELT studies	N/A	Reflective, inquiry-driven PD supports autonomy and innovation.

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp. 285-300
DOI: https://dx.doi.org/10.24093/awej/vol16no3.17

Preference Structure in Refusals by Chinese Female International Students

He Zhai
Department of English Language, Faculty of Languages and Linguistics, Universiti Malaya Kuala Lumpur, Malaysia
Email: s2037030@siswa.um.edu.my

Ridwan Wahid
Department of English Language, Faculty of Languages and Linguistics, Universiti Malaya Kuala Lumpur, Malaysia
Corresponding author: ridwanwahid@um.edu.my

Woo Wai Sheng
Department of Asian and European Languages, Faculty of Languages and Linguistics Universiti Malaya, Kuala Lumpur, Malaysia
Email: woows@um.edu.my

Received: 06/29/2025 Accepted:08/21/2025 Published: 09/20/2025

Abstract
This study aims to examine refusals from a sequential perspective guided by Politeness Theory. It advances politeness theory and contributes pertinent data to L2 pragmatics by exploring how female students utilise linguistic, social, and situational resources to construct preference structures and convey politeness. The research involved 60 Chinese female international students enrolled at a university in Malaysia. Refusals were elicited through Discourse Completion Tests, a questionnaire with 16 scenarios that varied in three controlled variables: social distance (acquaintances, strangers, intimates), eliciting acts (suggestions, offers, requests), and gender (addressing the same, opposite, or unspecified gender). The study analysed preference structures from the dimensions of content, order, and frequency based on refusal classification developed by Beebe et al. (1990), guided by Brown and Levinson's Politeness Theory. The findings revealed that participants varied in preference structures, with sequences prioritising multi-element patterns of more recycled sub-strategies, such as 'Reason' and 'Gratitude' to mitigate face-threat and encode politeness in their refusal production. The study offers pedagogical implications for pragmatic teaching in intercultural communication contexts.

Keywords: Chinese, Discourse Completion Tests, female international students, preference structure, refusals

How to cite: Zhai, H., Wahid, R., & Sheng, W.W. (2025). Preference Structure in Refusals by Chinese Female International Students. *Arab World English Journal, 16*(3): 285-300.
https://dx.doi.org/10.24093/awej/vol16no3.17

Introduction

Research on the competence of English learners in realising speech acts, such as refusals, has received vigorous attention (Ali Qassim et al., 2021; Boonsuk & Ambele, 2019; Taguchi & Roever, 2020). Refusal is recognised as a particularly complex speech act (Kwon, 2004). From a sequential perspective, refusals require not only long sequences of negotiation in a coherent, organised, and purposeful manner and cooperative achievements, but also 'face-saving manoeuvres' to accommodate the non-compliant nature of the acts (Félix-Brasdefer, 2006). To redress the threats to the faces of both interlocutors, learners should fine-tune direct strategies, such as 'no' and 'will not', as well as indirect strategies, including reasons and adjuncts like prefatory particles ('well', 'oh'), in different sequences to manage politeness in refusal formulation. These systematic sequences in permutations and combinations refer to preference structures that can effectively express rejection and lessen the offence to interlocutors (Wong & Waring, 2020). Politeness Theory, proposed by Brown and Levinson (1987), indicates that politeness plays a crucial role in communication, even in potentially sensitive situations, such as refusals. Refusals are inherently Face Threatening Acts (FTAs) and can pose a threat to interlocutors' positive or negative face. Refusing felicitously and politely presents a significant challenge for English learners, particularly when they use their preference structures, as it involves reducing the negative impact on the interlocutors and avoiding a clash in communication.

Although most students possess sufficient linguistic proficiency, they still face challenges in employing appropriate pragmatic means to establish a congenial relationship between interlocutors. In a foreign context such as Malaysia, students may struggle to mobilise effective strategies to negotiate refusals due to unfamiliarity with the local culture, cultural dynamics, and academic pressures. Thus, they are likely to overuse direct refusals in such a context (Lim, 2014), which could lead to displeasure or discomfort for the interlocutors (Bella, 2014). They may impinge upon the interlocutor's face due to a failure to mediate pragmatic choices in performing the speech act of refusal. This means that the problems can lead to pragmatic failure in communication (Németh, 2004). As such, realising the speech act of refusal requires considerable pragmatic competence to alleviate refusals with appropriate strategies. The selection of refusal strategies is not only different across cultures but also different among individuals. In Chinese culture, females are considered to be more delicate, modest, indirect, and polite, and these expectations have influenced the realisation of refusals, navigating how politeness is expressed (Wardhaugh, 2006). In such cases, the study is meaningful to look at how Chinese female international English learners in Malaysia deploy refusal strategies with preference structures as a redressive device. Researchers have gained insight into refusal patterns, highlighting differences between males and females in the use of refusal strategies (Abed, 2011; Alalawi, 2022). However, females' refusals have received little attention (Ren & Woodfield, 2016). The following research questions guide the present study, framed within Brown and Levinson's Politeness Theory.

1. What refusal strategies do Chinese female international students use?
2. How do Chinese female international students sequentially organise their refusals as preference structures?

Literature review

A substantial body of research in the field of pragmatics has explored various speech acts (Elsweiler, 2022; Wu & Roever, 2021). Among these studies, the ones on refusals have particularly attracted the attention of scholars (Allami & Naeimi, 2011). These foundational

achievements have provided valuable insights into how learners issue refusals using different strategies. As pioneers in the study of refusals, Takahashi and Beebe (1987) explored pragmatic transfer when Japanese English learners (at both lower and higher proficiency levels) produced refusals through a DCT. They coded data in terms of frequency, order, and content to identify the degree of pragmatic transfer. The results indicate that higher-proficiency learners showed a greater propensity for pragmatic transfer, although pragmatic transfer was present in learners at both lower and higher proficiency levels. Subsequently, Beebe et al. (1990) continued to use the same dataset to investigate refusals and developed a more refined coding scheme. Following this path, most follow-up studies resorted to DCTs to elicit refusals (Allami & Naeimi, 2011). However, some scholars have employed role-plays as a data collection instrument. For example, Taguchi (2013) investigated differences in the use of refusal strategies among English learners at higher and lower proficiency levels and native English speakers. The results show that the appropriateness of refusal strategy was affected by proficiency: learners at higher proficiency levels demonstrated a higher level of pragmatic competence. They could produce refusals more effectively and efficiently. Similarly, Bella (2014) drew on role-plays to collect data, comparing the pragmatic competence of learners at three proficiency levels with that of native speakers. She reported that learners' pragmatic competence improves as their proficiency increases, yet what advanced learners performed was still different from that of native speakers. These studies have underscored the role of proficiency in the selection of refusal strategies; however, few have examined refusals from the perspective of gender.

Recently, studies have explored fleeting behaviours in the use of refusal preference structures, relying on the fine-grained toolset of Conversation Analysis (CA). CA provides micro insights into how politeness unfolds in refusal production, thereby enhancing Brown and Levinson's Politeness Theory. For example, Al-Gahtani and Roever (2018) compared the refusal strategies applied by learners at advanced and intermediate levels, as well as native English speakers, from a sequential perspective. The findings show intermediate proficiency learners exhibited greater delay in refusal production through 'yes but' compared with low-proficiency learners. In contrast, advanced students demonstrated more pragmatic competence and did not show any delay in selecting a refusal strategy. For native speakers, they tended to use the prefatory particle 'well', which was not adopted by either proficiency level. Overall, all three groups attempted to work out an effective means of organising refusal sequences (Brown & Levinson, 1987). Similarly, Abbas and Chee (2023) discovered that the structure of 'regret+non-performative statement' is an indirect and effective way to maintain interpersonal relationships, claiming "learners tend to implicitly reject to protect face by avoiding explicit refusals" by examining refusal responses among Iraqi and Malaysian EFL learners.

Despite the numerous studies above, to date, limited research has probed into preference structures, and those on refusal preference structures involving Chinese females in L2 pragmatics have been even more scarce due to cultural taboos, which constrain females' genuine expressions and overdependence on the universal frameworks of refusals, overlooking the nuanced differences in gender. Students might overrely on L1-based preference structures such as 'regret+future acceptance' to formulate refusals in English (Liao & Bresnahan, 1996), likely hindering smooth communication. To fill the gap, this study aims to provide an understanding of how Chinese female international students in Malaysia issue refusals with proper preference structures to convey politeness via DCTs based on Beebe et al.'s classification.

Methodology

Building upon the theoretical principles, this section examines the study's methodology, including participants, instruments, and research procedures. This study took a mixed approach. After the DCT was employed to collect data, the refusal responses were categorised into sub-strategies of direct, indirect, and adjunct categories (content) proposed by Beebe et al. (1990) to address Research Question 1. To answer Research Question 2, the preference structures were examined using the dimensions of order and frequency and analysed through the lens of Brown and Levinson's Politeness Theory.

Participants

The participants were recruited using the snowball sampling technique, whereby existing participants referred others they knew (Albirini, 2016, p. 55) due to access challenges. It is important to note that the generalizability of findings is limited because the participants come from homogeneous social networks. Sixty Chinese female students (between 19 and 21 years old) from different majors (TESOL, Computer, English, Education…) at Universiti Malaya (UM) participated in this study. These students were Chinese heritage learners who had received formal English language education since primary school. All participants were admitted to UM with the requisite IELTS scores (IELTS 5.5+), indicating their sufficient proficiency in reading, comprehending, and completing the DCTs without difficulty.

All participants received a small gift as a token of appreciation for their involvement. They all signed informed consent forms prior to the study, allowing them to make informed decisions about their participation. They had the option to withdraw at any time. The experiment did not injure the participants in any manner, either intentionally or unintentionally, by inflicting pain. No identifiable information will be shared with or become accessible to anyone apart from the researchers and their recruiters (colleagues who assisted in identifying potential participants and administering DCTs for the researchers).

Research Instruments

The instrument was a Discourse Completion Test (DCT), including 16 situations with contextual prompts in each one. There were three controlled variables in each situation: social distance (acquaintances, strangers, intimates), eliciting acts (suggestions, offers, requests), and gender (addressing the same gender, opposite gender, unspecified gender). The variables within the DCT were set, as shown in Appendix A. The DCT was to elicit refusals and facilitate pragmatic analysis. Although DCT includes hypothetical situations that may not capture spoken nuances such as tone and hesitation (Golato, 2017), which might fail to provide a broader picture to look at refusal preference structures in natural communications, it can collect data rapidly and efficiently and control variables such as gender, age, and social distance, making it a reasonable method for researchers (Olshtain & Blum-Kulka, 1983). The DCT had two sections: the first part included personal information such as name and significant, and the second part presented 16 situations to elicit data.

The DCT was adapted from Beebe et al.'s (1990) research design, which had already been pilot-tested. The DCT was valid and reliable, helping the researcher control various contextual variables, as highlighted by Blum-Kulka and Olshtain (1984). To ensure its suitability in the Malaysian context, minor adjustments were made; for instance, 'KLCC' was substituted in situation 11 to incorporate a relevant Malaysian cultural factor. The instrument's validity was

further confirmed by three experts in pragmatics (Expert A is a professor in pragmatics at UM; Expert B is the editor-in-chief of the Journal of Guangdong University of Foreign Studies; Expert C is a senior lecturer in applied linguistics at YC University in Shanxi Province, China) (Contextual validity means relevance rating=4.7/5). Meanwhile, the pilot study established the reliability of DCT, ensuring it was well-suited for the Malaysian environment. The following is an example:

Situation 5: One day, you want to buy a skirt with your sister, but you do not have enough money to buy a branded one. Your sister says, "We can go to the bazaar to get a cheap one."
You: _____

Research Procedures

The DCT was administered via online and offline channels. Before proceeding with the task, participants were asked to read the instructions and write their answers in the blank space after 'You' in the DCT. There was no set time restriction, so they had sufficient time to finish the DCT without being urged. The first researcher gave students appropriate instructions but did not present any biased views. After data collection, data analysis was conducted based on the coding scheme developed by Beebe et al. (1990). All refusal responses were systematically parsed into three main strategies (direct, indirect, and adjuncts) and sub-strategies (see Appendix B). These strategies are composed of a set of expressions, which could be "a word(s), phrase(s), or sentence(s) which can function as a refusal" (Beebe et al., 1990). The preference structures were determined guided by Politeness Theory. For example, in [Sorry+I have made a promise to Penny], 'sorry' functions as an expression of 'Regret' within indirect category indexed with B, and 'I have made a promise to Penny' serves as a 'Reason' also within indirect category indicated with B. Hence, the preference structure of refusal is 'B+B' format. In such a way, any combination of the 36 sub-strategies in Appendix B, such as 'Regret, Wish, Reason…', can form a preference structure. This suggests that even with the same strategy, the sequence, for instance, [positive feeling+apology] versus [apology+positive feeling], brings about different preference structures. To ensure the reliability of coding, a discussion with a second coder was used to settle any disagreements. Two independent coders achieved high inter-rater reliability (Cohen's $\kappa = 0.87$, 95% CI: 0.82-0.92). Then, the frequency of each format was calculated by dividing its number of occurrences by the total number of formats within the corresponding pattern (one-element, two-element, three-element, or multi-element).

Results

Based on the methodology, the findings indicate that participants utilised distinct sub-strategies under the three main categories (direct, indirect, and adjuncts) to perform refusals thereby answering research question 1; students produced one-element patterns across three types, two-element patterns across nine kinds, three-element patterns across 20 kinds, and multi-elements patterns across 38 types in different orders as their preferred structures to make refusals thereby answering research question 2. The following section illuminates the findings across the 16 situations.

Refusal Strategy Types

The data (see Appendix B) shows that the participants adopted 36 sub-strategies across the three main refusal strategies: direct, indirect, and adjuncts. There were three sub-strategies within

the direct category: Flat no, Negative proposition, and Performative; 17 sub-strategies under the indirect strategy category: Regret, Reason, Alternative, Future acceptance, Negative consequences, Criticism, Letting off the hook, Proverb, Topic switch, Questioning, Unspecific reply, Asking for approval, Sarcasm, Wish, Principle, Lecture, Self-defence; 16 sub-strategies within adjuncts: Hedging, Compliment, Gratitude, Positive feeling, Pause filler, Empathy, Non-verbal strategies, Address terms, Curse, Polite remarks, Parenthesis, Imperative mood, Greeting, Intensifier, Complaint, Comfort.

Preference Structures

From the table in Appendix B, it can be seen that among the three main refusal strategies, direct (A: 1–3), indirect (B: 4–20), and adjuncts (C: 21–36), females produced different types of refusal sub-strategies (1–36) in the form of a single sub-strategy or a combination of two or more to formulate their refusals across the 16 situations. In the study, 'sub-strategy' and 'element' had the same meaning and thus could be used interchangeably. The following section categorises the preference structures into four groups (one-element, double-element, three-element, and multi-element) and explicates them using 21 examples. S1, S2 represent the 16 situations,...and S16, and so on, while the numbers with '+' signs in parentheses indicate the corresponding composition of preference structures. Given the diversity of each format within each structure, this paper selects representative examples for analysis, arranged without regard to order.

One-element Preference Structures

In this group, female learners adhere to only one type of sub-strategy (under three main strategies, A, B, and C) to communicate with finesse.

'A' formatted refusal: In situation 12, where a female student refuses to lend her notebook to another female student, the participant uses direct devices such as 'no' quickly and immediately in Example 1. The unequivocal reply 'nope' without mediated devices is a vigorous rejection sequence, signalling the looming danger of interaction. In this manner, the refuser steers the conversation into a closing ring and brings the exchanges to an end. 'A' format was used in 7.4% of one-element patterns, most frequently in response to a male acquaintance's request in situation 9 (13.3%).

Example 1：Nope. S12(1)

'B' formatted refusal: Another feature of one-element formatted refusal is the adoption of indirect strategies. In situation 8, where a female employee refuses her male boss's suggestion to have her car fixed at his brother's workshop, the refuser utters a free-standing reason, 'I have already made an appointment' in Example 2. Her rejection is embodied in this account, intended to lessen the face threat. 'B' format accounted for the largest share (87.3%) of the total, with the peak occurrence in reply to a suggestion from a male stranger in situation 7 (10.1%).

Example 2: I have already made an appointment for a workshop. S8(5)

'C' formatted refusal: The preference structure is preceded by a compliment to avoid an overt display of disaffiliation in Example 3. The participant does not refuse immediately. Instead, she uses the phrase 'do well' to camouflage her refusal as praise for the cake, thereby implying her unwillingness to lend her notebook to another female student for examination preparation in situation 12. As a result, the hidden core refusal disguised as praise influences the interaction and prompts her to give a straightforward 'answer'. The format was adopted in 5.4% of one-element patterns, with the highest rate rejecting a female stranger's offer in situation 13 (27.3%).

Example 3：You can do well on your own. S12(22)

Double-element Preference Structures

In this group, there are nine kinds of formats, each combining two sub-strategies, such as 'A+B', 'B+A', and 'B+B'. 'B+B' has the highest percentage (47.4%). The type not only includes two different sub-categories (A, B, and C), but also encompasses the concurrent use of the same sub-strategy covered within a single category. For instance, 'Regret+Reason' in the format of 'B+B' still falls into the type**.** The following section provides a detailed explanation.

'A+B' formatted refusal: 'No' is a symbol of bluntness. In situation 5, where the female participant rejects her sister's suggestion to buy a cheaper skirt from the bazaar, the participant uses 'but' which serves as a buffer zone to leave room for the refuser to state her principle after she utters a tone-strengthening marker, 'no' as seen in Example 4. Then, she offers a reason to desensitise the refusee and promote the negotiation of refusals, which demonstrates her greater interactional competence. This format amounted to 13.9% of double-element patterns, with the highest frequency in response to a gender-unspecified stranger's suggestion in situation 4 (14.0%).

Example 4: No, but I think this one is better. S5(1+5)

'B+A' formatted refusal: In situation 9, where a female English teacher refuses her male colleague who cannot handle his class and asks her for help, the refuser expresses her time occupation in the response 'I had other tasks at that time' and then shows her negative ability 'could not' in Example 5. This two-element sequence is used to soften the refusals and enhance solidarity. The format was discovered in 1.9% of double-element patterns, mainly for a male acquaintance's request in situation 9 (37.5%).

Example 5: I had other tasks at that time and could not help you. S9(5+2)

'B+B' formatted refusal: In situation 11, where a female refuses to lend her car to a male friend to drive to KLCC, the participant's response is illustrated in Example 6. A punchy 'sorry' without any colouring is to forestall cooperation so that the refusee has no chance to complain. The explanation, 'I have my cars fixed' further strengthens the rejection. The preference structure launched here is a typical evasive strategy in refusals. The 'B+B' format (47.4%) dominated among the structures, occurring most frequently in situation 12 (5.5%), in response to a female stranger's request.

Example 6: Sorry, I have my car fixed these days. S11(4+5)

'C+B' formatted refusal: In situation 2, where a female declines her female friend's suggestion for buying a research book in a bookstore, the participant takes an interest in the prefatory particle 'oh' and cites the reason 'it is out of date' in Example 7. Although the enacted, chatty, and relaxing tone of 'oh' and the implied meaning of 'apology' seem to cohere and make the interlocutors feel closer, a sense of refusal lingers in the air. The occurrence of 'Pause filler' and 'Reason' within two domains (C and B) is not intended to disengage her from her good friend, but rather to prevent the refusee from making another request. The 'C+B' format represented 29.2% of double-element patterns, with the top occurrence in situation 15 (18.5%), replying to an unfamiliar female's offer.

Example 7: Oh, but it is out of date. S2(25+5)

'A+C' formatted refusal: In situation 15, where a female refuses an unfamiliar female student who offers her a lift home, the participant aggravates the positive face of the refusee by uttering the direct strategy 'no' to refuse the stranger, as seen in Example 8. Surprisingly, the refuser subsequently uses a 'Curse'. The 'Curse' element serves as a special kind of performative verb for venting emotions. This structure, as a preference structure, is weak in maintaining good

relations, as politeness is at the lowest level. The 'A+C' format constituted a small percentage (1.4%) of double-element patterns, with the highest occurrence in situation 15 (50.0%), which involved refusing an offer from a female stranger.

Example 8: No. Go away! S15(1+29)

'C+A' formatted refusal: To refuse a female friend's suggestion for buying a research book in a bookstore in situation 2, the participant mediates her action by using a 'play it safe' strategy (thanks for…) without more effusive displays to preserve a good relationship with her schoolmate in Example 9. However, she is still willing to reveal her genuine feelings. So she indicates a change in stance through 'but' to pre-empt the refusee's further expectations about the offer. By deploying a greater range of sequential and linguistic elements, she displays much more ability to postpone the core refusal for politeness. The 'C+A' format was applied in 0.5% of double-element patterns, appearing once in situation 2, refusing a female friend's suggestion, and once in situation 14, declining a male friend's suggestion.

Thanks for your help, but I do not think so. S2(23+2)

'C+C' formatted refusal: In situation 13, where a female student refuses her female lecturer, who lives in the same neighbourhood, and offers to drive her to school, the participant employs an unstraightforward expression as a preference structure in Example 10. Instead of immediately acceding to the request and producing her refusal, the refuser adopts a positive agreement, saying 'thank you', and breaks the contiguity with the address term 'Dr'. This implies that the refuser tends to introduce a potential reason to turn down this offer. The 'C+C' format was only adopted in 1.4% of double-element patterns, most in situation 14 (66.7%), responding to a male acquaintance's suggestion.

Example 10: Thank you, Dr… S13(23+28)

Three-element Preference Structures

The data reveal that participants adopt three-element formats across 18 kinds to tackle the complications in refusal production, such as 'A+C+B', 'A+B+B', and 'C+A+C'. 'B+B+B' took up the most significant portion at 21.7%. The following is to illustrate them in detail.

'A+C+B' formatted refusal: In situation 2, where a female declines her female friend's suggestion for buying the recommended research book in a bookstore, the participant utters a direct strategy ('no'), which attracts the eyes of the targeted interlocutors in Example 11. The refuser turns to 'thanks' to lay out an alternative possibility, mitigating her refusal. Then she reframes her refusal by expressing a reason (this book may not relate to my topic) to build rapport. The 'A+C+B' format comprised 7.6% of three-element patterns, with the highest percentage in situation 10 at 26.7%, in response to a male stranger's offer being rejected.

Example 11: No thanks, this book may not relate to my topic. S2(1+23+5)

'A+B+B' formatted refusal: In situation 4, where a female lecturer teaching international students English refuses one student's suggestion to postpone a quiz to the following week, the participant's preference structure incorporates pragmatically suitable discourse markers in Example 12. 'Flat no', 'Principle', as well as 'Reason' are all effective devices deployed by the refuser to manage interactional contingencies and steer the conversation toward a targeted facet of the refusal domain. This format constituted 2.5% of three-element patterns, used once in situations 1, 3, 4, 5, and 16, respectively.

Example 12: No, you have to take this quiz today, because we have other plans for next week. S4(1+18+5)

'C+A+C' formatted refusal: In situation 13, where a female refuses her female lecturer

who lives in the same neighbourhood and offers to drive her to school, the participant combines three elements, including two adjuncts (C) discontinuously and one direct sub-strategy (A) in the middle, not to spark tensions in Example 13. After the gratitude "Thank you very much for…", the refuser turns to 'cannot' as a harbinger of denial to show her rejection. Then the refuser winds up the conversation with 'Compliment' (You are nice), as if giving the refusee a 'green light' to relax. The preference structure reveals the refuser's true thoughts, which are central to the act of refusal. This format accounted for the smallest share (0.5%) of three-element patterns, only once in situation 13, which involved responding to a female stranger's offer.

Thank you very much for the kind invitation, but I cannot take it. You are nice. S13(23+2+22)

'A+C+C' formatted refusal: In situation 4, where a female lecturer teaching international students English refuses one student's suggestion to postpone a quiz to the following week, the refuser assembles sources of 'Negative proposition' (not available) and 'Imperative mood' (come on) followed by to indicate her viable interactional actions in refusal in Example 14. The endearment of 'baby' at the end can be a 'last-ditch effort' to capture the refusee's attention. This is to attenuate the adverse effects and enhance group cohesion. The preference structure of 'Negative proposition', 'Imperative mood', and 'Address term' serves as an indirect and mitigating device to exonerate the refuser from blame. The format was adopted in 0.5% of three-element patterns, specifically in situation 4, where it rebuffed a gender-unspecified stranger's suggestion.

Example 14: I am not available to come, baby. S4(2+32+28)

'B+B+A' formatted refusal: Situation 11 demonstrates a scenario where a female refuses her male friend's request to use her car to go to KLCC. To address this, the participant makes a repentant utterance, such as 'sorry,' to seek forgiveness. The word 'sorry' serves as a way to soften the refusal for specific reasons, such as the risk involved in lending her car, and is intended not to insult but to comfort the requester, as she does not wish to distance herself from him, even though he is merely an acquaintance. Then she produces 'not dare' to further reinforce her refusal, achieving a positive outcome. This format was applied in 4.5% of three-element patterns, most frequently in situation 12 (33.3%), where a request from a female stranger was rejected.

Example 15: Sorry, lending you my car is a risk that I do not dare to take. S11(4+5+2)

'B+B+C' formatted refusal: In response to a male student's suggestion to provide additional practice activities in class during situation 7, the refuser engages in cohesive interactional activities, utilising the following preference structure to mitigate face threat from the refusee. In Example 16, 'Regret' (sorry), 'Principle' (this is my curriculum), and 'Imperative mood' articulate the refuser's position indirectly to divert the refusee's attention from the illocutionary force of the refusal. Through the structure, the outcomes of politeness and her intention not to accept the suggestion are pursued. This format was used in 5.6% of three-element patterns, with the top rate in situation 15 (27.3%), declining a female stranger's offer.

Example 16: Sorry, this is my curriculum. You need to follow this. S7(4+18+32)

'B+C+B' formatted refusal: In situation 13, where a female refuses her female lecturer's offer to drive her to school, the refuser starts with a reason description, uses the hedging 'I am afraid' in the middle, and goes back to reason description as an end in Example 17. The mitigated interactional sequence serves as a strategy to close off further expectations regarding the offer. The three-track indirectly supports the refuser's rejection. This is useful in avoiding inflicting damage that cannot be avoided on the interpersonal relationship. This format was used in 17.2%

of three-element patterns, most in situation 3 (20.6%), addressing a gender-unspecified acquaintance's request.

Example 17: I go to school by Grab with my good friend. I am afraid my friend cannot afford the high taxi fees. S13(5+21+5)

'C+A+B' formatted refusal: To refuse a female student's suggestion for bringing the additional class next Friday forward to finish the syllabus in situation 6, the refuser considers each move carefully so that the refusee understands her stance. The 'Empathy' (I understand your feelings…) and 'Negative proposition' (I cannot) structure provides a trajectory for the subsequent progression of refusal in Example 18. The final 'reason' reinforces the refusal by specifying an action, offering the refusee an effective exit strategy. The three factors are a key part of the glue that binds the refuser's actions together in coherent sequences to project deference and demeanour. This format was employed in 1.0% of three-element patterns, all in situation 6, refusing a female stranger's suggestion.

Example 18: I understand your feelings very well, but I cannot because changing the course is not a simple thing. I need to find out if there is an empty classroom available and if it is approved. S6(26+2+5)

Multi-element Preference Structures

These preference structures contain more than three sub-strategies. This format still includes recycled parts, if any, for instance, 'C' in the format of 'C+C+C+B'. They are represented in 38 types, including 'A+C+C+B', 'C+A+B+B', 'B+B+A+B+C', and 'C+C+C+A+B+B'. Both 'B+C+B+B' and 'C+C+B+B' had the highest proportion (11.1%).

'A+C+C+B' formatted refusal: the format appeared once in situation 14, where a male friend tends to borrow RM200 to buy a birthday present for his roommate and then make an alternative suggestion, the participant selects a direct strategy ('no'), 'thanks' (a convivial marker) and an intensifier 'actually' (a booster) to lessen the imposition on both interlocutors in Example 19. Then, she provides the reason, 'I am not familiar with him'. The refuser is not confined by her underdeveloped pragmatic competence in refusal production. Instead, she opts for effective redressive means to counteract the potential face damage to the refusee.

Example 19: No thanks, but I am not familiar with him. S14(1+23+34+5)

'C+A+B+B' formatted refusal: this format occurred once in situations 1, 9, and 16, respectively, accounting for 4.2% of the total. For example, in situation 1, where a female refuses her male friend who invites her to watch an action movie, the refuser selects a four-element preference structure to reject in Example 20. The prefatory 'oh' and direct strategy 'no' are mobilised as a counter to the refusee's expectation; the reason (I am watching the movie I like) and wish (I do not want to watch anything else) are assembled to further the possibility of refusing. The sequence and directionality of the refusal towards 'success' are key indicators for identifying preference structures to remedy offence.

Example 20: Oh no, I am watching the movie I like; I don't want to watch anything else. S1(25+1+5+17)

'C+C+C+B' formatted refusal: this format occurred once in situations 2, 5, and 13, respectively, and represents 4.2% of the total. To handle the refusal in situation 13, where participants are required to refuse a female lecturer who lives in the same neighbourhood and offers a ride to school, the participant uses 'Gratitude', 'Address term', 'Compliment' and 'Reason' in two domains (C and B) to shape the refusal in Example 21. The refuser begins with a gesture of gratitude, saying 'thank you so much,' to defuse potential tension. She does not immediately

produce a curt and abrupt refusal. Instead, she uses phrases such as 'dear lecturer' and 'you are so nice', and provides a reason (my classmates have invited me to attend class together) to soften the refusal, appearing caring and bounteous. This preference sequence allows interlocutors to maintain social harmony.

Example 21: Thank you so much, dear lecturer. You are so nice. However, my classmates have invited me to attend class with them. S13(23+28+22+5)

Discussion

The research offers a fresh perspective on the sequential dynamics of refusals, direct, indirect, or adjuncts, all of which contribute to mitigating potential conflicts. The research questions are restated here for convenience.

RQ1: What refusal strategies do Chinese female international students use?

RQ2: How do Chinese female international students sequentially organise their refusals as preference structures?

The study reveals that female participants possess a strong capability to engage in social interactions in refusals, ranging from one-element to complex multi-element forms. These strategies enable them to communicate their intentions effectively and minimise cognitive load. All female participants used various direct sub-strategies, such as 'Negative proposition' and 'Flat no', despite their bluntness and lack of politeness. These findings are consistent with previous research on sequence organisations (Bella, 2014; Taguchi, 2013) and second language refusals (Al-Gahtani & Roever, 2018). The prevalent use of indirect strategies among learners aligns with the findings of Al-Gahtani and Roever (2018), Bella (2014), and Taguchi (2013), which indicate that learners often perform their refusals by providing reasons. Within the same category, there existed two or more distinct sub-categories, such as 'Flat no' and 'Gratitude', which is similar to the findings by Scatter et al. (2011). 'Negative proposition' ('will not' and 'could not') co-occurred with adjuncts such as 'Positive feeling' ('Yes but' or 'okay but' structure) to form a preference structure. This structure might indicate that interlocutors realise the speech act of refusal indirectly and demonstrate a more substantial capacity to maintain social unity. Nevertheless, this contradicts Kentmen et al. (2023), who observed that female learners hardly ever utter the word 'no'. Females tend to employ 'no', possibly because they have become more assertive and independent-minded with the generation shifts. Another reason may be that they wish to clarify their understanding so they can better adapt to the academic context in Malaysia. The use of 'Apology' and 'Gratitude' is compatible with the discovery by Soutar and Mitchell (2018), who noted that expressive utterances in the form of 'thank you' and 'apologies' are chosen by female students in their research. 'Hedging' and 'Pause filler' displayed a sense of indirectness through delayed thinking. However, this case contradicts the finding of Sattar et al. (2011), in which no adjunct served as the head of the refusal. All the discoveries are consistent with those by Morkus (2014), who advocates the view that refusal strategies are recycled. The striking point is that the participants even have recourse to 'Curse' after 'Flat no' to refuse offers. This is compatible with Pan's (2012) view that indirect strategies (such as 'curse') increase the likelihood of being perceived as uncooperative or even impolite from the perspective of the hearer. This impolite phenomenon may reflect a shift in values in the workplace, from valuing respect for authority to prioritising individual achievement. With the rise of liberal ideas, people have now become more focused on individual fulfilment. Generation Z is rewriting the rules of the workplace (Racolţa-Paina & Irini, 2021). Politeness is less about conformity and more about

respect for individual autonomy.

Females tend to employ multi-element patterns. The combined use of two or more of the sub-strategies increases the range of lexical tools recruited for refusal formatting and reflects obvious lengthiness in refusal formulation. Although some direct strategies were used, the imposition was mitigated by indirect strategies and adjuncts, which resemble those by Elsweiler (2022). Students' orientation towards the 'waffling' phenomenon might stem from insufficiency in proficiency or pragmatic skills, constrained by limited exposure to an authentic English context. Furthermore, L1 transfer due to insufficient exposure to native English phrases and idioms renders them unconfident and hesitant, leading to verbosity as a form of psychological support in communication (Edmondson & House, 1981). The enhanced interactional competence of the learners is evidenced by their broader selection of refusal preference sequences, incorporating a combination of sub-strategies such as 'Negative proposition', 'Flat no', 'Regret', and 'Gratitude'. These strategies allow them to customise their refusals to specific contexts and interlocutors, thereby conveying politeness. The data reflect the refusal recipient designs proposed by Al-Gahtani and Roever (2018), who integrated various resources to structure sequences and prevent miscommunication. Beyond L1 transfer, the Malaysian culture and social norms, which value harmony and indirectness, might prompt students to use softening means. The use of composite patterns (C+A+B, C+A+B) suggests females attempt to prioritise adjuncts as heads of refusals to avoid detriment to interpersonal relationships. The findings of this study support Brown and Levinson's (1987) assertion that both positive and negative faces are crucial concerns, encouraging females to adopt more diverse means for softening refusals. The finding demonstrates that interlocutors are all better able to express their commitment to maintaining social cohesion and avoiding disaffiliation. These semantic elements, packaged in facework, represent a form of the female participants' psychological attitude towards the interlocutors. This suggests that Chinese culture inextricably shapes the widespread use of strategies. The gender expectations, such as consideration, delicateness, and sensitiveness, motivate them to minimise potential face threat from refusals. The case also mirrors the finding by Kentmen et al. (2023), in which female participants prefer gentle politeness as a way of refusing. The politeness investment may be held accountable for reconciling the conflicting goals of clarity and politeness use.

Conclusion

This study aims to empirically scrutinise how female learners effectively organise preference structure in refusal production via Discourse Completion Tests (DCTs). The sequences serve as a focal point for negotiating refusals. The female learners' simultaneous use of multiple sub-strategies within preference structures is attributed to their diverse social objectives, as well as their extensive familiarity with a repertoire of strategies for conveying politeness. When realising refusals, learners tend to gravitate towards specific sequence structures. In other words, female learners prioritise specific structures above others to achieve their communicative goals and avoid compromising the balance between clarity and politeness.

Although the data pool is limited and the results from females cannot be generalised to all, this study has significant importance for research into the growth of the complex L2 discursive environment, as well as for pedagogical implications in intercultural communication contexts. The study provides data on preference structures to mitigate refusals. This is significant in providing a reference for both female and male Chinese international students in Malaysia to deal with daily communication. Further research into the development of pragmatic competence among English

learners, particularly concerning various social behaviours, would be both compelling and beneficial. Additionally, the findings support the classification of different sub-strategies for refusals, thereby opening avenues for further exploration into more types of refusal strategies. Given that non-verbal cues in the production of speech acts remain unexplored, it is highly recommended that researchers address these pragmatic implications using complementary methods such as role-plays to capture non-verbal cues.

Funding: The research is not funded.
Acknowledgements: Not applicable
Conflicts of Interest: The authors declare no conflicts of interest.
Authenticity: This manuscript is an original work.
Artificial Intelligence Statement: Grammarly was used only for grammar correction and language editing. The authors take full responsibility for the integrity and accuracy of the content.

About the Authors:
He Zhai is currently pursuing her PhD in linguistics at Universiti Malaya's Faculty of Languages and Linguistics. ORCHID ID-https://orcid.org/0009-0001-1127-9751

Ridwan Wahid (PhD) is an associate professor of linguistics at Universiti Malaya's Faculty of Languages and Linguistics. ORCHID ID-https://orcid.org/0000-0002-5228-950X

Woo Wai Sheng (PhD) is a senior lecturer in Japanese Language at Universiti Malaya's Faculty of Languages and Linguistics. ORCHID ID-https://orcid.org/0000-0003-0074-9110

References

Abbas, N. F., & Chee, M. H. (2023). Refusal and politeness strategies favoured among Iraqi and Malaysian learners in marriage proposals. *Discourse and Interaction, 16*(2), 29–45. https://doi.org/10.5817/DI2023-2-29

Abed, A. Q. (2011). Pragmatic transfer in Iraqi EFL learners' refusals. *International Journal of English Linguistics, 1*(2), 166. https://doi.org/10.5539/ijel.v1n2p166

Al-Gahtani, S., & Roever, C. (2018). Proficiency and preference organisation in second language refusals. *Journal of Pragmatics*, 129, 140–153. https://doi.org/10.1016/j.pragma.2018.02.008

Alalawi, B. (2022). The Role of Gender on Saudi English Language Learners' Performance of Refusals. *Arab World English Journal, 13* (3) 220–235. https://doi.org/10.24093/awej/vol13no3.14

Albirini, A. (2016). Modern Arabic sociolinguistics: Diglossia, variation, codeswitching, attitudes and identity. Routledge. https://doi.org/10.4324/9781315714043

Ali Qassim, T., Abbas, N. F., Ahmed, F. F., & Hameed, S. (2021). Pragma-linguistic and socio-pragmatic transfer among Iraqi female EFL learners in refusing marriage proposals. *Arab World English Journal, 12*(1), 419–437. https://doi.org/10.24093/awej/vol12no1.28

Allami, H., & Naeimi, A. (2011). A cross-linguistic study of refusals: An analysis of pragmatic competence development in Iranian EFL learners. *Journal of Pragmatics, 43*(1), 385–406. https://doi.org/10.1016/j.pragma.2010.07.010

Beebe, L. M., Takahashi, T., & Uliss-Weltz, R. (1990). Pragmatic transfer in ESL refusals. In R. C. Scarcella, E. S. Andersen, & S. D. Krashen (Eds.), *Developing communicative competence in a second language* (pp. 55–73). Newbury House.

Bella, S. (2014). Developing the ability to refuse: A cross-sectional study of Greek FL refusals. *Journal of Pragmatics, 61*, 35–62. https://doi.org/10.1016/j.pragma.2013.11.010

Blum-Kulka, S., & Olshtain, E. (1984). Requests and apologies: A cross-cultural study of speech act realisation patterns (CCSARP). *Applied linguistics*, *5*(3), 196–213. https://doi.org/10.1093/applin/5.3.196

Boonsuk, Y., & Ambele, E.A. (2019). Refusal as a Social Speech Act among Thai EFL University Students. *Arab World English Journal, 10*(2), 213–224. https://doi.org/10.24093/awej/vol10no2.17

Brown, P., & Levinson, S. C. (1987). Politeness: Some universals in language usage. Cambridge University Press.

Edmondson, W. J., & House, J. (1981). *Let us talk, and talk about it: a pedagogic interactional grammar of English*. Urban & Schwarzenberg.

Elsweiler, C. (2022). Gender variation in the requestive behaviour of Early Modern Scottish and English letter-writers? A study of private correspondence. *Journal of Historical Sociolinguistics*, *8*(1), 55–88. https://doi.org/10.1515/jhsl-2022-2004

Félix-Brasdefer, J. C. (2006). Linguistic politeness in Mexico: Refusal strategies among male speakers of Mexican Spanish. *Journal of Pragmatics*, *38*(12), 2158–2187. https://doi.org/10.1016/j.pragma.2006.04.006

Golato, A. (2017). Naturally occurring data. In *The Routledge Handbook of Pragmatics* (pp. 21–26). Routledge. https://doi.org/10.4324/9781315737349

Kentmen, H., Debreli, E., & Yavuz, M. A. (2023). Assessing tertiary Turkish EFL learners' pragmatic competence regarding speech acts and conversational implicatures. *Sustainability, 15*(4), 3800. https://doi.org/10.3390/su15043800

Kwon, J. (2004). Expressing Refusals in Korean and American English*: Multilingual Matters*, *Kwon, J. (2004). Expressing refusals in Korean and American English* (23), 339–364.

Liao, C. H., & Bresnahan, M. J. (1996). A contrastive pragmatic study on American English and Mandarin refusal strategies. *Language Sciences*, *18*(3–4), 703–727. https://doi.org/10.1016/S0388-0001(96)00032-3

Lim, L. (2014). Cross-cultural communication and pragmatic failure in academic interactions: The case of international students in Malaysia. *Journal of Intercultural Communication*, (35), 1–15.

Morkus, N. (2014). Refusals in Egyptian Arabic and American English. *Journal of Pragmatics, 70*, 86–107. https://doi.org/10.1016/j.pragma.2014.06.009

Németh T, E. (2004). The principles of communicative language use. *Acta Linguistica Hungarica*, *51*(3-4), 379–418. https://doi.org/10.1556/ALing.51.2004.3-4.6

Olshtain, E., & Blum-Kulka, S. (1983). Cross-linguistic speech act studies: Theoretical and empirical issues. *Languages across cultures*, 235–248.

Pan, Y. (2012). Facework in refusals in Chinese survey interviews. *Journal of Politeness Research*, *8*(1), 53–74. https://doi.org/10.1515/pr-2012-0004

Racolţa-Paina, N. D., & Irini, R. D. (2021). Generation Z in the workplace through the lenses of human resource professionals–A qualitative study. *Calitatea*, *22*(183), 78–85.

Ren, W., & Woodfield, H. (2016). Chinese females' date refusals in reality TV shows: Expressing involvement or independence? *Discourse, Context & Media*, 13, 89–97. https://doi.org/10.1016/j.dcm.2016.07.003

Scatter, A., Chelah, S., & Suleiman, R. (2011). Refusal Strategies in English By Malay University Students. *GEMA Online™ Journal of Language Studies Volume 11 (3)*, 1–13.

Soutar, I., & Mitchell, C. (2018). Towards pragmatic narratives of societal engagement in the UK energy system. *Energy research & social science*, *35*, 132–139. https://doi.org/10.1016/j.erss.2017.10.051

Taguchi, N. (2013). Refusals in L2 English: Proficiency effects on appropriateness and fluency. In N. Taguchi & J. M. Sykes (Eds.), *Refusals in instructional contexts and beyond* (pp. 101–119). Brill. https://doi.org/10.1163/9789004253204_007

Taguchi, N., & Roever, C. (2020). *Second language pragmatics*. Oxford University Press.

Takahashi, T., & Beebe, L. M. (1987). The development of pragmatic competence by Japanese learners of English. *JALT journal*, *8*(2), 131–155.

Wardhaugh, R. (2006). Sociolinguistics. *Victoria: Blackwell Publishing*.

Wong, J., & Waring, H. Z. (2020). *Conversation analysis and second language pedagogy: A guide for ESL/EFL teachers*. Routledge. https://doi.org/10.4324/9780429295376

Wu, J., & Roever, C. (2021). Proficiency and preference organisation in second language Mandarin Chinese refusals. *The modern language journal*, *105*(4), 897–918. https://doi.org/10.1111/modl.12708

Appendices

Appendix A: Sixteen Situations in DCT

Situation No.	Eliciting Acts	Situation Prompts	Social Distance
1	suggestion	A male friend's suggestion to watch an action movie	acquaintance
2	suggestion	A female friend's suggestion to buy a practical book	acquaintance
3	request	A boss's request to spend extra time working	acquaintance
4	suggestion	A student's suggestion to postpone a quiz	stranger
5	suggestion	A sister's suggestion to buy a cheaper skirt	intimate
6	suggestion	A female student's suggestion to bring the class forward	stranger
7	suggestion	A male student's suggestion to give more practice	stranger
8	offer	A male boss's offer to fix a car in his brother's workshop	acquaintance
9	request	A male colleague's request for help in handling the class	acquaintance
10	offer	A male taxi driver's offer to his son's hotel	stranger
11	request	A male friend's request to borrow a car	acquaintance
12	request	A female student's request to borrow notes	stranger
13	offer	A female lecturer's offer to take a lift home	stranger
14	suggestion	A male friend's suggestion to borrow money from others	acquaintance
15	offer	A female student's offer to take a lift home	stranger
16	request	A mother's request to do homework	intimate

Appendix B: Refusal Substrategies by Females

Direct Sub-strategies	Indirect Sub-strategies	Adjuncts Sub-strategies
1 Flat no	4 Regret	21 Hedging
2 Negative proposition	5 Reason	22 Compliment
3 Performative	6 Alternative	23 Gratitude
	7 Future acceptance	24 Positive feeling
	8 Negative consequences	25 Pause filler
	9 Criticism	26 Empathy
	10 Letting off the hook	27 Non-verbal strategies
	11 Proverb	28 Address terms
	12 Topic switch	29 Curse
	13 Questioning	30 Polite remarks
	14 Unspecific reply	31 Parenthesis
	15 Asking for approval	32 Imperative mood
	16 Sarcasm	33 Greeting
	17 Wish	34 Intensifier
	18 Principle	35 Complaint
	19 Lecture	36 Comfort
	20 Self-defence	

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.301-314
DOI: https://dx.doi.org/10.24093/awej/vol16no3.18

Unpacking Learner Autonomy: Instructor Insights, Sustainable Benefits, and Challenges in Tertiary EFL Context

Lara Arafat
Center for Preparatory Studies, Sultan Qaboos University
Muscat, Sultanate of Oman
Corresponding Author: larafat@squ.edu.om

Zeinab Al Khshali
Center for Preparatory Studies, Sultan Qaboos University
Muscat, Sultanate of Oman
Email: khshali@squ.edu.om

Received: 06/19/2025 Accepted:09/03/2025 Published: 09/20/2025

Abstract
In a rapidly changing world, the field of English as a Foreign Language (EFL) is undergoing a significant transformation, with increasing emphasis on learner autonomy (LA). Theories of pedagogy have been maturing and developing at a fast pace, bringing increased attention to the mosaic of perspectives surrounding LA. In view of its sustainable benefits, EFL instructors at Sultan Qaboos University (SQU) have made efforts to cultivate self-directed learning in supportive educational environments. Although SQU has made these efforts, promoting a culture of autonomy remains a significant challenge. This study explores English language instructors' perceptions of learner autonomy at SQU, with a focus on its sustainable benefits and implementation challenges. It examines how instructors define and approach learner autonomy, the academic and personal advantages they associate with it, and the obstacles they encounter in fostering an LA environment. To gain insights into learner autonomy, the researchers interviewed twenty instructors from the Center for Preparatory Studies (CPS) at SQU. The findings reveal that most LA perceptions fall within a similar conceptual range. The study also highlights both the benefits of promoting learner autonomy for students and the ongoing difficulties instructors face in its implementation. The study concludes by recommending further research from the learner perspectives to deepen the understanding of LA, its perceived value, and the barriers to its adoption in higher education institutes (HEIs), as well as to examine the positive impact of blended learning on learners.

***Keywords*:** Center of Preparatory Studies, English as a foreign language, higher education institutes, instructor perceptions, learner autonomy, pedagogical challenges, self-directed learning, Sultanate Qaboos University, sustainable benefits

How to cite: Arafat, L., & Khshali, Z. (2025). Unpacking Learner Autonomy: Instructor Insights, Sustainable Benefits, and Challenges in Tertiary EFL Context. *Arab World English Journal*, *16*(3): 301-314. https://dx.doi.org/10.24093/awej/vol16no3.18

Introduction

Autonomous learning, as widely defined by language educators, refers to the competence to take responsibility for learning (Holec, 1981). According to Masouleh and Jooneghani (2012), learner autonomy is a dynamic process that varies in different learning contexts and often requires scaffolding by teachers to facilitate the learning process. They also underscore the importance of induction sessions and continuous support to help students become familiar with independent learning tools, materials, and resources, which not only contribute to student academic and personal development but also foster lifelong learning.

Despite its widely recognized advantages, LA remains a complex, multi-dimensional construct, shaped by a range of factors that go beyond student motivation or pedagogical approaches. Dang (2012) points out that instructors have explored numerous strategies to foster learner autonomy in foreign or second language education, accompanied by a range of theoretical frameworks. Nevertheless, significant obstacles to the effective implementation of autonomy-oriented pedagogy in higher education often arise due to the diversity in student learning styles and rates of progress. Al-Saadi (2011) notes that many students enter university settings without prior exposure to autonomy-supportive learning environments, mainly due to school systems that do not adequately equip them with the skills and strategies needed to become independent learners. Integrating learner autonomy into university curricula continues to be a considerable challenge for educators.

This study investigates the perceptions of English language instructors at SQU regarding learner autonomy, with a particular focus on its sustainable benefits and the challenges faced in cultivating an autonomy-supportive learning culture. It seeks to answer the following research questions:

What are instructor perceptions of learner autonomy, including its levels and modes of learning?

What are the benefits of learner autonomy at both academic and personal levels?

What key challenges do instructors face in promoting and implementing learner autonomy?

Literature Review

Educational theories and pedagogies have been continuously evolving in the field of English language education (Abilasha & Ilankumaran, 2014; Eaton, 2010). Teaching models are gradually changing from traditional teacher-led to more collaborative student-focused methods (Abilasha & Ilankumaran, 2014; Jacobs & Farrell, 2001). According to Soontiens (2004), the 21st-century students are motivated to learn a language not solely for communication purposes, but also to build networks, connect with others, and form collaborative relationships. Abilasha and Ilankumaran (2014) further emphasize that the latest teaching and learning trends focus on quality rather than quantity. These findings necessitate a shift in current practices to support meaningful improvement (Abilasha and Ilankumaran, 2014). English language instructors recognize that the twenty-first century has brought with it a bundle of educational trends to help students not only learn but also showcase their knowledge tangibly using their innate and acquired interpersonal skills to become self-empowered (Eaton, 2010). The priority now in the educational field is how students acquire sustainable knowledge that they can apply to both personal and career life.

Henri Holec, widely considered the father of learner autonomy, identified it as a significant trend in the EFL field. According to Smith (2015), the concept has received increased attention in recent years, yet it is not entirely new. Autonomous learners independently manage learning, take

full responsibility for the learning process and rely on themselves rather than external sources (Smith, 2015). Dickinson (1995) emphasizes that independent learners tend to be more motivated and produce higher-quality, more effective work. They are also responsible for organizing, arranging, and supervising their academic process (ElFalih & Lamrani-Hassani, 2024). Developing soft skills such as critical thinking, creativity, and problem-solving is essential for students to enhance LA, as they help students succeed in the modern, 21st-century world (Olugbenga, 2021). Such a perspective highlights the pressing challenges associated with implementing learner autonomy, particularly in contexts where traditional, teacher-centered approaches have long prevailed. Two key challenges relate to the students themselves and the academic institutions. First, the students lack a sense of autonomy in language learning (Masouleh & Jooneghani, 2012). Breen and Mann (1997) add that a key determinant of learner autonomy is the learner readiness to engage, especially since learner autonomy involves self- exploration. Olugbenga (2021) emphasizes that a shift in the curriculum from a traditional to a progressive approach is necessary to facilitate the implementation of learner-centered practices.

Recognizing both the advantages and challenges of learner autonomy, and in alignment with the strategic direction outlined by the Oman Vision 2040 Implementation Follow-up Unit (n.d.) as well as the vision, mission, and values of Sultan Qaboos University (n.d.-b), which emphasize the importance of preparing graduates with skills for the 21st-century labor market, the CPS at SQU has implemented active learning pedagogies to promote autonomous learning. As Thanasoulas (2000) emphasizes, fostering learner autonomy requires students to take responsibility for and manage their learning. To support this effort, initiatives such as the Coursera Project and the Mahrat (Potentiel.ly) Platform were introduced in 2023, offering students opportunities to develop both academic and personal portfolios that showcase their progress and achievements.

Method

Context Setting

The researchers conducted the study at the CPS, SQU, where instructors comprise both Omani and non-Omani faculty members who teach in the General Foundation Program (GFP). The GFPs, in Oman, prepare high school graduates with the skills needed for successful entry into higher education. The Oman Authority for Academic Accreditation and Quality Assurance (n.d.) states that programs offered by various HEIs aim to develop student competencies in four core areas: English Language, Computing, Mathematics, and General Study Skills. The CPS, the leading provider of GFP at SQU, is responsible for delivering this curriculum and ensuring that students acquire the foundational skills needed for academic success at the tertiary level.

At SQU, the CPS enrolls over 3,000 students annually and operates three academic departments: English for Humanities, English for Sciences and Mathematics, and IT. Foundation Programme English Language (FPEL) forms the most significant component of the GFP, consisting of six progressive proficiency levels, typically delivered as two levels per 15-week semester. Such a structure supports continuous language development and academic readiness. Additionally, the CPS offers English credit courses that help students develop foundational English skills relevant to their respective colleges, as well as IELTS preparation programs to expand undergraduate academic success further.

Institutional Mission

Aligned with its mission, the CPS aims to equip students with the skills needed to engage in and succeed in their undergraduate programs actively. Additionally, the CPS emphasizes delivering a high-quality educational experience that fosters academic knowledge, essential skills, and positive learning attitudes (Sultan Qaboos University, n.d.-a).

Participants

The study involved twenty English language instructors at the CPS. Teaching faculty comprises approximately 250 instructors representing around 30 countries, with a majority of Omani instructors. Diversity in professional and cultural backgrounds enriches the pedagogical approaches of the center and contributes to varied and innovative teaching strategies.

Instrument

This study employed a qualitative research approach, utilizing purposeful sampling that included twenty participants based on their research interests, years of teaching experience, and academic qualifications (Creswell, 2012). Researchers collected information through one-on-one interviews and focus group discussions.

The researchers transcribed and summarized all interviews, then analyzed the data and categorized it into three overarching themes: perceptions, benefits, and challenges. Insights from the participants were compared and corelated with existing literature in the field of learner autonomy to support the validity of the findings.

Results

Researchers organized the interview protocol into three key sections. In the first section, researchers investigated participants' overall perceptions of LA, including their definitions, the levels or degrees recognized, and views on the mode of learning. In the second section, they examined insights regarding the long-term benefits of LA, especially its impact on students' academic progress and personal development. In the final section, they explored perspectives of instructors on challenges students face in developing learner autonomy, with particular attention to the influence of prior educational experiences of students and institutional constraints.

Instructor Insights: Definition, Varying Degrees, and Modes of Learning

The participants provided varying definitions of learner autonomy (LA), which clustered into three main categories. The first group described it as the ability of learners to take charge of their learning (Holec, 1981). The second group expanded on this view by focusing on psychological aspects (Little,1991). The third group described a broader understanding of LA, incorporating technical, psychological, and political dimensions (Benson, 1997).

Participants commonly stated that learner autonomy varies in degree depending on the learners, and the educational context. A few participants noted that most learners exhibit partial autonomy based on contextual and curricular constraints rather than complete autonomy (Benson, 2006). Other participants reported the significance of implementing autonomy gradually, offering students opportunities to participate in decisions about their learning.

Most participants described the mode of autonomy as co-managed. To illustrate, while autonomy includes independent learning, it occurs in collaborative contexts where students engage with peers and instructors (Kharroubi & El Mediouni, 2024; Szőcs, 2017; Trebbi, 1990). They also

discussed the teacher's role as a facilitator who encourages learner independence while co-constructing knowledge with students (Dam, 1995; Yan, 2012).

Sustainable Benefits: Academic and Personal Growth

Instructors widely agreed that learner autonomy contributed to both academic and personal development. A number of participants reported that giving students more control over their learning led to higher levels of engagement, responsibility, and academic achievement (Saad & Abdullah, 2025; Little et al., 2000). At a personal level, participants noted that autonomous learners displayed greater resilience, confidence, and self-awareness (Lou, 2021). These participants described autonomy as a foundation for lifelong learning that prepares students for future academic and professional challenges.

Challenges to Autonomy: Student and Institutional -Related

The participants expressed concerns about the readiness of students for autonomous learning. Several participants observed that students often lacked the skills necessary to take charge of their education. The participants also noted that prior educational experiences of students (Al-Saadi, 2011; McNair, 1997) made it difficult for them to adapt to more autonomous learning models. Participants also highlighted institutional barriers that hindered the promotion of learner autonomy (Marzuki et al., 2023).

Discussion

Instructor Insights: A Tripartite Perspective

*Definition*The perceptions of the participants regarding LA fell into three broad categories. The first group agreed with Holec's (1981) foundational definition, viewing autonomy as the capacity of students to independently manage their learning. The second group found the definition by Holec (1981) insufficient, arguing that it primarily emphasized the logistical or theoretical aspects of the learning process, without addressing how learners enact autonomy in practice. Benson (2001) points out that Holec outlines what autonomous learners are capable of, but not how they achieve it. The same group emphasized the importance of integrating the conceptual framework of Holec with the focus of Little (1991) on learners' cognitive and metacognitive capacities (as cited in Jiménez Raya, Lamb, & Vieira, 2007). They also highlighted psychological aspects such as motivation and self-regulation, which Little (1991) describes as essential to autonomy, emphasizing that autonomy is evident not only in how learners acquire knowledge, but also in how they apply that knowledge in broader, real-world situations. Participants within this group also described autonomy as a personalized learning journey driven by intrinsic motivation or a natural love for learning (Dickinson, 1995). The third group expanded further by adopting a three-dimensional view of learner autonomy: technical, psychological, and political (Benson, 1997). Such a perspective suggests that autonomy involves not only independent learning and the required learner attributes, such as attitude, aptitude, and maturity, but also agency and control over what is learned (Holec, 1981). To these participants, learner autonomy includes both internal learner capacities and external control over learning content and direction (Saad & Abdullah, 2025).

Varying Degrees

The instructors interviewed in this study articulated a wide range of perspectives on the degrees or levels of LA observed and practiced in their classrooms. Findings from the interviews

suggest that LA is not a fixed concept but one that operates along a spectrum, rather than being a binary presence or absence. Nunan (1997) notes that autonomy exists across what he calls "degrees," a view that echoed throughout the interviews. Several participants emphasized the importance of adjusting lesson plans to offer students increased opportunities to express opinion and make selections (Pennycook, 1997). One participant articulated that this represents a gradual, bottom-up approach to developing autonomy, in which instructors progressively guide students to assume greater responsibility for their learning. Another instructor noted that this process prepares learners to act as stakeholders in their educational journey, which reinforces the idea that autonomy is developmental, gradual, and contextual.

The influence of learner characteristics and the educational context was also frequently mentioned. One participant with a PhD explained that learner autonomy can take many forms. According to Little (1991), the forms of learner autonomy are shaped by factors such as learners' age, needs, motivation, and the learning context. This complexity suggests that autonomy is situationally dependent, as Dickinson (1987) asserts that the learning situation can impose practical and psychological limitations on full independence.

Participants generally agreed that full autonomy, though ideal, is often difficult to achieve in practice. Holec (1981) describes genuine autonomy as self-managed learning, where the learner independently sets goals, selects materials, monitors progress, and evaluates outcomes. Theoretically, this level of autonomy allows learners to make decisions about what, how, when, and why to study (Holec, 1981); however, most instructors acknowledge that such autonomy is rare and typically emerges only under favorable conditions, such as a supportive curriculum and high learner motivation. In addition, as McNair (1997) notes, many students enter higher education with limited experience in autonomy due to past educational systems that did not encourage independent learning. Consequently, instructors see their role as gradually facilitating the development of autonomy, acknowledging that partial autonomy is often a more realistic and contextually appropriate goal.

Mode of Learning

When asked about the mode of learning associated with LA, whether it is primarily social or individual, and the role of the teacher, most participants described it as collaboratively managed (Holec, 1981). Trebbi (1990) also highlights the essential role of the social dimension in autonomous learning, which takes place collaboratively with others. The participants emphasized that learning is not a process of passively receiving information, but rather one that involves the active construction of meaning, which is often co-constructed through interaction with peers and instructors (Kharroubi & El Mediouni, 2024; Szőcs, 2017).

The participants agreed that the role of teachers in supporting LA is not to dominate or control the learning process, but to stimulate the interest of learners, assist in navigating difficult decisions, and support students in overcoming obstacles throughout their learning journey (Dam, 1995; Yan, 2012). The teacher acts as a facilitator, guiding students toward greater independence while maintaining a collaborative learning environment.

Ultimately, defining learner autonomy, exploring its degrees, and identifying its preferred mode of learning remain controversial areas within the field. Benson (2006) notes that these aspects of autonomy are often subjective, shaped by individual experiences, institutional contexts, and broader educational philosophies.

Sustainable Benefits: Intersecting Academic and Personal Development

Colleges and universities worldwide aim to equip students with the essential skills and experiences needed to thrive in the current century. Recent empirical studies highlight the positive impact of LA on language proficiency and academic performance, particularly at the higher education level. Autonomous learning is increasingly viewed as intrinsically rewarding due to its sustainable and long-term approach to education, empowering students to take control of their learning journey and grow both academically and personally (Mısır, Koban Koç, & Koç, 2018). The Sultanate of Oman emphasizes the importance of sustainable and autonomous learning through the vision of holistic education, which seeks to equip learners with the knowledge, values, and skills needed for current and future challenges (General Secretariat of the Education Council, 2017, p.23). Participants in this study strongly agreed that learner autonomy supports student responsibility and discipline in both academic and personal contexts, which eventually contributes to lifelong learning and success.

Academic Growth

Learner autonomy plays a critical role in improving academic outcomes. The study participants observed that autonomy supports academic discipline and responsibility, resulting in increased satisfaction, motivation, and achievement, as highlighted by Saad and Abdullah (2025). Active participation helps students develop self-confidence and accountability, which enhances their academic experience.

The three-phase study by Nguyen (2012), published in *New Zealand Studies in Applied Linguistics*, confirms a positive relationship between learner autonomy and language proficiency. According to Little et al. (2000), learners demonstrated clear academic improvement when they attended class regularly, participated actively, and took responsibility for their learning. ElFalih and Lamrani-Hassani (2024) also emphasize that autonomy-supportive learning environments lead to more engaged and independent learners.

Participants believed that involving learners as active participants in the learning process fosters a collaborative partnership with academic institutions. Such involvement empowers learners to set personal goals, select appropriate learning materials, manage their time effectively, and choose learning environments that align with their individual needs. These results align with the perspectives of Nunan (1997), Cotterall (1999), and Benson (2001), all as cited in Balçıkanlı (2010), who emphasize that goal setting is crucial for meaningful and personalized learning.

Furthermore, researchers such as Al-Saadi (2011) and Nupus et al. (2021) advocate recognizing students as co-agents in the learning process, particularly in second-language contexts where autonomy helps learners communicate more confidently. Participants also supported the view of Tuan and Chi (2020) that learners should be able to identify their needs, strengths, and weaknesses, monitor their progress, and evaluate their outcomes. The ability to reflect on and direct their learning enhances academic independence and long-term success. Participants in this study also emphasized that, although integrating learner autonomy into the curriculum is crucial, it requires curriculum reform, particularly addressing overloaded curriculum, to ensure learners gain the maximum academic benefit (ElFalih & Lamrani-Hassani, 2024). Other participants echoed that students who manage their learning are more inclined to apply academic skills in real-life situations and embrace lifelong learning as underscored by Nupus et al. (2021) and Alonazi (2017).

Personal Growth

Participants in this study emphasized the significant role of learner autonomy in promoting personal development. Participants generally believed that students should take ownership of their learning as this fosters self-discipline, commitment, and responsibility. A participant emphasized that engaging students in autonomous learning helps students strengthen self-awareness, refine goal setting abilities, and become more adaptable. Lou (2021) highlights that learner-centered approaches replace traditional cramming methods, encouraging students to collaborate and think critically. Previous research indicates that increased control over learning fosters higher levels of self-determination and motivation (Chan, 2001). Most participants generally observed that autonomous learners are more resilient and better equipped to navigate personal and academic challenges.

Moreover, autonomy fosters metacognitive engagement. One participant noted that learners become active by thinking autonomously, recognizing their responsibilities, making informed choices and becoming aware of the strategies they employ. Learners begin to personalize their approaches, adapting their learning strategies to suit their unique styles (Alonazi, 2017). Another participant explained that students who engage in autonomous learning are better able to appreciate the relevance of their studies, develop critical thinking skills, and become adaptable and self-sufficient. Participants further highlighted that autonomy nurtures lifelong learning by developing the mindset needed to update knowledge and develop new skills in a fast-changing world. It also promotes ongoing self-discovery and personal growth outside formal education.

In brief, learner autonomy contributes to both academic and personal growth, positioning it as a key element in sustainable education. Learners who take ownership of their learning develop critical skills such as responsibility, resilience, and self-efficacy, which help them become independent professionals committed to lifelong learning (Gupta et al., 2024) and success in the rapidly changing digital world. Although the role of instructors remains indispensable, it is ultimately the learners who must take responsibility for their educational journey. Noura and Yusuf (2023) note that those who experience autonomy in learning are more likely to continue growing, succeed academically, and thrive beyond the classroom.

Overwhelming Challenges: A Two-Dimensional Perspective

Autonomous learning, while pedagogically sound and increasingly advocated for, faces substantial barriers that are both student-related and institutionally driven. Oxford English Dictionary (Oxford University Press, 2025) defines a challenge as a task or situation that is either new or difficult and requires a person to demonstrate skill or ability to handle it. Participants in the current study offered rich insights into the fundamental challenges obstructing the successful implementation of LA. The interview questions explored these challenges across two dimensions: those related to the learners themselves and those embedded in institutional frameworks.

Student-Related Challenges

The participants, as reported by Al-Saadi (2011), emphasized that innate abilities of learners and their readiness for autonomous learning are critical to the success of the LA approach. They reached consensus that students must acquire both soft skills and metacognitive competencies to thrive autonomously. Learners need to demonstrate traits such as self-awareness (Dickinson, 1995), motivation, discipline, and confidence. They must also be metacognitively

prepared to plan, monitor, evaluate, self-regulate, and critically reflect to solve problems effectively.

A central question raised was whether students possess the maturity and mindset required to navigate their learning path, guided by their learning compass as autonomous learners. Based on their teaching experience, participants offered practical and realistic responses. As noted by Al-Saadi (2011, p.95), a participant observed that students are often unprepared for autonomy due to educational systems that "spoon-feed" them, limiting their ability to make independent decisions. Another participant reinforced this perspective by stating that students are not metacognitively ready to take on the responsibility of learning how to learn. Participants also highlighted that student engagement and investment in learning are essential factors. According to Nguyen and Nguyen (2020), learners need to stay active in their learning and feel comfortable asking questions. A few participants remarked that learners often displayed little interest in the material, which hindered the development of autonomy. Other participants pointed out that students lack the cognitive skills needed to plan, perform, control and correct their work in diverse communicative tasks, such as understanding, interacting, producing, and mediating (Dmitrenko et al., 2021).

Institution-Related Challenges

Challenges stemming from the learners themselves often sparked discussions about the institutional structures that further complicate the promotion of autonomy. Marzuki et al. (2023) describe these as related institutional obstacles. Participants identified prescriptive academic environments that hinder student self-actualization and the development of autonomy.

The participants emphasized that student perceptions of learner autonomy originate from their past educational experiences and the guidance they previously received. Several participants agreed that the hierarchical nature of the education system restricts opportunities for self-directed learning as Marzuki et al. (2023) noted. One participant stated that administrators set learning goals, outcomes, and assessment methods, instructors interpret these directives and students comply. The structure fosters a grade-oriented culture where success is measured primarily through exams, which often account for up to 75% of final grades, rather than through meaningful learning. Therefore, students tend to mimic expected answers instead of exploring their learning priorities or developing independent methods. One participant argued that imposing prescribed course content, learning materials, and testing systems prevents students from practicing autonomy (Alibakhshi, 2015). The participant explained that administrators require students to produce formulaic essays following a predetermined outline, rather than encouraging creativity

All in all, two interrelated dimensions, student readiness and institutional constraints, hinder the development of learner autonomy. On one hand, many students lack the maturity, motivation, and metacognitive skills required for autonomous learning. On the other hand, institutional practices, such as rigid curricula, standardized assessments, and top-down policies, limit opportunities for student agency and self-directed learning. Addressing the challenges requires a systemic shift that supports both the personal growth of learners and the structural flexibility necessary to foster genuine autonomy. A participant concluded optimistically that, despite substantial challenges, the potential for positive outcomes is even greater.

Conclusion

This study explored instructor perceptions, the sustainable benefits, and the key challenges associated with fostering an autonomous learning environment at CPS, SQU. Contextual and

philosophical factors shape definitions and degrees of autonomy, which remain fluid (Benson, 2006). Nevertheless, learner autonomy is consistently associated with essential life skills such as responsibility, resilience, and self-motivation. Such qualities are increasingly critical for success in a dynamic world. Findings from this study reinforce the view that learner autonomy is a key element of sustainable education, contributing meaningfully to both academic and personal growth.

However, the path toward developing autonomy is not without obstacles. Findings indicate that its progress is hindered by two interconnected dimensions: students' limited readiness and institutional constraints. Some learners lack the metacognitive, emotional, and motivational preparedness needed to take responsibility for their learning (Al-Saadi, 2011). In parallel, rigid curricula, exam-oriented assessment systems, and top-down institutional policies limit opportunities for authentic learner agency and decision-making.

Implications for practice include the need for policymakers and curriculum designers at educational institutions to re-evaluate rigid academic structures and to embed autonomy-supportive practices into curricula and assessment methods. Institutions could offer flexible learning options, encourage students to make choices, and train educators to guide rather than control the learning process. Furthermore, educational institutions should consider implementing instructor training programs that cultivate student self-regulation, goal-setting, and critical thinking. The Professional Development and Academic Support Committee at the CPS regularly offers workshops that introduce a range of pedagogical approaches to promote effective learner autonomy strategies.

Finally, the findings show that there is a pressing need for future research to find feasible and suitable ways to support autonomy in different learning environments. Instructors and administrations can create learning environments that not only support academic success but also empower students for lifelong learning by addressing both learner preparedness and institutional flexibility (Saad & Abdullah, 2025). Future studies could explore how technology-assisted learning and artificial intelligence can cultivate autonomy in blended learning environments, enhancing both flexibility and overall effectiveness.

Limitations of the Study

This study identifies several limitations despite offering valuable insights into the perceptions, benefits, and challenges surrounding learner autonomy (LA). First, the study relied on a relatively small sample size of twenty participants, which, although rich in qualitative depth, may not fully represent the broader educational landscape. Additionally, findings are context-specific and may not be generalizable to all institutions or student populations, especially those with differing cultural, curricular, or policy frameworks. Another limitation concerns the data collection method as interviews provided in-depth exploration but remained inherently subjective. The inclusion of quantitative survey methods could strengthen the findings by enabling a larger sample size. Crucially, by focusing primarily on teacher perspectives, the study overlooks student voices, thereby limiting a comprehensive evaluation of learner readiness and autonomy from the learner point of view.

These limitations underscore the need for further research involving comparative studies across diverse educational contexts, larger and more varied samples, and the inclusion of student perspectives to deepen and broaden our understanding of how learner autonomy can be effectively supported and implemented.

Funding: This research is not funded.
Conflicts of Interest: The authors declare no conflicts of interest.
Authenticity: This manuscript is an original work.
Artificial Intelligence Statement: During the preparation of this work, the authors used **ChatGPT (OpenAI)** to improve **readability and language**. The authors reviewed and edited the content as needed and took full responsibility for the publication after using this tool.

About the authors:

Lara Arafat is a senior instructor at the Centre for Preparatory Studies, Sultan Qaboos University, Oman, with over 24 years of teaching experience. She holds a BA in English from the University of Jordan, a PG Certificate from Coventry University, and a CELTA from Cambridge University. Her research interests include learner autonomy and student socio-academic rapport. ORCID ID: https://orcid.org/0009-0009-9686-819X

Zeinab Al Khshali is a senior instructor at the Centre for Preparatory Studies, Sultan Qaboos University, Oman. She received a Master of Professional Studies, Language Teaching Specialization from the University of Auckland, New Zealand, in 2003. She has over 20 years of teaching experience. She is interested in autonomous learning, and active learning techniques. ORCID ID: https://orcid.org/0009-0006-7515-2617

References

Abilasha, R., & Ilankumaran, M. (2014). Trends in English language teaching: A novel perspective. *International Journal on Studies in English Language and Literature*, *2*(11), 46–52. https://www.arcjournals.org/pdfs/ijsell/v2-i11/5.pdf

Alibakhshi, G. (2015). Challenges in promoting EFL learners' autonomy: Iranian EFL teachers' perspectives. *Issues in Language Teaching, 4*(1), 79–98. https://ilt.atu.ac.ir/article_3464_5f97acf1c82149f4bbcd3f411836d02e.pdf

Alonazi, S. M. (2017). The role of teachers in promoting learner autonomy in secondary schools in Saudi Arabia. *English Language Teaching, 10*(7), 183 –189. https://doi.org/10.5539/elt.v10n7p183

Al-Saadi, H. (2011). From spoon feeding to self-feeding: Helping learners take control of their own learning. *Arab World English Journal, 2*(3), 95–114. https://awej.org/images/AllIssues/Volume2/Volume2Number3Aug2011/5.pdf

Balçıkanlı, C. (2010). Learner autonomy in language learning: Student teachers' beliefs. *Australian Journal of Teacher Education, 35*(1), 90–103. https://doi.org/10.14221/ajte.2010v35n1.8

Benson, P. (1997). The philosophy and politics of learner autonomy. In P. Benson & P. Voller (Eds.), *Autonomy and independence in language learning* (pp. 1–20). Longman.

Benson, P. (2001). Teaching and researching autonomy in language learning. *Language Teaching, 34*(2), 103–117.

Benson, P. (2006). Autonomy in language teaching and learning. *Language Teaching, 40*(1), 21–40. https://doi.org/10.1017/S0261444806003958

Breen, M. B., & Mann, S. J. (1997). Shooting arrows at the sun: Perspectives on a pedagogy for autonomy. In P. Benson & P. Voller (Eds.), *Autonomy and independence in language learning* (pp. 132–149). Longman.

Chan, V. (2001). Readiness for learner autonomy: What do our learners tell us? *Teaching in Higher Education, 6*(4), 505–518. https://doi.org/10.1080/13562510120078045

Cotterall, S. (1999). Key variables in language learning: What do learners believe about them? *System, 27*(4), 493–513. https://doi.org/10.1016/S0346-251X(99)00047-0

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Pearson Education.

Dam, L. (1995). *Learner autonomy 3: From theory to classroom practice.* Authentik.

Dang, T. T. (2012). Learner autonomy: A synthesis of theory and practice. *Internet Journal of Language, Culture and Society*, *35*(1), 52–67. https://www.researchgate.net/publication/298092439_Learner_autonomy_A_synthesis_of_theory_and_practice

Dickinson, L. (1987). *Self-instruction in language learning*. Cambridge University Press.

Dickinson, L. (1995). Autonomy and motivation: A literature review. *System, 23*(2), 165–174. https://doi.org/10.1016/0346-251X(95)00005-5

Dmitrenko, N., Petrova, A., Podzygun, O., & Nikolaeva, S. (2021). Strategies in autonomous learning of professionally oriented English communication. *The Journal of Teaching English for Specific and Academic Purposes, 9*(3), 527–540. https://doi.org/10.22190/JTESAP2103527D

Eaton, S. E. (2010). *Global trends in language learning in the twenty-first century*. Onate Press.

ElFalih, T., & Lamrani-Hassani, L. (2024). Empowering learner autonomy: EFL teachers' perspectives and classroom dynamics. *International Journal of Linguistics and Translation Studies, 5*(4), 178–210. https://doi.org/10.36892/ijlts.v5i4.532

General Secretariat of the Education Council. (2017). *Philosophy of education in the Sultanate of Oman* (1st ed.). https://www.squ.edu.om/Portals/0/pdf/Philosophy%20of%20education%20oman.pdf?ver=2019-03-05-115833-697

Gupta, N., Ali, K., Jiang, D., Fink, T., & Du, X. (2024). Beyond autonomy: Unpacking self-regulated and self-directed learning through the lens of learner agency- a scoping review. *BMC Medical Education, 24*, 1519. https://doi.org/10.1186/s12909-024-06476-x

Holec, H. (1981). *Autonomy in foreign language learning*. Pergamon.

Jacobs, G. M., & Farrell, T. S. C. (2001). Paradigm shift: Understanding and implementing change in second language education. *TESL-EJ, 5*(1), 1–18. https://files.eric.ed.gov/fulltext/ED574144.pdf

Jiménez Raya, M., Lamb, T., & Vieira, F. (2007). *Pedagogy for autonomy in language education in Europe: Towards a framework for learner and teacher development*. Authentik.

Kharroubi, S., & El Mediouni, A. (2024). Towards a glocalized learner autonomy: A systematic review of Moroccan higher education. *Journal of English Language Teaching and Applied Linguistics, 6*(2), 158–166. https://doi.org/10.32996/jeltal.2024.6.2.17

Little, D. (1991). *Learner autonomy 1: Definitions, issues and problems*. Authentik.

Little, D., Dam, L., & Timmer, J. (Eds.). (2000). *Focus on learning rather than teaching: Why and how? Papers from the IATEFL conference on Learner Independence, Kraków, 14–16 May 1998*. Trinity College Dublin. https://files.eric.ed.gov/fulltext/ED441341.pdf

Lou, L. (2021). Cultivation of students' autonomous learning ability in application-oriented universities. *Theory and Practice in Language Studies, 11*(4), 2. https://tpls.academypublication.com/index.php/tpls/article/view/859/624

Marzuki, I., Indrawati, I., & Yunus, H. I. (2023). Teachers' challenges in promoting learner autonomy: The socio-cultural barriers of Indonesian EFL teachers. *Pioneer: Journal of Language and Literature, 15*(1), 119–137. https://doi.org/10.36841/pioneer.v15i1.2853

Masouleh, N. S., & Jooneghani, R. B. (2012). Autonomous learning: A teacher-less learning! *Procedia - Social and Behavioral Sciences, 55*, 835–842. https://doi.org/10.1016/j.sbspro.2012.09.570

McNair, S. (1997). *Getting the most out of HE: Supporting learner autonomy*. Department for Education and Employment.

Mısır, H., Koban Koç, D., & Koç, S. E. (2018). An analysis of learner autonomy and autonomous learning practices in massive open online language courses. *Arab World English Journal Special Issue on CALL, 4*(4), 24–39. https://doi.org/10.24093/awej/call4.3

Nguyen, C., & Nguyen, T. (2020). Non-language-major students' autonomy in learning English in Vietnam. *Arab World English Journal, 11*(3), 184–192. https://doi.org/10.24093/awej/vol11no3.11

Nguyen, T. C. L. (2012). Learners' autonomy in language learning: How to measure it rigorously. *New Zealand Studies in Applied Linguistics, 18*(1), 50–65. https://www.academia.edu/15021189/LEARNER_AUTONOMY_IN_LANGUAGE_LEARNING_HOW_TO_MEASURE_IT_RIGOROUSLY

Noura, I., & Yousuf, H. (2023). Enhancing learners' autonomy in vocabulary learning in EFL Libyan classroom. *East Asian Journal of Multidisciplinary Research, 2*(11), 4616. https://docs.google.com/document/d/17A8hvJ4zIrjOrzlowhTvs8AehwPNkA805MBkh0ls05c/edit?tab=t.0

Nunan, D. (1997). Designing and adapting materials to encourage learner autonomy. In P. Benson & P. Voller (Eds.), *Autonomy and independence in language learning* (pp. 192–203). Longman.

Nupus, H., Manalullaili, M., & Riznanda, W. A. (2021). Teachers' beliefs in learners' autonomy in English language learning: A case study at MA Masdarul Ulum. *Ta`dib : Jurnal Pendidikan Islam, 26*(1), 55–67. https://doi.org/10.19109/td.v26i1.11190

Olugbenga, M. (2021). The learner centered method and their needs in teaching. *International Journal of Multidisciplinary Research and Explorer, 1*(9), 64-69. https://ijmre.com/index.php/IJMRE/article/view/62

Oman Authority for Academic Accreditation and Quality Assurance (OAAAQA). (n.d.). *General foundation programme accreditation.* *https://www.oaaaqa.gov.om/public/external-activities/cheqa#general*

Oman Vision 2040 Implementation Follow-up Unit. (n.d.). *People and society (Pillar 1)*. In *Oman Vision 2040*. https://www.oman2040.om/pillar/1?lang=en

Oxford University Press. (2025). Challenge. In *Oxford English Dictionary*. https://www.oed.com/search/dictionary/?scope=Entries&q=challenge

Pennycook, A. (1997). Cultural alternatives and autonomy. In P. Benson & P. Voller (Eds.), *Autonomy and independence in language learning* (pp. 35–53). Longman.

Saad, S., & Abdullah, A. (2025). Exploring learner autonomy: A conceptual perspective on self-directed learning in higher education. *International Journal of Modern Education*, *7*(24), 1304–1315. https://doi.org/10.35631/IJMOE.724092

Smith, S. (2015). Learner autonomy: Origins, approaches, and practical implementation. *International Journal of Educational Investigations, 2*(4), 92–91. http://www.ijeionline.com/attachments/article/41/IJEIonline_Vol.2_No.4_2015-4-07.pdf

Soontiens, W. (2004). When in Rome: The realities of skill development in an 'Anglo' educational environment. *Journal of Research in International Education, 3*(3), 301–317. https://www.researchgate.net/publication/234769116_When_in_Rome_The_Realities_of_Skill_Development_in_an_Anglo_Educational_Environment

Sultan Qaboos University. (n.d.-a). *Mission & vision – Center for Preparatory Studies.* https://www.squ.edu.om/cps/About-Us/Mission-Vision

Sultan Qaboos University. (n.d.-b). *Mission, vision and values*. https://www.squ.edu.om/About/Mission-Vision-and-Value

Szőcs, K. (2017). Teachers' and learners' beliefs about language learning autonomy and its implications in the classroom: A mixed-method study. *Apples – Journal of Applied Language Studies, 11*(2), 125–145. https://doi.org/10.17011/apples/urn.201708233542

Thanasoulas, D. (2000). What is learner autonomy and how can it be fostered? *The Internet TESL Journal, 4*(11). http://iteslj.org/Articles/Thanasoulas-Autonomy.html

Trebbi, T. (Ed.). (1990). *Third Nordic Workshop on developing autonomous learning in the FL classroom, Bergen, August 11–14, 1989: Report* (Mimeographed report). Institutt for praktisk pedagogikk, Universitet i Bergen. https://warwick.ac.uk/fac/soc/al/research/collections/elt_archive/research_projects/dahla/archive/trebbi-1990.pdf

Tuan, V. V., & Chi, H. T. K. (2020). English language teachers' beliefs and practices to promote learner autonomy. *TNU Journal of Science and Technology, 225*(12), 93–100. https://jst.tnu.edu.vn/jst/article/view/3566/pdf

Yan, S. (2012). Teachers' roles in autonomous learning. *Journal of Sociological Research, 3*(2), 557–562. https://doi.org/10.5296/jsr.v3i2.2860

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp. 315-328
DOI: https://dx.doi.org/10.24093/awej/vol16no3.19

Needs Analysis of English-Language Use of Tourism Personnel in Xi'an, China

Boyu Pang
Faculty of Education, Universiti Teknologi MARA, Puncak Alam, Malaysia
Department of Public Courses, Shaanxi Polytechnic University, Shaanxi, China
Email: pangboyu3995@gmail.com

Faizah Binti Abd Majid
Faculty of Education, Universiti Teknologi MARA, Puncak Alam, Malaysia
Corresponding Author**:** faiza404@uitm.edu.my

Geethanjali A/P Narayanan
Faculty of Education, Universiti Teknologi MARA, Puncak Alam, Malaysia
Email: geetha@uitm.edu.my

Received:06/30/2025 Accepted:08/26/2025 Published: 09/20/2025

Abstract
This study investigated the English language needs and self-perceived proficiency of tourism personnel in Xi'an to improve the design of ESP courses for vocational education. It makes a significant contribution to bridging the gap between the current Tourism English course and the real-world work demands of China's tourism industry. A mixed-methods approach was employed. Quantitative data were gathered from 104 tour guides and office staff using a structured questionnaire, and qualitative data were obtained through semi-structured interviews with five experienced practitioners. Both instruments were self-developed and validated through standard procedures to ensure content reliability and validity. Descriptive statistics were used to examine the survey results, and thematic analysis was applied to interpret the interview transcripts. The results reported that listening and speaking are the most commonly used and valued skills, while reading and writing were considered secondary. Participants expressed confidence in performing template-based tasks; nevertheless, they reported difficulties with spontaneous communication situations, professional reading, and creative writing. It is recommended that future Tourism English courses should emphasize oral communication training, such as scenario-based practice, real-world simulations, and role-play activities. Instructors are encouraged to provide timely feedback and facilitate students to communicate more spontaneously. These findings provide practical guidance for designing more targeted and workplace-oriented English instruction for tourism professionals in similar contexts.

Keywords: English for Specific Purposes (ESP), listening and speaking skills, scenario-based instruction, social needs analysis, Tourism English, tourism personnel
How to cite: Pang, B., Majid, F.B.A., & Narayanan, G. A. (2025). Needs Analysis of English-Language Use of Tourism Personnel in Xi'an, China. *Arab World English Journal, 16*(3):315-328. https://dx.doi.org/10.24093/awej/vol16no3.19

Introduction

English, as the world's lingua franca, has become increasingly prominent in the tourism sector (Zahedpisheh et al., 2017). It has emerged as an essential tool for improving service quality and enhancing the international visitors' experiences. The significance of English in tourism has become even greater over recent years due to the growing demand for global mobility in the post-pandemic era. As one of the world's leading tourist destinations, China has welcomed a large number of international visitors after the pandemic, and the slogan 'China Travel' has attracted much attention and become an international buzzword (Chinese Tourism Academy., 2025). Notably, China adopted a 240-hour visa-free transit policy for visitors of 54 nations and completed reciprocal visa exemption agreements with countries such as Malaysia, Singapore, and the UK (National Immigration Administration., 2024), which significantly boosts inbound tourism. According to data released by the State Council of China, the country welcomed 131.9 million inbound visitors in 2024, representing a 61% year-on-year increase. Cities like Xi'an, known for its historical and cultural attractions such as the Terracotta Warriors and the ancient city wall, have particularly benefited from the policy. These shifts have further highlighted the significant role of English proficiency in delivering high-quality service in the tourism industry.

Tourism personnel, especially English tour guides, are important links between people from different cultures. Their proficiency in English directly impacts the delivery of high-quality tourism services and the overall satisfaction of international visitors (Ren & Wong, 2021). With the development of tourism, tourism personnel who welcome foreign guests are required to communicate with tourists in English in various situations, such as clarifying cultural distinctions, providing instructions and information, and handling emergencies (Munusamy & Kaliappen, 2023).

However, studies have shown that the English proficiency of many tourism personnel in China remains insufficient to meet the growing needs of diverse and changing work conditions (Hong et al., 2022). Tourism personnel, in fact, have difficulty dealing with tourists from foreign countries and encounter unexpected situations. The reason is that the existing English courses currently offered in Tourism Management Programs do not meet the actual needs of the industry (Bo, 2023). Many courses have failed to adjust to the dynamics of the post-pandemic world of international tourism and do not adequately meet the communication needs of actual situations of tourism (Fan, 2022). This disconnection has led to deficiencies in the practical application of English among tourism personnel.

To address this gap, English for Specific Purposes (ESP) courses are designed to meet the communication needs of tourism professionals. In common sense, the first step to developing an effective ESP course is conducting a systematic needs analysis (Long, 2005). This analysis identifies the specific language needs of target learners and helps design instructional content that is relevant to their context (Alshabeb et al., 2017; Richards, 2001). Over the past few decades, several theoretical models for needs analysis have been proposed. The models designed by Hutchinson and Waters (1987) and Dudley-Evans (1998) are the most widely used. Building on these previous foundations and considering the Chinese educational system, Chen (2010) proposed a needs analysis model that includes both personal learning needs and broader societal expectations.

Given that most tourism personnel in Xi'an have received local vocational English education and accumulated practical work experience, this study adopts Chen's (2010) social needs model to systematically investigate their English language needs in real-world contexts. The objective of

this study is to explore the English language needs of tourism personnel in Xi'an by examining their workplace use of English, their self-perceived language proficiency, and their recommendations for future ESP course design, in order to inform more context-relevant and needs-based Tourism English instruction. Accordingly, the study addresses the following research questions: (1) What English language skills are required by tourism personnel in their daily workplace? (2) How do tourism personnel evaluate their English language proficiency in terms of workplace demands? (3) What are their recommendations for the future Tourism English course?

Literature Review

Needs Analysis and Needs Analysis Models

As defined by Nunan and Carter (2001), "Needs analysis is fundamental to an EAP approach to course design and teaching" (p. 127), and it "should be the starting point for devising syllabuses, courses, materials and the kind of teaching and learning that takes place" (Jordan, 1997, p. 22). It refers to techniques and methods, such as introspection, interviews, observation, and questionnaires, which are used to identify learners' actual learning needs (Hyland, 2014). As Long (2005) proposed, "Just as no medical intervention would be prescribed before a thorough diagnosis of what ails the patient, so no language teaching program should be designed without a thorough needs analysis" (p. 1). It can be concluded that needs analysis is an essential step in developing a language course. As mentioned above, to achieve more effective needs analysis, numerous scholars have continually refined analytical models, resulting in a relatively mature and well-developed system (Chen, 2010). From the perspective of the entire development of foreign language needs analysis models, there are mainly four types, each representing a different stage of development: such as Munby's Target Situation Analysis model (1981); Allwright's Present Situation Analysis model (1982); Hutchinson and Waters' (1987) Needs Analysis model and Dudley-Evans & St. John's (1998) Needs Analysis model. In the Chinese context, Chen (2010) adapted these well-designed models and proposed a localized needs analysis framework that classifies learners' needs into two sub-models: a personal needs analysis model and a social needs analysis model, which is based on Shu's (2004) classification. This study adopts Chen's social needs model, which includes three key parts to examine the English needs of tourism personnel in Xi'an: employees' English use at the workplace, employees' self-evaluation of English proficiency at the workplace, and their recommendations for future ESP courses.

Needs Analysis Studies in Tourism English

Unlike previous studies that examine tourism professionals as a general group, this study focuses on two specific roles—tour guides and office staff who regularly interact with international visitors. This targeted approach enables a more detailed analysis of English language needs in frontline tourism services that are tailored to their specific job roles.

Numerous studies across Southeast and West Asia have emphasized the importance of oral communication skills in tourism contexts. For example, Prachanant (2012) found that speaking and listening were the most important skills for employees of international travel agencies in Thailand. Chumphong and Chuai-in (2020) also found that frontline staff in hotels and tourism agencies put more value on listening and speaking than on reading and writing. Namtapi (2022) further identified that it was challenging to understand different accents and keep up with fast-paced speech, as well as that vocabulary and grammar were limited.

Similar trends have been observed in other places. In Malaysia, Kholidi et al. (2022) reported that tourism staff found it difficult to listen and speak, while they thought written skills were less frequently required in work. In Indonesia, Farida (2022) emphasized the value of speaking and writing among certified tour guides and agency managers. In Saudi Arabia, Al-Malki et al. (2022) found that oral skills were necessary for giving information about tourist spots and answering tourists' questions, while grammar and vocabulary were not that essential. Mohammed (2024) reaffirmed this finding by pointing out that key speaking skills, such as giving directions, apologizing, and offering advice, are viewed as fundamental in guiding work. In the study by Aysu and Özcan (2021), which investigated Turkish tourism graduates and employers, the importance of oral communication clarity and efficiency was also highlighted.

Although these studies provide valuable insights, most focus on tourism contexts outside mainland China. There is a clear knowledge gap regarding the practical language requirements of tourism personnel in China, as most research has focused on English instruction for tourism students rather than working professionals (Sun, 2019; Wu, 2020). Given the growth of inbound tourism and Xi'an's ongoing popularity as a heritage destination, it is both appropriate and necessary to look into the specific English communication needs of Chinese tourism personnel who interact directly with tourists.

At the same time, few studies have applied Chen's (2010) social needs analysis model to this context. To address this gap, the present study applies Chen's social needs model to conduct a detailed analysis of the English language needs of tourism personnel in Xi'an. The aim is to provide both theoretical perspectives and practical suggestions that can help improve ESP courses for future tourism management students and tourism personnel training in China's vocational colleges, ensuring that these programs are more closely aligned with industry requirements.

Method

Research Design

This study adopted a mixed-methods approach to investigate the English language needs of tourism personnel in Xi'an, a major tourist destination in China. A combination of quantitative and qualitative methods was used to get a wide range of patterns and in-depth insights. The quantitative part used a structured questionnaire, and the qualitative part consisted of semi-structured interviews. This design can offer a comprehensive understanding of the language challenges and work demands faced by tourism personnel.

Instruments

The questionnaire was developed based on Chen's (2010) social needs model and followed China's national competency standards for English-speaking tour guides. It has three parts: the participants' demographic background, their perceived importance of English language skills in the workplace, and their self-assessed proficiency in those skills. A pilot study was conducted to verify that the instrument was clear, reliable, and valid in content. For the qualitative phase, a semi-structured interview protocol was used. Twenty open-ended questions explore participants' workplace English use, perceived strengths and challenges, and suggestions for future course improvement. Interviews were conducted in Mandarin using WeChat or Tencent Meeting and lasted between 20 and 30 minutes.

Participants

A total of 104 tourism personnel participated in the quantitative part. This included 79 frontline English-speaking tour guides and 25 office-based staff who interacted with international visitors. All participants held valid tour guide licenses and had at least one year of working experience with foreign tourists. The demographic information of participants is listed in Table 1 as demonstrated below:

Table 1. *Demographic Information*

Demographic Variables	Categories	Frequency	Percentage
Workplace	Private company	79	76.0%
	State-owned company	25	24.0%
Working year	Less than 3 years	17	16.4%
	More than 3 years	87	83.7%
Job position	Office staff	25	24.0%
	Tour guide	79	76.0%

In the qualitative phase, five participants were purposively selected. This group included two office staff, two scenic site guides, and one local guide. To ensure confidentiality, the names of T1 to T5 were used during the data analysis and reporting process.

Data Analysis

Data analysis proceeded in two stages. For quantitative data, SPSS v28 was used. Prior to analysis, all data were screened and cleaned, resulting in 67 valid responses. Descriptive statistics, including means, standard deviations, and percentages, were used to analyze participants' perceptions of the importance of English skills and their self-evaluated proficiency. For qualitative data, thematic analysis was conducted, and it followed Braun and Clarke's (2006) six-phase procedure. Interview transcripts were translated into English, coded by the first author using NVivo 15, and reviewed by co-authors and domain experts. Discrepancies in coding were discussed until consensus was reached to ensure trustworthiness and consistency. The research procedure is demonstrated in Figure 1 below.

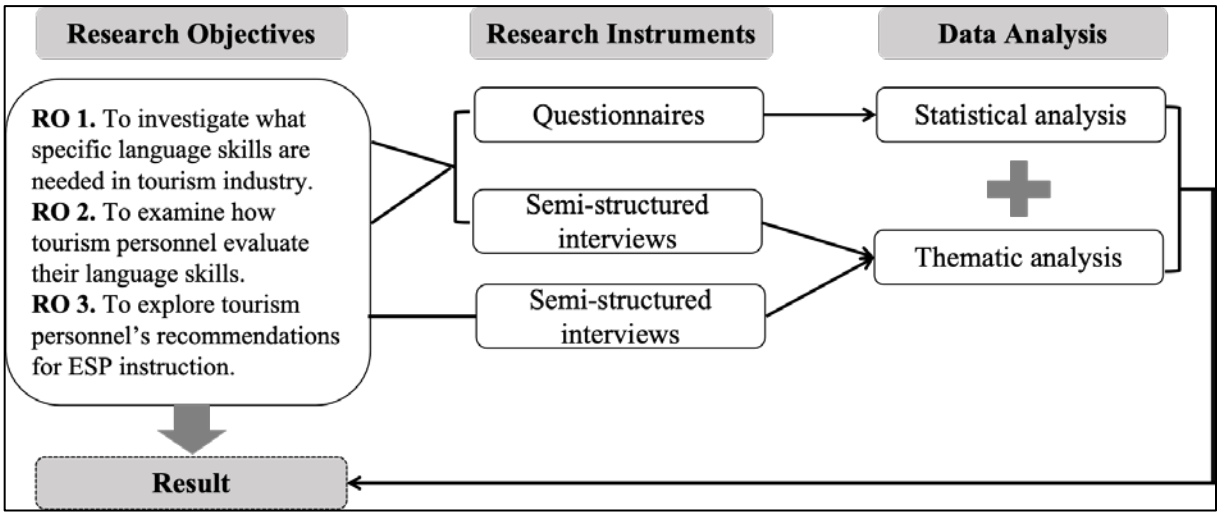


Figure 1. Research Design

ISSN: 2229-9327

Ethic Consideration

Ethical approval for this study was obtained from the Research Ethics Committee of the university in which the present study was registered. All participants provided informed consent and were assured of confidentiality and anonymity. Participation was voluntary, and participants could quit at any time without consequence.

Results

This section reports the main findings of the study, focusing on three aspects: the specific English language skills required in tourism work, the self-evaluated proficiency of tourism personnel, and their recommendations for future ESP instruction.

Specific Language Skills Needed in The Tourism Industry

To start with, the first research objective explores participants' perceptions of the importance of listening, speaking, reading, and writing skills in their daily professional work. The questionnaire results show that listening and speaking were rated as the most important English skills in daily tourism work (M = 3.24), followed by reading (M = 3.21), with writing receiving the lowest rating (M = 2.17) (see Table 2).

Table 2. *Mean score of four English skills*

Skills	*M*
Listening	3.24
Speaking	3.24
Reading	3.21
Writing	2.17

M mean

Subskill-level data are presented in Table 3. For listening, the most important tasks identified were understanding announcements and promotional videos, identifying keywords, and recognizing emotions or attitudes. In the interviews, participants reported that capturing keywords and the emotions of tourists are essential for resolving complaints and managing on-site situations. For example:

> *'Sometimes tourists are emotional and speak quickly. If I can pick up on keywords like 'late' or 'wrong place,' I can quickly understand what has happened.'* (T1)
>
> *'I have to focus on the main words; otherwise, a misunderstanding could affect the arrangements. We don't need to catch every detail, but we must grasp the key emotions and meanings.'*(T2)

For speaking skills, participants place high importance on frequently used expressions and spontaneous communication, such as self-introduction, giving unscripted explanations, and responding to tourists' questions. Interviews reconfirm the practical need for adaptability and flexible expression, which shows that communication skills have a direct impact on service quality and work efficiency. For instance, T5 remarked, *Self-introduction is the most basic skill, and it sets the first impression every time I meet a new group of foreign tourists.* T3 shared, *'When we explain attractions in English, we normally have prepared scripts, but more often we have to adjust based on tourists' reactions, so fast reactions are really important.'* T4 added, *'Sometimes tourists ask unexpected questions during explanations, and if I can't answer clearly in English, it affects their impression of the service.'*

Regarding reading, the participants placed high emphasis on handling work-related practical texts, such as attraction descriptions, itineraries, and tickets, as well as emails and instant messages.

Interviews further corroborate the quantitative data signified in Table 3 below. T1 explained, *'It's crucial to accurately understand the key information in English itineraries and tickets—like time and place—otherwise, mistakes can easily happen.'* T5 noted, *'The most important content is directly related to the foreign group's itineraries. If you miss any detail, it can disrupt the whole trip.'* T3 highlighted the use of English signs inside attractions: *'Even though there are Chinese versions, I often check the English ones in advance to help explain things to foreign tourists.'* Regarding the importance of understanding emails and instant messages, all interviewees mentioned it during the interviews.

As summarized in Table 3, writing received a relatively lower overall rating. However, practical tasks, such as responding to tourists' inquiries, preparing itineraries, and filling out forms all in English, are recognized as core skills in improving the participants' efficiency and professionalism. As T1 shared, *'We often have to reply to tourists' questions about schedules, timings, or costs by email.'* T3 mentioned, *'Besides giving explanations, I also frequently answer questions via WeChat or email.'* In addition, T5, who works as a tour guide, emphasized the value of making itineraries and giving important notices to the guests.

Table 3. *Important Skills at Work*

Skills	Subskills	*M*	*SD*
Listening	Understand English announcements and videos	3.30	0.78
	Identify keywords	3.28	0.71
	Recognize attitudes and emotions	3.16	0.79
Speaking	Give self-instructions	3.36	0.60
	Give unscripted explanations	3.30	0.76
	Respond to tourists' questions	3.25	0.73
Reading	Understand attraction descriptions	3.31	0.78
	Read itineraries and tickets	3.30	0.84
	Read email and instant messages	3.21	0.77
Writing	Write responses to tourists' inquiries in English	3.30	0.80
	Prepare travel itineraries	3.18	0.76
	Fill in English forms	3.15	0.70

M mean *SD* Standard Deviation

Tourism Personnel's English Language Proficiency at Work

The second research objective focuses on the participants' self-assessment of their English language proficiency in daily work. Based on the data from questionnaires and interviews, the findings reveal that participants are generally confident when handling structured and routine tasks; however, they face challenges in more complex and unpredictable communication situations. The quantitative data are demonstrated in Table 4.

Table 4. *Performance of skills at work*

Skill	Best-performing subskills	*M*	*SD*	Lowest-performing subskills	*M*	*SD*
Listening	Identify keywords	2.79	0.99	Identify different accents	2.51	0.99
	Understand announcements /videos	2.76	0.96	Recognize attitudes and emotions	2.69	0.99
Speaking	Give self-introductions	2.96	1.01	Give unscripted explanations	2.58	0.99
	Give scripted explanations	2.78	0.95	Participate in discussions	2.64	0.83
Reading	Read itineraries and tickets	2.85	1.03	Read tourism-related articles	2.57	0.82
	Read forms and invoices	2.82	0.91			

Writing	Write responses	2.79	0.91	Design travel brochures	2.55	0.99
	Fill out forms	2.79	0.98			

In terms of listening, participants reported the highest mean score for identifying keywords. Nevertheless, they struggled with different English accents, and interview data further support this trend. It showed that participants frequently mentioned unfamiliar accents are a primary barrier to listening comprehension. T1 stated, 'When the accent is quite strong, I sometimes get lost.' Similarly, T5 explained, 'If a tourist has a heavy accent, I really have to listen several times to understand.' In fast-paced conversations, several participants reported relying on keywords to guess the whole meaning.

For speaking, participants showed more confidence in prepared content and through regular practice, such as self-introductions and giving scripted explanations in descriptive analysis. However, their confidence drops in unscripted, spontaneous expression—for example, giving unscripted explanations or participating in discussions. T3 expressed that, although he is proficient in giving prepared explanations of tourism spots, he often encounters difficulties during unexpected Q&A sessions due to his limited vocabulary and grammar. T4 added, *'If tourists simply ask for directions or ticket information, I can respond quickly. But when it comes to complex topics like the history of a site or itinerary suggestions, I usually need to organize my thoughts first, and look for the proper words.'*

In reading, participants rated themselves as more capable when handling structured texts such as itineraries, tickets, forms, and invoices. In contrast, reading tourism-related articles received the lowest score in this skill area. Interviewees similarly emphasized their familiarity with simple documents. As T1 mentioned, *'Sometimes, I receive formal notices from travel agencies in English. I just don't understand the professional terms.'* T2 echoed this, stating, *'Legal documents or official tourism articles can be challenging, mostly because of the professional vocabulary.'*

In writing, the tasks that were performed with the most confidence were writing responses to tourists and filling out forms. Designing promotional materials, such as brochures, received the lowest mean score. Interviewees confirmed this pattern. For example, T5 reported, *'Replying to daily emails and completing English forms is basically no problem—even if my language isn't perfect, the communication is still clear.'* T2 added that most complaints or inquiries can be handled smoothly thanks to the established templates that the company offers. However, when it comes to less familiar writing tasks, such as preparing the brochures, as T1 stated, *'I haven't had systematic training in promotional writing, and my vocabulary isn't idiomatic. Usually, the marketing or external communications staff handle those tasks.'*

Tourism Personnel's Recommendations for ESP Instruction

Interview participants provided several suggestions regarding the content and delivery of future ESP courses for tourism professionals.

First, all participants emphasized the importance of prioritizing listening and speaking skills. They highlighted challenges in responding spontaneously to tourists' questions and their desire to improve both fluency and pronunciation. As T2 explained, *'Understanding tourists' questions is just the starting point; you also need to be able to answer right away, even if it's not a complicated response, as long as you can say it clearly and quickly.'* T3 added, *'Sometimes I know how to say something in my mind, but when I speak, I hesitate, especially when tourists ask unusual questions or express themselves in unexpected ways. I want to learn how to respond quickly in English.*

ISSN: 2229-9327

Pronunciation is also an issue—I often feel that my pronunciation isn't accurate enough and hope to improve it.'

Secondly, participants recommended that course content be closely linked to real workplace situations. Many suggested that typical tasks, such as simulated tourism attraction explanations, complaint handling, and route guidance, can help improve practical language skills, and they can be reinforced through scenario-based training, like role-plays. In terms of course design, they proposed including specific exercises in site interpretation and practical response techniques, so that students could master commonly used expressions and apply them immediately on the job after completing the course. Selected interview excerpts include:

> *'The course could include more practical listening and speaking exercises, like simulating the tourists asking questions and we can respond.'* (T1)
>
> *'It would be best to have practical simulations, such as airport reception, site explanations, or handling complaints handling—specific situations where we need to use English.'* (T5)

Additionally, participants prefer a more interactive classroom and more engaging learning experiences. In this way, they can have opportunities for sharing practical insights. They viewed group work and role-plays as more engaging and effective than traditional exercises. T2 mentioned, *'It would be very useful if some experienced guides could share how they deal with real issues. The course should be more interactive—not just filling in blanks or translating sentences. After starting work, you realize that on-the-spot problem-solving is most important.'* T4 agreed, *'Role-play activities during practical training are very interesting. I think these kinds of exercises really helpful.'*

Finally, several participants emphasize the importance of teacher feedback and encouragement and which can enhance effective communication. T3 said, *'Teachers could encourage us to speak more, for example, through group exercises and role-plays.'* T5 added, *'Sometimes we don't realize when our English is unnatural. If teachers can correct us, we can improve next time.'*

Discussion

The findings of this study reaffirm the widely recognized importance of listening and speaking skills in tourism-related communication. This trend is consistent with previous research (Aysu & Özcan, 2021; Farida, 2022; Ijabah & Amrullah, 2023; Prachanant, 2012; Sermsook et al., 2021), which emphasizes the centrality of oral interaction in tasks such as handling inquiries, providing explanations, and managing complaints. In the present study, however, the significance of oral skills extends beyond general fluency to include more specific subskills. Participants placed particular emphasis on identifying key information and interpreting emotional cues during conversations—subskills that are often overlooked in previous studies, which have focused primarily on broad communicative competence. This suggests a shift in emphasis from generic speaking proficiency to situational responsiveness in real-time interaction.

Although participants rated reading and writing as relatively less important than oral skills, their role in everyday tourism tasks remains clear. Practical literacy activities—such as understanding itineraries, processing instant messages, and preparing short written documents—were frequently cited. Challenges arose primarily about longer, professional, or creative texts, for example, formal reports or promotional brochures. These observations align with Chumphong and Chuai-in (2020), Namtapi (2022), and Wahyuningtyas et al. (2022), who found that while tourism workers are capable of navigating functional texts, they often lack confidence in handling complex

written genres. The distinction between transactional literacy and extended writing ability highlights the need for differentiated support across reading and writing tasks in course design.

Another notable pattern concerns the participants' greater confidence in structured, predictable tasks, and they lack confidence in spontaneous and unfamiliar communication. Participants have confidence in high-frequency tasks, such as self-introductions, email replies, and scripted explanations. In contrast, they found it more difficult to respond to questions, adapt to varying accents, or engage in open-ended discussion. These findings are consistent with Ka-kan-dee and Nonthapot (2020), Oktavinanda et al. (2022), and Giri and Boonrattanakittibhumi (2022), who also noted challenges in flexible oral communication among tourism professionals. The present study also found that the "template-based approach" is mostly applied in language teaching, as language is built on repeated, predefined patterns. This approach is actually rooted in both education and workplace practice. It may reflect a broader trend in vocational English education in China. As suggested by Bo (2023), this approach can prioritize accuracy and structural control over spontaneity. However, while templates facilitate clarity and reduce error in routine contexts, they may also hinder adaptability in dynamic interactions.

Beyond skill assessment, this study also offers insight into participants' expectations for course design. Interviewees expressed a strong preference for training that focuses on real-world communication tasks, especially those involving quick responses and accurate pronunciation. These perspectives reinforce the need to move beyond general teaching practice to a more scenario-based training and on-the-spot responsiveness, as also emphasized by Farida (2022) and Namtapi (2022). In terms of instructional methods, participants favored scenario-based activities such as role-plays and simulations. These preferences resonate with pedagogical models proposed by Aysu and Özcan (2021) and Al-Malki (2023), which emphasize contextually grounded, interaction-rich learning.

Additionally, participants emphasized the importance of teacher feedback and guided correction, which supports a more dynamic and responsive classroom environment. This aligns with the learner-centered approaches advocated by Lertchalermtipakoon et al. (2021). They recommend a reduced reliance on the grammar-translation method and an increased use of interactive strategies.

Pedagogical Implication

Based on the findings and existing literature, several implications for designing tourism English courses are proposed and stated below.

First, ongoing and context-specific needs analysis is essential for effective ESP course development (Graves & Xu, 2000; Hyland, 2014). As the tourism context evolves, particularly with recent policy changes such as China's expanded visa-free entry and the rebound of inbound travel, language demands on frontline staff continue to shift. Regular updates based on dynamic needs assessments can ensure that course content remains aligned with actual workplace communication tasks.

Second, course content should be relevant and closely aligned with the actual language needs of the learners. The present findings from tourism-related professionals in Xi'an indicate a clear emphasis on oral communication skills, and they show more focus on the ability to use the language more flexibly. In terms of teaching methods, practical activities such as role-plays and real-scenario simulations are recommended to help students gain more experience in language use. Additionally, the use of digital tools and multimedia resources can further enhance engagement

and practicality, particularly for those digital-native learners. Finally, it indicates that instructors play a pivotal role in the success of the ESP course. Effective instructions should motivate students to communicate confidently in realistic settings and offer timely, constructive feedback. Based on interview insights, inviting experienced tour guides to share workplace language experiences and strategies can boost course relevance and foster students' professional identity.

To summarize, these strategies may promote meaningful learning and better prepare students to meet the high communication standards of the post-pandemic tourism industry.

Conclusions and Limitations

This study examined the English-language needs and proficiency of tourism personnel in the Xi'an area through a mixed-methods approach. Results show that oral communication skills are considered essential by practitioners, particularly in understanding the keywords of conversation, interpreting emotions, and responding spontaneously. While literacy received less emphasis, the participants still valued the ability to handle workplace texts, for example, itineraries and inquiries. Overall, practitioners expressed confidence in routine, template-based communication but were less prepared for dynamic, unpredictable ones. Furthermore, the participants strongly recommended that tourism English courses prioritize listening and speaking and be grounded in real-world situations, such as free communication. They showed a preference for role-plays, simulations, and real-time feedback in class to gain practical communication skills. These insights highlight that course design should be closely aligned with industry demands in order to improve the relevance and impact of ESP training.

This study has several limitations. First, the sample is geographically confined to tourism personnel in Xi'an. Future studies could expand to include other regions to explore regional variations in English language needs. Second, the research focuses solely on practitioners' perspectives, and lacks input from tourists in terms of service quality. Finally, while this study focused on general language competencies, it did not examine micro-level skills, such as accuracy in specialized vocabulary (e.g., historical, culinary, or transportation terms) or grammatical precision. Future research could address these dimensions to inform more targeted and professionally aligned ESP curriculum design.

Funding: This research is not funded.
Acknowledgements: Not applicable.
Conflicts of Interest: The authors declare no conflict of interest.
Authenticity: This manuscript is an original work from field research.
Artificial Intelligence Statement: During the preparation of this work, the authors used Grammarly and ChatGPT to check the grammar and improve language expressions. After using the tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

About the Authors:

Boyu Pang is a PhD candidate at the Faculty of Education, Universiti Teknologi MARA, Malaysia. She holds a MA. from the University of Edinburgh and currently teaches in a vocational university in China. Her research interests include English as a Foreign Language (EFL), English for Specific Purposes (ESP), and English language instruction in higher education. https://orcid.org/0009-0005-7175-6782

Prof. Dr. Faizah Binti Abd Majid is a Professor in Adult and Higher Education at the Faculty of Education, Universiti Teknologi MARA, Malaysia. She has vast experiences in research and publication in the areas of TESL, teaching and learning.
https://orcid.org/0000-0002-9642-6342

Dr. Geethanjali A/P Narayanan is a senior lecturer at the Faculty of Education, Universiti Teknologi MARA, Malaysia. Her research interests are ESP teaching and teaching English as a foreign language. https://orcid.org/0000-0003-4118-5622

References:

Al-Malki, E. A. (2023). English Language and Communicative Proficiency of Saudi Tourism and Hospitality Students: A Present Situation Analysis. *Arab World English Journal (AWEJ)*, *14*, 458-475. https://doi.org/http://dx.doi.org/10.2139/ssrn.4432878

Al-Malki, E. A., Javid, C. Z., Farooq, M. U., Algethami, G. F., & Al-Harthi, A. A. (2022). Analysis of the English Language Needs of the Saudi Tourism Workforce: A First Step into Designing ESP Teaching Materials. *International Journal of Learning, Teaching and Educational Research*, *21*(2), 72-88. https://doi.org/https://doi.org/10.26803/ijlter.21.2.5

Allwright, R. (1982). Perceiving and Pursuing Learners' Needs. In M. Geddes & G. Sturtridge (Eds.), *Individualisation* (pp. 24-31). Oxford: Modern English Publications.

Alshabeb, A., Alsubaie, F. H., & Albasheer, A. Z. (2017). English for Specific Purposes: A Study Investigating the Mismatch between the 'Cutting Edge' Course Book and the Needs of Prince Sultan Air Base Students. *Arab World English Journal (AWEJ)*, *8*, 208-219. https://doi.org/https://dx.doi.org/10.24093/awej/vol8no3.24

Aysu, S., & Özcan, F. H. (2021). Needs Analysis in Curriculum Design: Language Needs of Tourism Students. *Sakarya University Journal of Education*, *11*(2), 305-326. https://doi.org/https://doi.org/10.19126/suje.854993

Bo, D. (2023). The Application of Schema Theory in Listening and Speaking Teaching of Tourism English in Higher Vocational Colleges. *Curriculum and Teaching Methodology*, *6*(7), 88-93. https://doi.org/10.23977/curtm.2023.060715

Braun, V., & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative research in psychology*, *3*(2), 77-101. https://doi.org/https://doi.org/10.1191/1478088706qp063oa

Chen, B. (2010). Theoretical Construction of a College English Needs Analysis Model. *Journal of Foreign Languages (Shanghai International Studies University)*(2), 120-123.

Chinese Tourism Academy. (2025). *Report on the Development of China's Inbound Tourism 2024 - 2025 Released —— Demand for Inbound Tourism to China Continues to Grow*. The State Council of the People's Republic of China. Retrieved 28 March from https://www.gov.cn/yaowen/liebiao/202503/content_7016031.htm

Chumphong, O., & Chuai-in, P. (2020). The Use of English Language in Tourism Workplaces in Pakmeng Beach Area, Trang Province. *ABAC ODI Journal Vision. Action. Outcome*, *7*(1), 132-144. https://core.ac.uk/download/pdf/287162844.pdf

Dudley-Evans, T., & St John, M. J. (1998). *Developments in English for Specific Purposes*. Cambridge University Press.

Fan, W. (2022). Study on the Characteristics, Problems and the Counter Measures to Teaching Contents in Colleges English Course of 'Tour Guiding' under a New Tourism Environment.

International Journal of Humanities, Social Sciences and Education, *9*, 94-99. https://doi.org/10.20431/2349-0381.0911011

Farida, M. (2022). Developing Local-Based English for Tourism Material for Religious and Cultural Tour Guide Program Students: A Needs Analysis. *ENLIT Journal*, *2*(1), 74-82. https://www.academia.edu/download/103996903/935.pdf.

Giri, A., & Boonrattanakittibhumi, C. (2022). A Needs Analysis of English Communication Skills for Local Youth Guides for Promoting Tourism in Pathio District, Thailand. *The International Journal of Interdisciplinary Educational Studies*, *17*(1), 85-103. https://doi.org/https://doi.org/10.18848/2327-011X/CGP/v17i01/85-103

Graves, K., & Xu, S. (2000). *Designing language courses: A guide for teachers*. Heinle & Heinle Boston, MA.

Hong, Y., Chen, L.-G., Huang, J.-H., Tsai, Y.-Y., & Chang, T.-Y. (2022). The Impact of Cooperative Learning Method on the Oral Proficiency of Learners of the Training Program for English Tourist Guides. *Frontiers in Psychology*, *13*, 1-13. https://doi.org/https://doi.org/10.3389/fpsyg.2022.866863

Hutchinson, T., & Waters, A. (1987). *English for Specific Purposes*. Cambridge University Press.

Hyland, K. (2014). English for Academic Purposes. In C. Leung (Ed.), *The Routledge Companion to English Studies* (pp. 392-404). Routledge.

Ijabah, N., & Amrullah, Q. L. (2023). Need Analysis of English Tourism: The Evidence from The Employees of Tourism Industry. *Acitya: Journal of Teaching and Education*, *5*(2), 214-226. https://doi.org/10.30650/ajte.v5i2.3590

Jordan, R. R. (1997). *English for Academic Purposes*. Cambridge University Press.

Ka-kan-dee, M., & Nonthapot, S. (2020). Language communication barriers among tourism services personnel in the Upper Northeast, Thailand. *International Journal of Management*, *11*(9), 613-622. https://doi.org/https://ssrn.com/abstract=3710469

Kholidi, M. A., Azhan, N. A. S., & Ariawan, S. (2022). Needs analysis of English language use in tourism industry in Selangor, Malaysia. *LSP International Journal*, *9*(1), 123-136. https://doi.org/10.11113/lspi.v9.18285

Lertchalermtipakoon, P., Wongsubun, U., & Kawinkoonlasate, P. (2021). Need Analysis: English Language Use by Students in the Tourism and Hospitality and Industry. *English Language Teaching*, *14*(3), 59-71. https://doi.org/https://doi.org/10.5539/elt.v14n3p59

Long, M. H. (2005). *Second Language Needs Analysis*. Cambridge University Press.

Mohammed, A. S. (2024). Designing a Needs-Based English Syllabus for Tour Guides in Saudi Arabia. *Journal of Language Teaching and Research*, *15*(3), 972-983. https://doi.org/10.17507/jltr.1503.31

Munby, J. (1981). *Communicative Syllabus Design: A Sociolinguistic Model for Designing the Content of Purpose-specific Language Programmes*. Cambridge University Press.

Munusamy, P., & Kaliappen, N. (2023). English for Tourism and Hospitality Purposes. In A. Sharma (Ed.), *International Handbook of Skill, Education, Learning, and Research Development in Tourism and Hospitality* (pp. 1-13). Springer.

Namtapi, I. (2022). Needs Analysis of English for Specific Purposes for Tourism Personnel in Ayutthaya. *LEARN Journal: Language Education and Acquisition Research Network*, *15*(1), 409-439. https://so04.tci-thaijo.org/index.php/LEARN/article/view/256730

National Immigration Administration. (2024). *China's Transit Visa-free Policy Has Been Comprehensively Relaxed and Optimized.* The State Council of the People's Republic of

China. https://www.nia.gov.cn/n897453/c1688948/content.html

Nunan, D., & Carter, R. (2001). *The Cambridge Guide to Teaching English to Speakers of Other Languages*. Cambridge University Press.

Oktavinanda, G., Rahma, E. A., Mursawal, A., & Syahputri, V. N. (2022). Voices from Coastal Communities: English Need Analysis of Local Tour Guide. *Lingua Didaktika: Jurnal Bahasa dan Pembelajaran Bahasa*, *16*(2), 155-168. https://doi.org/10.24036/ld.v16i2.119741

Prachanant, N. (2012). Needs Analysis on English Language Use in Tourism Industry. *Procedia-Social and Behavioral Sciences*, *66*, 117-125. https://doi.org/https://doi.org/10.1016/j.sbspro.2012.11.253

Ren, L., & Wong, C. U. I. (2021). Tour Guides' Dilemma under the Pandemic: To Stay or Not To Stay? An Investigation of Occupational Loyalty Among Tour Guides in Macao. *Asia Pacific Journal of Tourism Research*, *26*(11), 1270-1283. https://doi.org/https://doi.org/10.1080/10941665.2021.1980411

Richards, J. C. (2001). *Curriculum Development in Language Teaching*. Cambridge University Press.

Sermsook, K., Nakplad, R., & Jantawong, L. (2021). Problems and Needs Analysis of English for Tourism Industry: A Case of Hotel Staff in Phang Nga Province, Thailand. *Education Quarterly Reviews*, *4*(3), 73-82. https://doi.org/https://ssrn.com/abstract=3887154

Shu, D. (2004). *Foreign Language Teaching Reform: Problems and Countermeasures*. Shanghai Foreign Language Education Press.

Sun, T. (2019). *Research on the Optimization of Tourism English Curriculum Design in Secondary Vocational Schools Based on Needs Analysis* [Master Thesis, https://kns.cnki.net/kcms2/article/abstract?v=tMRSZR5ycIvv-lOEiW1nFg0n4p4vxQPZtMx8g8CYNunMuF7jQreIBEeV7MX6zECVXHUqi4WEYHS6W9TDkRvIvRHyJiMAymfiORfNV4YTa0-1KfUZ64sQKQ2L9dphp87yclur_lCUCCoNrQJpsMUyTiNTffrUi3r6Z6nWJrkcnRXhS3dF1khLrjCgqrY7kefSZDS5m0zzaBw=&uniplatform=NZKPT&language=CHS

Wahyuningtyas, D., Sumarno, W. K., Asih, R. A., Endriana, F., Windiarti, K. R., & Mafruudloh, N. (2022). An Analysis of the English Speaking and Writing Skill Level of Tour Guides in Trenggalek's tourist villages. *Journey: Journal of English Language and Pedagogy*, *5*(2), 293-303. https://doi.org/https://doi.org/10.33503/journey.v5i2.574

Wu, T. (2020). *Research on the Curriculum Design of Tourism English in Secondary Vocational Colleges Based on Needs Analysis* [Master Thesis, https://link.cnki.net/doi/10.27036/d.cnki.ggxsu.2020.000692

Zahedpisheh, N., Abu Bakar, Z. B., & Saffari, N. (2017). English for Tourism and Hospitality Purposes (ETP). *English Language Teaching*, *10*(9), 86-94.

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.329-344
DOI: https://dx.doi.org/10.24093/awej/vol16no3.20

Investigating the Difficulties Encountered by Saudi English Majors in Transcribing Schwa Sound \ə\

Ghaida Khalid Albahlal
Department of English, College of Science and Humanities, Kharj, Prince Sattam bin Abdulaziz University, Saudi Arabia
Email: ghaida.albhlal@gmail.com

Ansa Hameed
Associate Professor, Department of English, College of Science and Humanities, Kharj, Prince Sattam bin Abdulaziz University, Saudi Arabia
Corresponding Author: a.zahoor@psau.edu.sa

Received:06/29/2025 Accepted: 09/08/2025 Published:09/20/2025

Abstract

English language learning poses various challenges to Arab learners including the transcription of various unfamiliar sounds. This quantitative study examines the challenges faced by English majors in transcribing words that contain the schwa sound /ə/. The prime objective is to identify which positions of the schwa sound poses the greatest difficulty, analyze student performance patterns, and suggest possible solutions. The participants were randomly selected female English majors from Prince Sattam bin Abdulaziz University (PSAU) during the 2025 academic year. A transcription test was used to collect data that was administered to 68 participants. The SPSS software was used to analyze statistical data. The results showed that schwa sounds in initial positions were the hardest to transcribe, followed by medial and final positions. However, the differences in error rates across positions were relatively small. The findings indicate that the main issue Saudi English majors face in schwa transcription is their difficulty in accurately pronouncing the sound. Based on these results, the study recommends targeted strategies to improve learners' pronunciation and transcription of the schwa sound. Additionally, further research is encouraged to explore how the Arabic language influences Saudi EFL learners' ability to recognize and transcribe the schwa sound. The research is significant in the Saudi EFL context to understand the pronunciation issues and focus the weak areas in pedagogy.

Keywords: Phoneme, Phonetic transcription, Schwa sound, Schwa sound challenges, Saudi EFL learners

How to cite: Albahlal, G. K., & Hameed, A. (2025). Investigating the Difficulties Encountered by Saudi English Majors in Transcribing Schwa Sound \ə\. *Arab World English Journal, 16*(3): 329-344. https://dx.doi.org/10.24093/awej/vol16no3.20

Introduction

In many languages, certain sounds are pronounced differently depending on their positions within words or the surrounding phonetic context. As such, teaching the sounds of the English language alongside its alphabet is essential for learners to produce words with native-like pronunciation. Accurate pronunciation supports clear understanding and effective communication, whereas mispronunciation can lead to confusion and misunderstanding. Among the most challenging phonemes for learners to transcribe, pronounce, and distinguish are vowels, particularly the schwa sound (Elzaki, 2017). The schwa is the most frequently occurring vowel in English and is often described as an unstressed and reduced vowel. It lacks consistent letter representation, making it especially difficult to identify and articulate (Beghoul, 2017). Thus, this sound is a challenge for English language learners and especially for those who learn it as a foreign language.

Being in line with the needs of the globally connected world, English language learning is also an essential requirement for the academic and corporate institutions in Saudi Arabia. As it is a foreign language in this context, so Arab learners face many difficulties in learning. From listening to speaking and from reading to writing, researchers have identified several areas where Saudi EFL learners face serious issues (Mohanna, 2024). Shaibani (2023) also comments that teaching pronunciation is one of the major weak areas in the kingdom. English sounds and their transcriptions are also one of these serious problems. Alzinaidi and Abdel Latif (2019) identified the issues Saudi FL learners fac wile dealing with English consonant sounds and proposed that the EFL pedagogical practices need to focus the pronunciation with more practical approach. Along the lines, for Saudi learners of English as a foreign language, the schwa presents a unique challenge, as there is no exact equivalent sound in the Arabic language. As a result, learners often substitute it with other vowel sounds (Kondo, 1995).

In pronunciation instruction, transcription is considered one of the most essential skills. It enables students to recognize phonetic symbols, allowing them to read and pronounce words accurately without relying on audio input. However, learners often face difficulties with vowel sounds, as there is no one-to-one correspondence between letters and vowel phonemes, particularly in the case of the schwa (Ocampo, 2017). Through classroom observation, the researcher noted that some female Saudi English majors demonstrated noticeable challenges and weaknesses in transcribing words that contain the schwa sound. This research aims to examine the difficulties faced by Saudi English majors in transcribing the schwa sound /ə/, with a specific focus on identifying which positional context of the schwa is most challenging for participants to transcribe.

This study, therefore, aims

- To investigate the difficulties that Saudi English majors encounter in transcribing the schwa sound
- To examine the positions (initial, medial or final) of the sound where they have major difficulty
- To propose the possible solutions for overcoming these challenges

This study therefore seeks to answer the following question: Which positional context of the schwa sound (initial, medial, or final) presents the greatest difficulty in transcription for Saudi English majors at PSAU?

The significance of this study lies in identifying the specific challenges English majors face when transcribing words that contain the schwa sound, particularly by determining which positional context of the schwa is most challenging to recognize. By addressing these difficulties

and proposing practical solutions, the study aims to enhance learners' transcription and pronunciation skills. Additionally, the findings hold value for educators, as they provide insights into the specific areas where students struggle. This can help instructors refine their teaching methods for transcription and pronunciation, ultimately supporting more effective language instruction.

Literature Review

This section reviews the relevant literature on learners' difficulties in recognizing and transcribing the schwa sound. It is divided into two main parts. The first part defines and explains key terms, including phoneme, schwa sound, phonetic transcription, and difficulties related to the schwa sound. The second part discusses previous studies regarding students' challenges with schwa transcription and pronunciation.

The schwa sound

In every language, speech is composed of a variety of sounds, including vowels and consonants, which are organized into fundamental units known as phonemes. Phonemes serve as the basic building blocks of spoken language and differ from letters in that they are abstract units that can be represented in multiple ways (Roach, 2018). They are defined as distinct sound units that differentiate meaning between words (Delahunty & Garvey, 2010). In English, different letters or letter combinations may represent a single phoneme. For example, the long /iː/ vowel sound appears in words such as *each*, *achieve*, *me*, and *bleed*, despite being spelled differently in each case (Yavas, 2011). In other words, every phoneme is made up of combination of vowel and consonant sounds. These sounds are essential to learn while learning pronunciation. As the present study focuses one of the short vowels sounds that is schwa, the next paragraphs describe it in detail.

The schwa is a mid-central, unrounded, and lax vowel sound produced with the tongue in a flat and relaxed position. It commonly occurs in unstressed syllables and often replaces other vowel sounds in stressed syllables (Roach, 2018). The term "schwa" originates from the Hebrew word "Shewa," which means "no vowel" or "emptiness," reflecting its characteristic as a reduced or neutral vowel sound (Ariyani, 2011). Young and Messum (2022) define schwa "articulated as a minimal, reduced, vowel-type sound, and in other contexts it is not actively articulated and is instead no more than an incidental noise that occurs when the speaker passes from one consonant to another one" (p. 2). Phonetically, the schwa is transcribed as /ə/ and is the shortest and most frequently used vowel sound in English speech. While schwa typically appears in unstressed syllables, when it occurs in stressed syllables, it is transcribed as /ɜː/ (Philip, 2016). For example, unstressed syllables: community /kəˈmjuː·nə·ti/, about /əˈbaʊt/, teacher /ˈtiː·tʃər/, and stressed syllables: shirt /ʃɜːt/, hurt /hɜːt/, etc.

Wilkins's sound chart (1668), as cited in Heselwood (2013, p. 1050), depicts the schwa sound as /Y/rather than/ə/. Wilkins and Tucker regarded the schwa and /h/ as fundamental vowel sounds, describing them as the 'sonorous' and 'aspirate' roots from which "all of our other vocal sounds are made to spring" (ibid). The schwa can correspond to various vowel qualities, including /a/, /u/, /o/, /e/, and /i/. It can appear at the beginning, middle, or end of English words (Hudson, 2018).

Pronunciation and the challenge of phonetic transcription

Pronunciation is one of the critical skills to acquire while learning any language (Faraj, 2023). One of the most effective methods for teaching pronunciation is through transcription. Transcription involves writing the phonemic symbols of words between slashes, with each phoneme represented by a specific symbol. Square brackets are sometimes used to indicate variations in pronunciation between dialects, such as [ɛə] changing to [ɛː] or [ɒ] changing to [ɔ] (Lindsey, 2019). There are two main types of transcription: broad and narrow. Broad transcription provides a basic representation of sounds, for example, *people* /ˈpiːpəl/ and *goal* /gəʊl/. Narrow transcription, on the other hand, includes diacritic marks that add detailed articulatory information to the basic phoneme symbols, such as vowel nasalization in *ban* /bãn/ (Zsiga, 2014).

A crucial skill for improving transcription and pronunciation is understanding the articulators, analyzing phonemes, identifying places of articulation, and developing listening skills. These abilities work together to enhance students' capacity to distinguish between sounds, especially vowels, and support them in transcription tasks (Mompéan, 2017).

The specific challenge of the schwa

The schwa is one of the weak vowels and can occur at the beginning, middle, or end of words—for example, *banana* /bəˈnɑːnə/ and *camel* /ˈkæməl/. It is considered as one of the difficult sounds to acquire for both second and foreign language learners. In the foreign language context, it becomes more problematic if the first language of the learners does not contain any similar sound. Difficulties with the schwa often arise when the learner's native language does not include this sound, making it hard to recognize both in pronunciation and transcription. Incorrect pronunciation and interference from the native language commonly led to errors in transcribing the schwa (Ocampo, 2017). Kapranov (2021) has also listed some major factors that make the learning of schwa sound problematic, for example, the complicated orthographic system of English that lacks schwa sound, the phonological properties that are not easy to comprehend by the EFL learners.

The schwa often appears in combination with other vowels to form diphthongs, which are complex vowel sounds, making transcription more challenging. Diphthongs are gliding vowels characterized by a smooth transition between two vowel sounds, such as /eɪ/, /ɪə/, /eə/, /ʊə/, and /əʊ/ (Kurtineine, 2017). According to Vera Deittes (cited in Ocampo, 2017, pp. 38–39), learners are often influenced by the spelling of words when producing phonemes. Because schwa is not represented by specific letters, this can confuse. Various vowels (a, e, i, o, and u) can be pronounced as /ə/. Another source of transcription errors occurs when schwa replaces the /r/r/sound at the end of words, for example, in *farmer*, where the final letter "r" is pronounced as /ə/ (Tench, 2011).

Review of the empirical studies

Many studies have explored the difficulties that English majors face with phonetic transcription, particularly concerning vowel sounds. Among the most challenging phonemes to transcribe is the schwa sound. In English speech, the schwa is especially complex due to its high frequency and the fact that it can be represented by different letters with varying spellings (Beghoul, 2017). A review of some related studies has been presented below in a thematic order.

L1 interference

One of the major reasons that hinder the non-native speakers of English to be fluent is the interference from the mother tongue. Studies across different contexts (e.g., Chouchane, 2016 for Arabic; Ariyani, 2011 for Indonesian) consistently point to L1 interference as a primary cause of schwa transcription errors. In his study, Chouchane (2016) examined the pronunciation difficulties of English sounds among Arab learners, focusing on two Omani students as his sample. His analysis revealed that learners' mother tongue has a strong influence on their second language pronunciation. He noted that English has a larger vowel inventory than Arabic. Unlike Arabic, which lacks the schwa as an individual vowel, English uses the schwa as its most common vowel sound. This difference contributed to the students' difficulties with the schwa, mainly since it can be represented by various letters in English. In contrast, Arabic has a more consistent one-to-one correspondence between letters and sounds. Additionally, the students struggled to identify and produce the schwa correctly due to a lack of awareness of stressed and unstressed syllables, commonly referred to as the strong and weak forms in English speech.

Ariyani (2011) conducted a specialized study on Indonesian students learning English as a foreign language (EFL), focusing on their difficulties with schwa pronunciation. In her research, she administered a test to assess the learners' ability to identify and pronounce the schwa sound in English words. The results revealed that the students struggled to recognize and produce the schwa accurately. This difficulty was attributed to negative transfer from their native language, as they instinctively applied Indonesian pronunciation habits to English. Based on her findings, Ariyani recommended that students focus on learning phonetic symbols and transcription using the International Phonetic Alphabet (IPA) to improve their understanding and correct pronunciation of English sounds.

The role of stress and awareness

Few studies have focused the role of stress in pronunciation of schwa sound. As the students are not aware of stress patterns in English, they often make mistakes. In this regard, a study by Kapranov (2021) investigated whether the schwa sound poses difficulties for intermediate EFL learners whose native language is Norwegian. Participants were given phonemic transcription tasks, during which they frequently replaced the schwa with other vowel phonemes such as /e/, /ɒ/, /æ/, /ɑː/, and /ɜː/. The researcher emphasized the importance of teaching phonemic transcription to EFL learners, highlighting the schwa as a primary challenge for many students. According to Kapranov, this difficulty largely stems from learners' limited awareness of English word stress, which leads them to substitute the schwa with other vowel sounds

Cassimiro (2011) from Estadual de Goiás University conducted a study on the importance of phonetic transcription for achieving native-like pronunciation. The research highlighted the influence of the mother tongue on English speech, particularly regarding the neutral schwa sound as a significant challenge. For example, in the word *dangerous* /ˈdeɪndʒərəs/, the vowel "a" is in a stressed syllable, while the vowels "e" and "ou" occur in unstressed syllables and are neutralized to schwa sounds. Regardless of which vowel is neutralized, it should be transcribed as a schwa. However, the study found that English majors often do not follow this rule and tend to replace the schwa with other phonemes, such as /e/ and /ɪ/.

Another study by Weber (2018) highlights how listening tasks focused on schwa pronunciation help learners develop their phonological and morphological awareness. Since the schwa is the most frequently occurring sound in English and commonly appears in words with

three or more syllables, especially in academic texts, accurate pronunciation is essential. Weber emphasizes that mastering the schwa enables learners to recognize, pronounce, and transcribe effectively.

Pedagogical interventions

Finally, few studies have explored the role of pedagogy in improving the pronunciation. For example, Trzaskowski (2016) from Nicolaus Copernicus University conducted a study exploring the relationship between phonemic transcription and pronunciation skills among Polish English majors. The research involved comparing students' errors in transcription tasks with their average pronunciation mistakes, using a phonemic transcription test as the primary tool. The findings showed that most pronunciation errors were directly reflected in the students' inaccurate transcription performance. Based on these results, the researcher recommended that teaching pronunciation and transcription should be integrated from primary school onward to develop precise skills in both areas before advancing to higher levels.

Mahmood et al. (2015) conducted a comparative study in which the experimental group was exposed to dictation exercises using nonsense words over a three-month period to enhance their recognition of English phonemes. This group demonstrated better performance in phonetic transcription compared to the control group. However, despite this improvement, the schwa sound remained challenging for the learners due to its complex nature.

A recent study by Bamigbade and Raji (2022) explored errors in phonemic transcription among Nigerian learners of English as a second language (L2). The researchers gave transcription tests to the participants and discovered that 22% of the errors involved the schwa sound, mainly because learners struggled to recognize the schwa in words. Their study highlighted the importance of teaching transcription to English learners and suggested using digital apps to improve both transcription and pronunciation skills.

The above listed studies show how the problems associated with schwa sound among EFL learners have been focused within various contexts, however, there is no such research conducted within the Saudi context. The present study aims to fill the gap.

Research Methodology

This section outlines the methodology employed by the researcher to achieve the study's objectives. It covers the research approach, participant selection, the instruments used, pilot testing for validity and reliability, data collection procedures, and data analysis techniques.

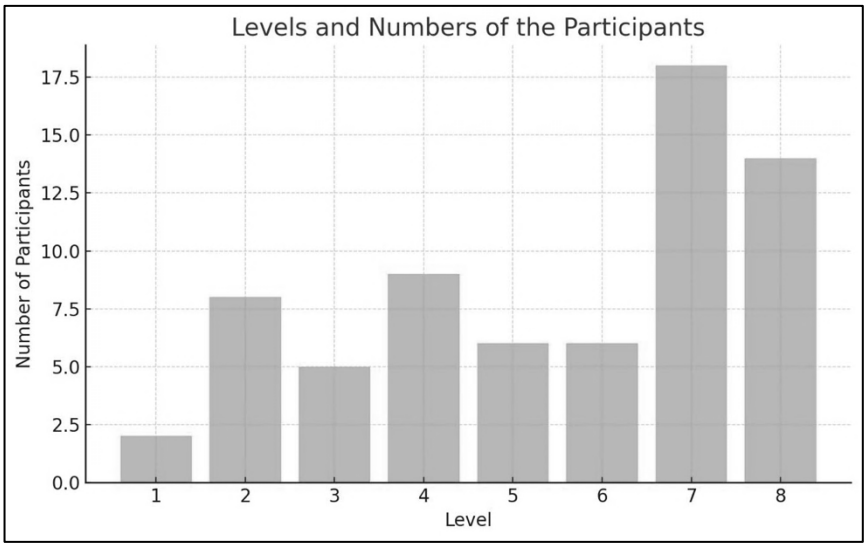
Method

The researcher employed a quantitative method for this study that primarily focuses on the numerical data. As the study aims to identify the issues of schwa sound through transcription, thus this method suits the purpose.

Participants of the Study

The participants in this study were 68 female English majors from different levels at PSAU during the 2025 academic year. These participants were selcted using random sampling metod. Random sampling metod is a probability-based sampling method which gives equal chances to all

the participants in the process. Figure one below shows that all the participants were from various



levels. The most prominent participants were from the seventh and eighth levels.

Figure 1: Semester levels of the selected participants

Research Instrument

The researcher employed a test as the primary instrument to determine which schwa sound location posed the most significant challenge for the participants to recognize. The test was administered online by the researcher, as it was difficult to assemble the students from various levels at one time and place. The test was prepared using Google forms and the time allocated for the test was 15 minutes, one minute for each item.

The researcher tested English majors, asking them to identify the vowel pronounced as a schwa sound in different word positions. The test included fifteen questions, evenly split among three schwa locations: front, middle, and final positions (test questions are provided in appendix-A). This allowed the researcher to determine which schwa location was the most difficult for Saudi English majors to recognize.

Piloting trials are essential for refining research procedures, obtaining necessary approvals, and assessing the feasibility of the study. Additionally, they help researchers gather more reliable data and improve the overall quality of the research. In this study, the pilot phase involved assessing the test's reliability through inter-rater agreement from an expert in the field.

Research Procedures

The test was distributed to 68 participants, and the collected data were analyzed using SPSS. This research employed statistical analysis to examine and interpret the numerical data collected from the research tool. It presents the participants' test performance results and offers a detailed analysis of their outcomes.

Findings

This section presents the results obtained from the research instrument. It includes the data collected from the students along with their analysis, highlighting which schwa position is the most challenging for English majors to identify.

The results were collected from female English majors across different academic levels. They are categorized according to the schwa sound's position in words: front, middle, and final. The findings are presented in the form of figures.

Figure 2 below presents the numbers and percentages of the correct and incorrect answers. The test included five questions to choose the vowel that is pronounced as schwa in a front location in words. The mean of the correct answers is 48.59% while the mean of the incorrect answers is 51.47%. The results show that the students' performance in identifying schwa in the front position is very low, as only half of the students answered the front position questions correctly. (For more details, see Table 1 in the Appendix-B)

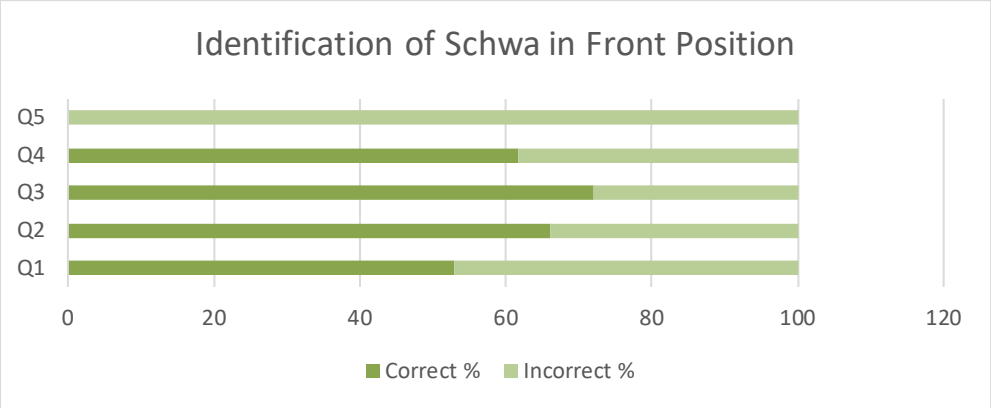


Figure 2. Identification percentage for front position of schwa sound

Figure 3 below shows the numbers and percentages of the correct and incorrect answers for each question. The test contained five questions to select the vowel that is pronounced as schwa in mid-position in words. The mean of the correct answers is 60% while the mean of the incorrect answers 40%. The mean represents that more than half of the students answered correctly. Also, it shows that their recognition of schwa in the mid location is higher than theirs in the front position. (For more details, see Table 2 in the Appendix-B)

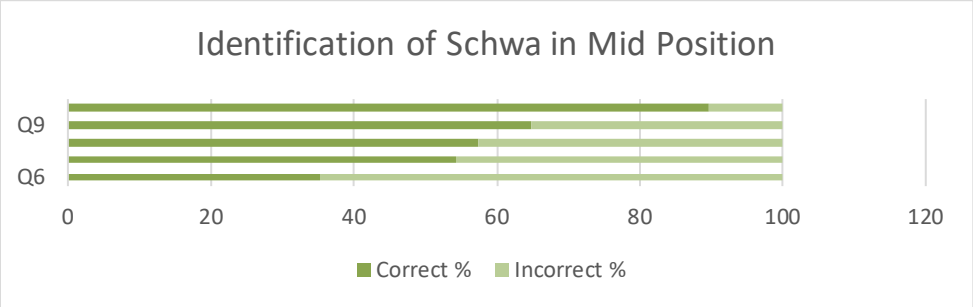


Figure 3. Identification percentage for medial position of schwa sound

Figure four below represents the numbers and percentages of the correct and incorrect answers. The test included five questions to choose the vowel that is pronounced as schwa in the final place in words. The mean of the correct answers of schwa in the final location is 54.41% while the mean of the incorrect answers is 45.59%. The results show that participants performed well in identifying schwa in the final position compared to the front and mid locations. (For more details, see Table 3 in the Appendix-B)

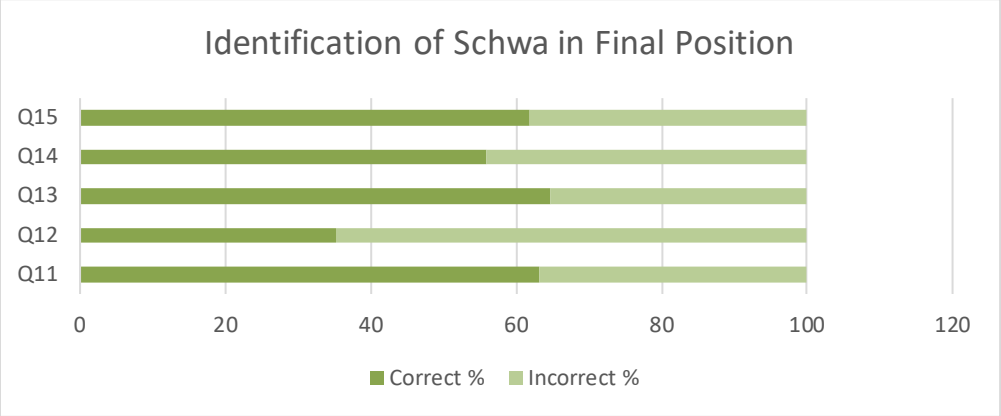


Figure 4. Identification percentage for final position of schwa sound

As all the above figures present the data for each single position with respect to five questions, the figure 5 below offers a comprehensive summary. The figure clearly shows that students have problems in identifying the position of schwa sound in variety of words. Learners showed a relatively balanced performance, with slightly more correct identifications (54.41%) than incorrect (45.59%) for the initial position. On the other hand, for the medial position, the highest accuracy was observed (60% correct), indicating that learners were more successful at identifying the schwa sound when it appeared in the middle of words. Performance in the final position was like the initial position, with a marginally higher rate of correct responses. Overall, the results indicate 54.3% of the correct identifications and 45.7% of incorrect identifications that means that almost half of the responses were not correct. This reveals that students have not much good understanding of the schwa sound with respect to various positions and need further guidance and practice.

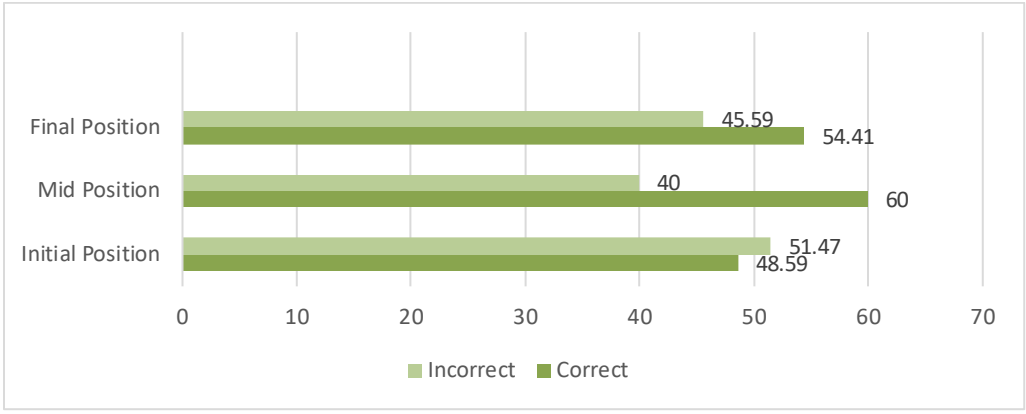


Figure 5. Comparison of all three positions

ISSN: 2229-9327

Discussion

This quantitative study provided precise data on English majors' ability to recognize the schwa sound in different word positions. The research used an equal number of questions for each schwa location to ensure a balanced assessment of students' performance. The results indicate that the participants struggled most with identifying schwa in the front position, with only half of the students answering correctly. Recognition in the mid-position was also weak but showed improvement compared to the front. Identification of schwa in the final position was the highest among the three, though still not very strong. Therefore, the most challenging schwa locations for English majors to identify, in order, are the front, mid, and finally, the final position.

Moreover, the data reveals significant insights into the challenges faced by Saudi EFL learners in identifying the schwa sound across different positions in words. The nearly equal distribution of correct and incorrect responses (51.47% correct) suggests that learners may struggle with the schwa sound when it occurs at the beginning of words. This could be attributed to the influence of the learners' native language, where such sounds may not exist or may be less prominent. The highest correct response rate (60%) in the mid position indicates that learners may have more exposure to the schwa sound in this context, perhaps due to its frequent occurrence in multi-syllable words. This finding suggests that instructional strategies focusing on word structure and syllable stress might enhance learners' abilities to identify schwa in various contexts. The performance in the final position (54.41% correct) mirrors that of the initial position. The slight improvement could reflect some familiarity with schwa sounds in word endings, but the overall rates still indicate a need for targeted phonetic training.

The higher error rate particularly in the initial position could be attributed to misperception of the Arab students where they confused it with long /a/ vowel. This could be due to their native language where /a/ sound in beginning is usually a long one and in their language, there is no concept of a short vowel like schwa. This finding agrees with Chouchane (2016) that attributed L1 interference an important factor for mispronunciation in the Arab learners. Another major reason could be the orthographic system where each single word (given in the test) like about, amount, etc. star with letter a but not pronounced with the same sound.

Overall, the results indicate that the participants need to improve their ability to identify the schwa sound in all positions, particularly in front of other sounds. This finding aligns with Bamigbade and Raji's (2022) study, which highlighted that Nigerian L2 learners of the English struggle with identifying and transcribing the schwa sound. Their research emphasized the importance of teaching phonetic transcription, particularly by utilizing applications to enhance this skill. Similarly, Mahmood et al. (2015) found that even after practicing transcription with nonsense words, students' recognition of the schwa remained low. Additionally, Kapranov's (2021) study on Norwegian learners supports these results, demonstrating that English learners generally face significant challenges in transcribing and recognizing the schwa sound.

Difficulties with pronouncing and transcribing the schwa sound are common challenges for English learners. To address these issues, several solutions are recommended. First, learners should focus on mastering the correct pronunciation of the schwa sound so they can accurately identify and produce it in words. Second, consistent practice transcribing words containing the schwa in various positions, whether in class with teachers, individually, or with peers, is essential. Ultimately, utilizing language learning applications that provide pronunciation guidance and transcription exercises can be highly beneficial. Improving the ability to recognize and identify the

schwa sound will significantly help English majors enhance their transcription skills. The results found that the front position was the most challenging location for recognizing schwa.

This research highlights the difficulties faced by female English majors in transcribing the schwa sound. However, the study had some limitations, particularly in developing the research tool and collecting data from participants, as outlined: the time allocated for this study was limited. Conducting the test on-site rather than online would have been preferable, as it could have allowed for a variety of question types. Specifically, the test could have included transcription and listening tasks, rather than multiple-choice questions, which might have provided more accurate insights. However, it is important to note that our methodology tested recognition rather than production. A test requiring active transcription may yield different results. Additionally, the online test was more accessible to students in the upper levels, but less so for those in the lower levels. Only two participants from the first level completed the test. Although the researcher aimed to obtain enough responses from each level, it was not achievable.

Conclusion and Recommendations

According to Shaibani (2023), "Although pronunciation is one of the most difficult aspects of English language teaching in Saudi Arabia, there has been very little attention paid to speech instruction" (p. 9). This research examined the most difficult schwa position to transcribe among 68 female English majors. The researcher used an online test to identify which schwa location posed the most significant challenge for the participants. The results revealed that recognition of the schwa sound was weakest in the front position, somewhat better in the mid position, and strongest in the final position. Nevertheless, participants showed low recognition across all positions and need to improve their skills. It is recommended that they apply the solutions suggested in this study and explore additional methods for improvement. The findings supported the research hypothesis and answered the research question. However, the study faced some limitations, including time constraints and limited access to participants.

The researcher recommends that students focus more on learning the sounds of English, especially vowel sounds, as this will help them pronounce words correctly and recognize each sound in the language. Additionally, it is recommended that English teachers place greater emphasis on teaching phonetics at the early stages of learning, enabling learners to acquire these skills more effectively than they would at later ages.

The researcher recommends further studies in this field, particularly focusing on vowel sounds. Given the students' difficulties with vowel pronunciation and transcription, the following areas are suggested for future research: Research should be conducted to explore how the Arabic language influences Saudi EFL learners' pronunciation, recognition, and transcription of the schwa sound. This would help identify clear strategies to assist learners in accurately recognizing, transcribing, and pronouncing schwa in English words. The researcher suggests employing an experimental method to investigate practical approaches to improve young learners' recognition of vowel sounds. Since these students learn English as a foreign language with limited practice opportunities, such research could provide valuable insights into enhancing their phonetic skills from an early stage.

Funding: This study is supported via funding from Prince Sattam bin Abdulaziz University, project number (PSAU/2025/R/1447).
Acknowledgment: This study is supported via funding from Prince Sattam bin Abdulaziz University, project number (PSAU/2025/R/1447).
Conflicts of Interest: The authors declare that they have no conflict of interest.
Authenticity: The manuscript is an original work.
Artificial Intelligence Statement: During the preparation of this work, the author used Grammarly AI to check the grammar.

About the Authors

Ghaida Albahlal is currently enrolled in the MA Linguistics Program at the Department of English, College of Science and Humanities, Kharj, Prince Sattam bin Abdulaziz University, Saudi Arabia. She has a keen interest in the Applied Linguistics domain, in particular ELT in the Saudi context. ORCID: https://orcid.org/0009-0006-1973-4727

Dr. Ansa Hameed is an Associate Professor in the Department of English, College of Science and Humanities, at Prince Sattam bin Abdulaziz University, Saudi Arabia. She is the author of numerous research papers published in various journals. Her areas of interest include Applied Linguistics, Discourse Analysis, Eco and Corpus Linguistics. ORCID: https://orcid.org/0000-0003-3477-156X

References

Alzinaidi, M. H., & Abdel Latif, M. M. M. (2019). Diagnosing Saudi students' English consonant pronunciation errors. *Arab World English Journal, 10* (4) 180- 193. https://dx.doi.org/10.24093/awej/vol10no4.14

Ariyani, L. D. (2019, March 30). Pronunciation error analysis of the English SCHWA sound produced by the students of UNSIQ. *Metaphor 1*(1), 1–15. https://ojs.unsiq.ac.id/index.php/metaphor/article/view/630

Bamigbade, W. A., & Raji, Z. O. (2022). Phonemic transcription error analysis of undergraduate English language students in southwestern Nigeria. *European Journal of English Language Teaching*, *6*(5). https://doi.org/10.46827/ejel.v6i5.4182

Beghoul, Y. (2017). Using the international phonetic alphabet in teaching English oral production and reception to Algerian learners as a potentially efficient solution to the spelling incoherence. *مجلة العلم والإيمان*, 63. https://doi.org/10.34174/0079-000-048-075

Cassimiro, A. J. J. (2011). Difficulties in dealing with English pronunciation. *Universidade Estadual de Goiás, Building the Way–Revista do Curso de Letras da UnU-Itapuranga, 1*(1), 48–57. https://www.revista.ueg.br/index.php/buildingtheway/article/view/159/131

Chouchane, A. (2016). Pronunciation difficulties for Arab learners of English. *Global English-Oriented Research Journal (GEORJ)*, *2*(2), 205–215.

Delahunty, G. P., & Garvey, J. J. (2010). *The English Language: From Sound to Sense.* https://doi.org/10.37514/per-b.2010.2331.

Elzaki, A. M. A. (2017). Investigating Difficulties Encountered by EFL Learners in Transcribing Words Including Diphthongs (Doctoral dissertation, Sudan University of Science and

Technology). https://www.academia.edu/96428320/Investigating_the_Difficulties_Encountered_by_EFL_Learners_in_Transcribing_Words_Including_Dip

Faraj, D.M. (2023). The effectiveness of learning activities in pronouncing the English vowel schwa by EFL college level students. *ZANCO Journal of Humanity Sciences, 27(5), 344–360.* https://doi.org/10.21271/zjhs.27.5.22

Heselwood, B. (2013). *Phonetic Transcription in Theory and Practice*. Edinburgh University Press.

Hudson, J. (2018). *The Sound of English: A Practical Course in British English Pronunciation.* Pronunciation Studio.

Kapranov, O. (2021). The English vowel schwa as a difficulty for Intermediate EFL students: Evidence from phonetic transcription. *Belgrade English Language and Literature Studies, 13,* 59–92. https://doi.org/10.18485/bells.2021.13.5

Kurtinienė, G. B. (2017). *Production, Recognition, Description, and Transcription of English Sounds.* Vilnius University Press: Lithuania.

Lindsey, G. (2019*). English after RP: Standard British Pronunciation Today*. Palgrave Macmillan: London.

Mahmood, A., Uzair, M., & But, M. I. (2015). Effectiveness of nonsense words in teaching phonemic transcription. *Gomal University Journal of Research, 31*(1), 155–164.

Mohanna, A. D. A. (2024). Difficulties and challenges encountered by Saudi EFL learners: A Diagnostic study. *Scholars International Journal of Linguistics and Literature, 7*(10), 288–299. https://doi.org/10.36348/sijll.2024.v07i10.002

Mompéan, J. A. (2017). Doing phonetic transcription in a modern language degree. *Estudios de Filología Inglesa: Homenaje al Profesor Rafael Monroy, Murcia: EditUM,* 479–505.

Ocampo, V. P. (2017). A Dual-Dimension Phonetic Analysis to Recognize Common Errors in the Use of the Sound schwa in Trained and Untrained Teachers. Universidad de Piura: Piura.

Philip, C. (2016). *English Phonetics and Phonology: An Introduction*. Wiley-Blackwell: New Jersey.

Roach, P. (2018). *English Phonetics and Phonology: A Practical Course*. Cambridge University Press: Cambridge.

Kondo, Y. (1995). *Production of schwa by Japanese speakers of English: A cross-linguistic study of co-articulatory strategies* (Unpublished Dissertation). https://era.ed.ac.uk/handle/1842/20618

Shaibani, A. A. (2023). Challenges in teaching pronunciation to Saudi female learners at TAIF University. *Arab World English Journal, 292,* 1–23. https://doi.org/10.24093/awej/th.292

Tench, P. (2011). *Transcribing the Sound of English*. Cambridge University Press: Cambridge.

Trzeciakowska, J. (2016). Mistakes in phonemic transcriptions made by Polish EFL teacher training college students. *Journal of Young English Philology Thought and Review.* https://repozytorium.umk.pl/handle/item/4091

Weber, R.–M. (2018). Listening for schwa in academic vocabulary. *Reading Psychology, 39*(5), 468–491. https://doi.org/10.1080/02702711.2018.1464531

Yavas, M. (2011). *Applied English Phonology.* Wiley-Blackwell: New Jersey.

Young, R. & Messum, P. (2022). Teaching schwa: Using intentional 'stuttering' on consonants to improve English pronunciation. In J. Levis & A. Guskaroska (eds.), *Proceedings of the*

12th Pronunciation in Second Language Learning and Teaching Conference, held June 2021 virtually at Brock University, St. Catharines, ON. https://doi.org/10.31274/psllt.13367

Zsiga, C. E. (2014*). The Sounds of Language: An Introduction to Phonetics and Phonology*. Wiley-Blackwell: New Jersey.

Appendix- A

Test Questions

Initial position of schwa

1. Choose the vowel that is pronounced as schwa in the word Advise
2. Choose the vowel that is pronounced as schwa in the word Ago
3. Choose the vowel that is pronounced as schwa in the word Again
4. Choose the vowel that is pronounced as schwa in the word Across
5. Choose the vowel that is pronounced as schwa in the word Amount

Medial position of schwa

6. Choose the vowel that is pronounced as schwa in the word Forget
7. Choose the vowel that is pronounced as schwa in the word Album
8. Choose the vowel that is pronounced as schwa in the word Understand
9. Choose the vowel that is pronounced as schwa in the word Husband
10. Choose the vowel that is pronounced as schwa in the word Happiness

Final position of schwa

11. Choose the vowel that is pronounced as schwa in the word Mother
12. Choose the vowel that is pronounced as schwa in the word Learner
13. Choose the vowel that is pronounced as schwa in the word Smaller
14. Choose the vowel that is pronounced as schwa in the word Horror
15. Choose the vowel that is pronounced as schwa in the word Picture

Appendix -B

Table 1: Front position

Questions	**Frequency of Correct Answers**	**Percentage %**	**Mean**	**Frequency of Incorrect Answers**	**Percent %**	**Mean**
Choose the vowel that is pronounced as schwa in the word *Advise*	36	52.9%	48.59%	32	47.1%	51.47%
Choose the vowel that is pronounced as schwa in the word *Ago*	45	66.2%		23	33.8%	
Choose the vowel that is pronounced as schwa in the word *Again*	49	72.1%		19	27.9%	

Choose the vowel that is pronounced as schwa in the word *Across*	42	61.8%		26	38.2%	
Choose the vowel that is pronounced as schwa in the word *Amount*	0	0.0%		68	100%	

Table 2: Mid position

Questions	**Frequency of Correct Answers**	**Percent %**	**Mean**	**Frequency of Incorrect Answers**	**Percent %**	**Mean**
Choose the vowel that is pronounced as schwa in the word *Forget*	24	35.3%	60%	44	64.7%	40%
Choose the vowel that is pronounced as schwa in the word *Album*	37	54.4%		31	45.6%	
Choose the vowel that is pronounced as schwa in the word *Understand*	39	57.4%		29	42.6%	
Choose the vowel that is pronounced as schwa in the word *Husband*	44	64.7%		24	35.3%	
Choose the vowel that is pronounced as schwa in the word *Happiness*	61	89.7%		7	10.3%	

Table 3: Final position

Questions	**Frequency of Correct Answers**	**Percent %**	**Mean**	**Frequency of Incorrect Answers**	**Percent %**	**Mean**
Choose the vowel that is pronounced as schwa in the word *Mother*	43	63.2%	54.41%	25	36.8%	45.59%
Choose the vowel that is pronounced as schwa in the word *Learner*	24	35.3%		44	64.7%	
Choose the vowel that is pronounced as schwa in the word *Smaller*	44	64.7%		24	35.3%	
Choose the vowel that is pronounced as schwa in the word *Horror*	38	55.9%		30	44.1%	
Choose the vowel that is pronounced as schwa in the word *Picture*	42	61.8%		26	38.2%	

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.345-358
DOI: https://dx.doi.org/10.24093/awej/vol16no3.21

Investigating the Relationship Between Self-Regulated Writing Strategies and Academic Writing Performance

Xiaolan Wang
Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Malaysia
Corresponding Author: p134016@siswa.ukm.edu.my

Supyan Hussin
Institute of Ethnic Studies (KITA), Universiti Kebangsaan Malaysia, Bangi, Malaysia
E-mail: supyan@ukm.edu.my

Azlina Abdul Aziz
Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Malaysia
E-mail: azlina1@ukm.edu.my

Received:06/02/2025 Accepted: 08/14/2025 Published: 09/20/2025

Abstract

This study investigates the relationship between self-regulated writing strategies (SRWS) and academic writing performance among Chinese undergraduate students learning English as a Foreign Language (EFL). Its primary purpose is to determine which dimensions of SRWS—cognitive, metacognitive, motivational and social—most strongly predict writing performance, thereby informing pedagogy in an exam-oriented context where learner autonomy remains under-developed. Employing a convergent parallel mixed-methods design, 300 students completed the Writing Strategies for Self-Regulated Learning Questionnaire (WSSRLQ) and a standardised academic writing task. Semi-structured interviews were conducted with 10 students. Quantitative analyses, including correlation and regression tests, revealed significant positive associations between SRWS dimensions—particularly goal-setting, monitoring, and motivational self-talk—and writing performance. Students with high SRWS demonstrated notably superior writing outcomes (Cohen's d = 2.04). Qualitative insights from interviews and text analysis indicated that high-performing students employed deliberate planning and strategic revision, while low-performing peers exhibited reactive behaviours and limited emotional regulation. Demographic variables such as academic level, prior training, and frequency of Automatic Writing Evaluation (AWE) tool use significantly influenced SRWS engagement. The findings highlight the predictive value of SRWS and underscore the need to incorporate strategy-based instruction and scaffolded feedback into EFL writing curricula. Pedagogical implications suggest fostering metacognitive and affective regulation to support autonomous and proficient academic writers. Future work should embed explicit, sustained SRWS instruction—augmented by digital feedback and collaborative tasks—into EFL curricula to cultivate autonomous, strategic writers.

Keywords: Academic writing performance, automated writing evaluation, English as a foreign language, self-regulated writing strategies

How to cite: Wang, X., Hussin, S. & Aziz, A. A. (2025). Investigating the Relationship Between Self-Regulated Writing Strategies and Academic Writing Performance. *Arab World English Journal, 16*(3): 345-358. https://dx.doi.org/10.24093/awej/vol16no3.21

Introduction

Academic writing is a cognitively demanding activity that poses significant challenges for second language learners, particularly those in EFL contexts. It requires not only linguistic competence but also the ability to manage complex thinking, structure ideas logically, and regulate the writing process effectively. In response to these demands, the concept of self-regulated learning (SRL) has gained increasing traction in applied linguistics and educational psychology. SRL is defined as the degree to which learners are metacognitively, motivationally, and behaviourally active participants in their own learning process (Zimmerman, 2000). Within the context of writing, this encompasses a specific subset of behaviours known as SRWS, which include planning, goal-setting, self-monitoring, revising, and emotional control (Teng & Zhang, 2016).

In second language writing research, SRWS are increasingly acknowledged as crucial determinants of writing performance. Learners who actively engage in goal-oriented planning, monitor their progress, revise their drafts critically, and regulate their emotions tend to produce more coherent and effective texts (O'Dowd et al., 2020). These strategies enable writers to cope with the cognitive load of composing in a second language, thereby facilitating greater autonomy and higher achievement. Moreover, the development of SRWS is seen as an essential component of academic literacy, preparing learners for more complex writing tasks in academic and professional contexts. Despite growing scholarly attention, however, research on SRWS remains fragmented and contextually limited.

Much of the existing literature conceptualises SRL as a general framework and tends to overlook the nuanced and multidimensional nature of SRWS in actual writing tasks. Several studies have examined the impact of individual strategies, such as planning or self-monitoring, yet few have addressed how different dimensions—cognitive, metacognitive, motivational, and social—interact and influence writing outcomes in tandem (Huang et al., 2023; Granado-Peinado et al., 2019). Additionally, while SRWS instruments such as the WSSRLQ developed by Teng and Zhang (2016) offer a robust measurement framework, empirical applications are still limited in scale and scope, particularly in non-Western contexts.

In the Chinese EFL context, research on SRWS is particularly scarce. Given the exam-oriented nature of the education system and its emphasis on teacher-centred instruction, students may have fewer opportunities to engage in strategic learning behaviours (Shen & Bai, 2024). This instructional culture may constrain learners' development of autonomy, particularly in writing classes where process-oriented instruction and peer collaboration are often underutilised. Similar instructional challenges have also been noted in other EFL contexts, such as the Arab region, where teacher-fronted practices and limited student autonomy have been shown to hinder the development of strategic writing behaviours (Alrabai, 2014). Furthermore, while studies have acknowledged the increasing use of AWE tools such as Pigai or Grammarly, there is limited understanding of how these digital tools influence or interact with students' use of SRWS. As digital feedback becomes more prevalent in writing instruction, understanding its relationship with learners' self-regulatory behaviours is essential for informed pedagogical decision-making (Teng et al., 2022).

Another limitation of previous studies is their predominant reliance on correlational designs. Although useful, these designs offer limited explanatory power regarding how students apply SRWS in actual writing tasks or how these strategies vary across learners of differing proficiency or academic backgrounds. Few studies have compared high and low strategy users in meaningful ways or triangulated data from multiple sources, such as writing samples,

questionnaires, and interviews. This gap is significant, as understanding not just the presence but the application of SRWS can offer more actionable insights for instruction.

To address these limitations, the current study investigates the relationship between SRWS and academic writing performance among Chinese EFL university students using a convergent parallel mixed-methods design. This approach integrates quantitative data from the WSSRLQ and writing assessments with qualitative data from semi-structured interviews, thereby providing a comprehensive and context-sensitive account of learners' strategic behaviours. Specifically, the study seeks to examine which SRWS dimensions are most strongly associated with writing performance, whether SRWS can serve as predictors of academic achievement, and how high and low SRWS users differ in their writing practices.

This study is both timely and significant. Theoretically, it contributes to the refinement of SRL and SRWS frameworks by situating them within real academic writing contexts in under-researched educational settings. It also responds to calls for more ecologically valid, mixed-methods research that can capture the dynamic and situated nature of writing strategy use. Practically, the findings have the potential to inform writing pedagogy in Chinese universities by identifying which strategies most strongly influence performance and should therefore be explicitly taught and scaffolded. In an era of increasing reliance on digital tools and growing emphasis on learner autonomy, this study offers critical insights into how strategy instruction and feedback can be integrated to support student writers. To guide the investigation, the following research questions were formulated:

1. What are the levels of SRWS among Chinese EFL learners?
2. Are different dimensions of SRWS significantly associated with academic writing performance?
3. Can SRWS serve as predictors of students' writing achievement?
4. Are there significant performance differences between students with high and low SRWS?

By addressing these questions, the study aims to contribute to a deeper understanding of the role of SRWS in academic writing development and to provide empirical evidence to support more effective instructional practices in EFL writing classrooms.

Literature Review

To address these questions, it is necessary to examine prior studies that have explored the theoretical foundations and empirical patterns of SRWS in second language contexts. SRL, grounded in social cognitive theory (Bandura, 1986), emphasises learners' active control over their cognition, motivation, and behaviour through goal-setting, monitoring, and reflection (Zimmerman, 2000). Zimmerman's triadic model—forethought, performance, and self-reflection—has been widely applied to second language writing research, as academic writing demands sustained strategic engagement. In particular, the role of SRWS has gained attention, referring to learners' ability to apply cognitive (e.g., planning), metacognitive (e.g., monitoring), and affective (e.g., emotion control) strategies throughout the writing process (Teng & Zhang, 2016).

Pintrich (2004) further expanded SRL by including contextual and motivational regulation, emphasising the multidimensional nature of writing strategies. In EFL contexts, this means writers must manage not only linguistic accuracy but also psychological and environmental demands. Zhang and Zhang (2025) revealed "distinct profiles of writing self-regulation, ranging from high-

strategy users to disengaged learners" (p. 1545), indicating that individual variability plays a critical role in shaping the effectiveness of SRWS. As Teng (2022) observed, effective second language writing is shaped by learners' regulatory competence across multiple domains.

Empirical studies have confirmed the importance of SRWS in second language writing. For instance, Teng and Huang (2021) found that metacognitive and motivational strategies significantly enhance writing fluency. Likewise, Teng and Zhang (2020) demonstrated that learners' emotional control and confidence-building practices are strongly associated with performance in high-stakes academic writing. However, not all findings are consistent. While Dong and Zhang (2022) emphasised the predictive power of motivational regulation, Granado-Peinado et al. (2019) found planning and revision to be more influential, especially in argumentative writing. These discrepancies suggest a need to examine SRWS not in isolation but as interacting dimensions.

Recent research has also expanded the SRWS framework by integrating reading-to-write approaches in second language academic contexts (Bai & Wang, 2023), highlighting the interdependence of comprehension and production strategies. Differences in SRWS application across writing formats (e.g., paragraph vs essay writing) further underscore the contextual variability of strategic writing behaviours (Umamah & Cahyono, 2020). As Abdelaty et al. (2025) observed, "feedback quality and task complexity jointly mediated the relationship between self-regulation and L2 writing achievement" (p. 6), underscoring the importance of instructional design in activating learners' strategic engagement. Building on this perspective, Stell and Iwashita (2024) argued that "co-regulated writing strategies emerge through collaborative planning, joint monitoring, and negotiated revision, which scaffold learners' individual regulation over time" (p. 6), reinforcing the value of peer-based regulation in EFL writing contexts.

Several validated instruments, such as the WSSRLQ (Teng & Zhang, 2016), have been developed to measure SRWS. The WSSRLQ captures a comprehensive range of strategies—text processing, feedback handling, motivational self-talk, and emotional regulation. However, many studies using such tools focus solely on correlation analyses, overlooking how different learners apply strategies in practice or across performance levels. As Dörnyei and Ryan (2015) argue, self-regulation is "a dynamic construct that connects strategic capacity, intent, and learning behaviour" (p. 169), underscoring the need for SRWS research to move beyond simple statistical associations and explore how strategies unfold in real writing contexts.

Moreover, most prior studies have concentrated on single-strategy training or treated SRWS as a homogeneous construct, often neglecting cultural and contextual influences. In Chinese EFL classrooms, for example, peer learning and revision may be underutilised due to exam-oriented instruction (Wang, 2023). This contextual variation highlights the need for research that captures how SRWS operate across diverse learners and instructional settings. Similar patterns have been observed in other cultural contexts, such as Saudi Arabia, where course memory and text processing emerged as the most powerful SRWS predictors of writing performance (Alanazi, 2024).

Method

This study employed a convergent parallel mixed-methods design (Creswell & Clark, 2017) to examine the relationship between SRWS and academic writing performance among Chinese EFL university students. Quantitative and qualitative data were collected and analysed independently before being integrated to provide a comprehensive understanding of students'

strategic writing behaviours. This design allowed for triangulation and ensured both breadth and depth of interpretation (Chen et.al, 2023).

Participants

The participants were 300 students from three comprehensive universities in northern China, selected via purposive sampling to ensure recent exposure to academic writing instruction. Eligibility criteria included completion of at least two semesters of English courses and willingness to complete a questionnaire and a timed writing task. Of the sample, 180 were female and 120 male, comprising 270 undergraduates, 24 Master's students, and six doctoral candidates. Additionally, 184 participants reported receiving formal academic writing training, and proficiency levels were self-reported based on College English Test and Test for English Major standards. Table 1 presents a summary of participant characteristics :

Table 1. *Demographics of the respondents*

Variable	Group	N
Gender	Female / Male	180 / 120
Academic Level	Undergraduate / Master / PhD	270 / 24 / 6
Writing Training	Yes / No	184 / 116

Instruments

Three main instruments were used for data collection. Firstly, the WSSRLQ (Teng & Zhang, 2020), a validated 40-item Likert-scale instrument, was used to assess SRWS across cognitive, metacognitive, motivational, and social dimensions. It has demonstrated strong internal reliability (Cronbach's α = .73–.87) and is widely used in Chinese EFL contexts.

Secondly, a timed expository writing task was administered to assess students' academic writing performance. Essays (250–300 words) were rated by two trained raters using an analytic rubric measuring coherence, organisation, language use, and content development. Inter-rater reliability was confirmed with an ICC of .87.

Lastly, semi-structured interviews were conducted with 10 students—five from the highest and five from the lowest scoring quartiles—to explore their perceptions and actual application of SRWS. Interview protocols were based on SRL theory (Zimmerman, 2000) and conducted in Mandarin for clarity.

These three instruments were selected to comprehensively address the research questions by capturing the reported use of SRWS (RQ1), their relationship with writing performance (RQ2–RQ3), and learners' perceptions and behaviours (RQ4).

Procedures

The study was conducted in four phases over eight weeks. In Phase one, participants completed the WSSRLQ online. In Phase two, they undertook the standardised writing task under exam-like conditions. In Phase three, writing scores were used to stratify students into high and low-performing groups. Finally, Phase four involved conducting and transcribing follow-up interviews with selected participants. All qualitative data were analysed using Braun and Clarke's (2006) thematic analysis framework, supported by NVivo 12. Quantitative data were analysed using SPSS 27, employing Pearson correlations, multiple regression, t-tests, and ANOVA.

Ethical approval was obtained from the university's research board. All participants provided informed consent and were assured of confidentiality. This rigorous design ensured the reliability and replicability of the findings.

Results

This section presents the main findings of the study, based on the analysis of quantitative and qualitative data. The results are organised into five subsections aligned with the research questions: (1) levels of SRWS among students, (2) correlations between SRWS and writing performance, (3) predictive power of SRWS, (4) performance differences between high and low SRWS groups, and (5) differences across demographic variables.

Levels of SRWS

The descriptive statistics revealed moderate-to-high levels of SRWS among Chinese EFL learners. The overall mean score on the WSSRLQ was 172.76 (SD = 25.78). Among the nine subscales, Motivational Self-Talk had the highest mean (M = 34.97, SD = 5.66), followed by Goal-Oriented Monitoring and Evaluating (M = 25.89, SD = 4.55) and Text Processing (M = 25.72, SD = 4.80). Lower scores were recorded for Peer Learning (M = 12.23, SD = 3.15), Course Memory (M = 12.71, SD = 2.40), and Idea Planning (M = 12.89, SD = 2.42), indicating limited engagement with collaborative and foundational cognitive strategies.

Table 2. *Descriptive statistics for SRWS dimensions (N = 300)*

SRWS Dimension	Mean	SD
Motivational Self-Talk	34.97	5.66
Goal-Oriented Monitoring & Evaluating	25.89	4.55
Text Processing	25.72	4.80
Feedback Handling	17.87	3.37
Interest Enhancement	17.45	3.26
Emotional Control	13.03	2.37
Course Memory	12.71	2.40
Idea Planning	12.89	2.42
Peer Learning	12.23	3.15

Correlations Between SRWS and Writing Performance

Pearson correlation analysis showed that all SRWS dimensions were positively associated with writing scores. The strongest correlations were found for Goal-Oriented Monitoring and Evaluating (r = .41), Course Memory (r = .39), and Idea Planning (r = .38), suggesting the importance of metacognitive and cognitive regulation in writing performance.

Table 3. *Pearson correlations between SRWS dimensions and writing performance (N = 300)*

SRWS Dimension	r
Goal-Oriented Monitoring & Evaluating	0.41
Course Memory	0.39
Idea Planning	0.38
Emotional Control	0.36
Motivational Self-Talk	0.34
Interest Enhancement	0.33
Feedback Handling	0.32
Text Processing	0.30
Peer Learning	0.27

These findings affirm that stronger self-regulatory skills are associated with better writing outcomes. However, the comparatively weaker correlations of social strategies such as Peer Learning highlight possible cultural or instructional constraints limiting collaborative engagement.

Predictive Power of SRWS on Writing Performance

Stepwise multiple regression identified five SRWS dimensions as significant predictors of writing scores. Course Memory emerged as the strongest predictor ($\beta = .56$, $p < .001$), followed by Feedback Handling ($\beta = .45$, $p < .001$), Motivational Self-Talk ($\beta = .26$), Goal-Oriented Monitoring and Evaluating ($\beta = .22$), and Peer Learning ($\beta = .19$). These results highlight the importance of both internal and external regulatory mechanisms.

The inclusion of Peer Learning as a significant, albeit weaker, predictor is noteworthy, suggesting that even marginal levels of social interaction contribute to performance. Nonetheless, the absence of stronger effects raises critical questions about the underdevelopment of collaborative writing culture in Chinese EFL classrooms.

Differences Between High and Low SRWS Groups

Independent samples t-test compared students with high and low SRWS. The high-SRWS group (N = 124) scored significantly higher in writing (M = 15.47, SD = 1.52) than the low-SRWS group (N = 105, M = 11.23, SD = 2.53), $t(178.34) = -16.32$, $p < .001$, with a large effect size (Cohen's $d = 2.04$).

Table 4. *Writing performance by SRWS group*

Group	N	Mean	SD	t(df)	p	d
Low SRWS	105	11.23	2.53	-16.32(178.34)	<.001	2.04
High SRWS	124	15.47	1.52			

The considerable performance gap suggests that fostering SRWS may be essential for narrowing achievement disparities. Nevertheless, the nature of strategy use—whether strategic planning or mechanical application—requires deeper exploration. Some low-performing students may imitate surface strategies without internalising them, indicating a need for qualitative insights.

Demographic Differences in SRWS Use

ANOVA and t-tests revealed no significant gender differences in SRWS use. However, academic level influenced SRWS scores significantly, $F(2, 297) = 9.71$, $p < .001$. Postgraduate students reported higher SRWS engagement than undergraduates. Students with prior writing training scored significantly higher on SRWS (M = 181.46) than those without (M = 168.96), $t(298) = -2.38$, $p = .018$.

Additionally, Pearson correlation analysis indicated that both self-reported English proficiency ($r = .212$, $p < .001$) and frequency of AWE tool use ($r = .224$, $p < .001$) were positively related to SRWS scores.

Table 5. *Comparison of SRWS scores across key variables*

Variable	Group	M	SD	Test	p
Gender	Male / Female	171.91 / 172.97	52.61 / 39.45	t = -0.17	.865
Academic Level	UG / Master / PhD	170.85 / 206.17 / 243.40	40.99 / 41.00 / 44.52	F(2,297) = 9.71	<.001
Writing Training	Yes / No	181.46 / 168.96	41.69 / 42.08	t = -2.38	.018

Table 6. *Correlation between SRWS and proficiency variables*

Variable	r	p
English Proficiency	.212	<.001
AWE Tool Use Frequency	.224	<.001

Frequent users of the AWE tool outperformed occasional and non-users in writing performance. Significant differences were observed between Group 0 (non-users, M = 12.74), Group 1 (occasional, M = 13.46), and Group 2 (frequent, M = 15.38), with effect sizes ranging from d = 0.72 to 0.89.

Table 7. *Writing scores by AWE tool use frequency*

Group	M	SD	Comparison	t	p	d
Non-users (0)	12.74	3.37	0 vs 1	-1.63	.104	-0.24
Occasional (1)	13.46	2.76	0 vs 2	-5.81	<.001	-0.89
Frequent users (2)	15.38	2.47	1 vs 2	-5.31	<.001	-0.72

While these differences suggest AWE tools support strategic revision, they also invite further inquiry: do high-frequency users engage more reflectively, or do they merely respond to surface-level prompts? Future studies may benefit from triangulating AWE logs with revision narratives to clarify the depth of student engagement. Moreover, reliance on self-reported frequency may overlook qualitative variation in tool use, reinforcing the value of mixed-method approaches.

Discussion

Interpretation of SRWS Levels, Patterns, and Influencing Factors among second language Writers

The overall SRWS scores suggested a moderate level of strategy use among Chinese EFL university students, with variations across demographic and experiential factors. While gender did not significantly influence SRWS levels, academic level emerged as a meaningful variable; postgraduate and doctoral students reported substantially higher SRWS engagement than undergraduates. This pattern may reflect the cumulative impact of academic experience on metacognitive awareness and strategic sophistication. These findings resonate with those of Teng and Zhang (2016), who also reported moderate-to-high SRWS usage among EFL students, suggesting growing recognition of the importance of planning, monitoring, and motivational regulation in academic writing.

Notably, prior exposure to formal academic writing training was associated with significantly greater SRWS use, underscoring the potential benefits of targeted instructional support. Furthermore, weak but significant positive correlations were found between SRWS and both self-reported English proficiency and frequency of AWE tool usage. The positive association between AWE tool frequency and SRWS echoes Teng (2022), who found that digital tools support learners' metacognitive behaviours and strategy engagement.

Cultural factors may also influence SRWS tendencies. Within the Chinese EFL context, where exam-oriented instruction dominates, learners may underutilise strategies related to planning and revising, while focusing more on local-level accuracy. Additionally, this study identified particularly low use of social strategies such as peer learning, which suggests a persistent individualistic orientation in Chinese writing pedagogy—despite policy efforts to promote communicative approaches (Teo, 2017). This aligns with observations from other teacher-centred

contexts (Al-Ghamdi & Alsehibany, 2021), where metacognitive development is hindered by limited student interaction. The observed variability in SRWS patterns thus appears to be shaped not only by individual proficiency and academic maturity but also by systemic instructional norms. This highlights the importance of explicitly fostering strategic writing behaviours through pedagogical intervention.

Linking SRWS Dimensions with Writing Performance

The relationship between SRWS performance was further illuminated by the statistical and qualitative findings of this study. Quantitative analysis indicated that students who reported higher levels of SRWS—particularly in dimensions such as goal-setting, monitoring, and revising—tended to achieve better writing scores. This aligns with the theoretical perspectives of Pintrich (2004) and Zimmerman (2000), both of whom emphasise the interplay of cognitive, metacognitive, and motivational processes in SRL. Consistent with these models, this study found that metacognitive strategies, including goal monitoring and evaluation, demonstrated the strongest correlations with writing performance, reinforcing prior findings that strategic self-regulation is integral to academic success (Teng & Huang, 2021).

From a cognitive and metacognitive standpoint, successful writers in this study demonstrated greater engagement with strategies that involved pre-task planning, self-monitoring during writing, and reflective post-task evaluation. These behaviours are consistent with Zimmerman's cyclical model, where forethought, performance control, and self-reflection support improvement. Students who showed high SRWS scores appeared to engage more fully in all three phases, suggesting more developed self-regulatory routines that translate into higher-quality written output.

Furthermore, the results highlighted the predictive power of both internal (e.g., course memory, motivational self-talk) and external (e.g., feedback handling) strategies. Notably, course memory emerged as the strongest predictor—a finding not widely reported in previous SRWS studies, possibly reflecting the critical importance of content internalisation and knowledge retrieval in exam-based writing contexts (Granado-Peinado et al., 2019).

The data also revealed that affective regulation—though often overlooked—played a role in differentiating writing performance. Students who managed writing-related anxiety, maintained confidence, or reported positive affect toward revision were more likely to demonstrate sustained engagement in the writing process. This supports Pintrich (2004)'s view that self-regulation includes not only cognition and motivation but also emotional control, which may be especially relevant in high-stakes academic writing tasks.

The role of the AWE tool appears to reinforce this self-regulatory cycle. Frequent users of the tool reported more strategic revisions and higher writing scores, suggesting that AWE feedback, when meaningfully integrated into students' metacognitive routines, can scaffold the development of self-regulation (Teng, 2022). The tool may thus serve as a form of external regulation that, over time, supports the internalisation of strategic writing behaviours.

Disparities between High and Low SRWS Groups

The comparison between high and low SRWS groups revealed substantial differences in writing performance. This result highlights the strong association between self-regulated writing behaviours and academic writing achievement. The size of the performance gap reinforces the critical role of SRWS in shaping outcomes (cf. Zimmerman & Schunk, 2011), while also pointing

to systemic inequities in writing instruction. Students in the low-SRWS group may lack explicit strategy training, particularly in Chinese EFL contexts where instruction remains exam-focused and teacher-centred (Teo, 2017).

High-SRWS students likely engaged in deliberate planning, active monitoring, and critical post-task evaluation—processes aligned with Zimmerman's cyclical model of SRL. In contrast, low-SRWS students may have focused on surface-level corrections rather than strategic content development or organisation. These disparities suggest that writing success is influenced not only by language proficiency, but by the extent to which learners apply metacognitive and cognitive strategies throughout the writing process.

Demographic patterns identified in earlier analyses suggest that undergraduates and those without prior academic writing training were more likely to fall into the low-SRWS group. This reinforces the importance of early instructional interventions to foster self-regulatory habits. As highlighted in this study, learners benefit from strategy-based instruction that fosters awareness, habit formation, and confidence in managing writing tasks.

Pedagogical Implications

The findings of this study underscore the importance of integrating strategy-oriented approaches into academic writing instruction within EFL contexts, particularly in Chinese higher education. The evident performance disparity between high and low SRWS groups illustrates that writing success is not merely a function of linguistic knowledge but closely tied to learners' ability to regulate their writing processes cognitively, metacognitively, and affectively. These findings align with earlier work emphasising that skilled writers are characterised by sustained goal monitoring, self-evaluation, and motivational self-talk (Teng & Huang, 2021).

While strategy use is often assumed to develop naturally with experience, the data indicate that without structured instructional support—such as that received through formal academic writing training—many students, especially undergraduates, lack the awareness or initiative to engage in planning, monitoring, and revision. This suggests that metacognitive behaviours must be explicitly taught rather than expected to emerge incidentally. Recent intervention studies confirm that combining teacher feedback with structured self-assessment tasks significantly enhances SRWS among EFL learners (Vasu et al., 2020).

The positive association between tool usage frequency and both SRWS levels and writing performance further illustrates the potential of AWE tools to function as external scaffolds for internal regulatory processes. Frequent tool users may have developed greater awareness of revision patterns and error correction over time, which in turn reinforced their strategic engagement. As this study found, learners who engaged regularly with AWE tools reported stronger strategy use and achieved higher performance outcomes (Teng, 2022). Online SRL tools such as self-monitoring charts have similarly been shown to support task engagement and metacognitive awareness in digital learning environments (Seker & Karagul, 2022).

In addition, the results highlight the need to consider affective regulation as a meaningful component of strategy use. Emotional responses to feedback, self-confidence, and anxiety management were all implicated in students' ability to persist through drafting and revising, aligning with established self-regulation models.

Limitations and Directions for Future Research

This study has several limitations that should be acknowledged. The sample was limited to one university in China, which may affect the generalisability of the findings to broader EFL populations. Differences in institutional writing practices and student demographics may have influenced both SRWS use and writing performance. Additionally, the reliance on self-reported data to measure strategy use may have introduced bias, as students may not always accurately reflect their cognitive and metacognitive behaviours. While writing scores served as performance indicators, they were based on a single post-intervention task and may not fully capture long-term writing development.

Future research should consider more diverse and representative samples across institutions and academic disciplines. Longitudinal studies could offer insights into how SRWS evolve and in response to instructional interventions. Moreover, integrating process-tracing methods such as think-aloud protocols or screen-capture tools could provide richer data on students' actual strategy use. Further studies may also explore how strategy-based instruction can be effectively integrated into EFL curricula, particularly through the use of digital tools. Notably, strategy-based instruction models grounded in SRL theory have been shown to outperform traditional writing instruction in improving academic writing skills (Anggraeni et al., 2025). In addition, greater attention to affective and motivational factors may yield deeper understanding of how emotional regulation supports or hinders learners' strategic writing engagement.

Conclusion

This study examined the relationship between SRWS and academic writing performance among Chinese EFL university students. The findings confirmed that SRWS play a significant role in shaping writing outcomes. Students with higher levels of cognitive, metacognitive, and affective regulation performed substantially better than their peers, highlighting the importance of fostering strategic writing behaviours.

The predictive power of SRWS—especially strategies such as goal-oriented monitoring, feedback handling, and course memory—suggests that writing instruction should go beyond teaching language accuracy and incorporate strategy training explicitly. These strategies are essential for building learners' autonomy and improving their ability to plan, evaluate, and revise effectively.

The study also found that academic level, writing training, and frequency of AWE tool use significantly influenced SRWS engagement. These findings emphasise the need for scaffolded instruction, particularly for undergraduates and novice writers. Moreover, the supportive role of AWE tools in enhancing writing strategy use reinforces their value when integrated into pedagogy alongside reflective guidance.

This research contributes to filling the gap in SRWS literature in non-Western EFL contexts by adopting a mixed-methods approach that offers a holistic view of learners' behaviours. It also suggests that writing pedagogy in EFL settings must be reoriented to develop not just writing skills, but also the strategic habits that underpin long-term writing success.

Future research should explore how SRWS evolve over time and how different instructional contexts influence their development. A greater focus on process-oriented assessments and teacher training will be essential to support the cultivation of self-regulated writers in diverse educational environments.

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 in order to proofread. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the content of the publication.

About the Authors

Xiaolan Wang is currently pursuing her PhD in TESL at the Faculty of Education, Universiti Kebangsaan Malaysia (UKM). She graduated with a Master's degree from the University of York. Her research interests are in second language acquisition and cross-cultural communication.
https://orcid.org/0009-0002-9878-6453

Professor Dr Supyan Hussin is a Principal Fellow at the Institute of Ethnic Studies, Universiti Kebangsaan Malaysia (UKM). His main research interests include language education, teacher training, educational technology, and Malay civilisation studies.
https://orcid.org/0000-0003-0544-0649

Dr Azlina Abdul Aziz is a senior lecturer in the TESL programme at the Centre for Teaching and Learning Innovations, Universiti Kebangsaan Malaysia (UKM). Her main research interests include the teaching of literature in TESL, teacher education, and the use of narrative and literary texts to explore social and cultural issues.
https://orcid.org/0000-0002-7800-3688

References

Abdelaty, Y., Alghamdi, E., & Alamer, A. (2025). The effects of self-regulation on L2 writing achievement: Examining the mediating roles of feedback quality and task complexity. *System, 132,* 103715. https://doi.org/10.1016/j.system.2025.103715

Alanazi, M. H. (2024). Predictive effects of self-regulated writing strategies on writing performance among Saudi EFL learners. *Journal of Educational and Psychological Studies, 14*(4), 668–678. https://doi.org/10.53543/jeps.vol14iss4pp668-678

Al-Ghamdi, W. S., & Alsehibany, R. A. (2021). A study on classroom discourse in enhancing metacognition to develop writing skill in Saudi EFL students. *Arab World English Journal, 12*(3), 450–463. https://doi.org/10.24093/awej/vol12no3.31

Alrabai, F. (2014). Motivational practices in English as a foreign language classes in Saudi Arabia: Teachers' beliefs and learners' perceptions. *Arab World English Journal*, *5*(1), 180–196. https://awej.org/images/AllIssues/Volume5/Volume5Number1March2014/15.pdf

Anggraeni, C. W., Mujiyanto, J., Rustipa, K., et al. (2025). Effects of utilizing self-regulated learning-based instruction on EFL students' academic writing skills: a mixed-method investigation. *Asian Journal of Second and Foreign Language Education*, *10*, 15. https://doi.org/10.1186/s40862-024-00317-6

Bai, B., & Wang, J. (2023). Conceptualising self-regulated reading-to-write models for L2 academic contexts. *Language Teaching Research, 27*(5), 1193-1216. https://doi.org/10.1177/1362168820971740

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Hoboken: Prentice Hall.

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

Chen, Q., Thwaite, A., & Moon, B. (2023). Chinese-speaking undergraduates in Australia: A lexical approach to teaching academic writing. *The Journal of Asia TEFL*, *20*(2), 375–396. https://doi.org/10.18823/asiatefl.2023.20.2.9.375

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research.* Thousand Oaks: Sage Publications.

Dong, L., Ji, T., & Zhang, J. (2023). Motivational understanding of MOOC learning: The impacts of technology fit and subjective norms. *Behavioral Sciences*, *13*(2), 98. https://doi.org/10.3390/bs13020098

Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. London: Routledge.

Granado-Peinado, M., Mateos, M., & Martín, E. (2019). Teaching to write collaborative argumentative syntheses in higher education. *Reading and Writing*, *32*, 2037–2058. https://doi.org/10.1007/s11145-019-09939-6

Huang, H.-W., Teng, D. C.-E., & Wu, R. (2023). Teaching chatbot prompt strategies in EFL essay writing instruction. *In 2023 5th International Workshop on Artificial Intelligence and Education (WAIE)* (pp. 12–17). IEEE. https://doi.org/10.1109/WAIE60568.2023.00010

O'Dowd, R., Sauro, S., & Spector-Cohen, E. (2020). The role of pedagogical mentoring in virtual exchange. *TESOL Quarterly*, 54, 146–172. https://doi.org/10.1002/tesq.543

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, 16(4), 385–407. https://doi.org/10.1007/s10648-004-0006-x

Seker, M., & Karagul, B. I. (2022). Assisting higher education learners online to acquire self-regulated writing strategies during Covid-19. *Journal of Learning and Teaching in Digital Age*, *7*(1), 64–75. https://doi.org/10.53850/joltida.989005

Shen, B., & Bai, B. (2024). Enhancing Chinese university students' writing performance and self-regulated learning (SRL) writing strategy use through a strategy-based intervention. *System*, 122, Article 103249. https://doi.org/10.1016/j.system.2024.103249

Stell, A., & Iwashita, N. (2024). Enhancing collaboration: Exploring regulated learning strategies in the co-regulatory processes of collaborative L2 writing. System, 125, 103410. https://doi.org/10.1016/j.system.2024.103410

Teo, P. (2017). Bridging policy and practice: A study of EFL teacher talk in China. *TESL-EJ*, 21(3). https://tesl-ej.org/wordpress/issues/volume21/ej83/ej83a1/

Teng, L. S. (2022). Explicit strategy-based instruction in L2 writing contexts: A perspective of self-regulated learning and formative assessment. *Assessing Writing*, 53, 100645. https://doi.org/10.1016/j.asw.2022.100645

Teng, L. S., & Zhang, L. J. (2016). A questionnaire-based validation of multidimensional models of self-regulated learning strategies. *Modern Language Journa*l, 100(3), 674–701. https://doi.org/10.1111/modl.12339

Teng, L. S., & Zhang, L. J. (2020). Empowering learners in the second/foreign language classroom: Can self-regulated learning strategies-based writing instruction make a

difference? *Journal of Second Language Writing*, 48, 100701. https://doi.org/10.1016/j.jslw.2019.100701

Teng, L. S., Wang, C., & Zhang, L. J. (2022). Assessing self-regulatory writing strategies and their predictive effects on young EFL learners' writing performance. *Assessing Writing*, 51, 100620. https://doi.org/10.1016/j.asw.2021.100573

Teng, M. F., & Huang, J. (2021). The effects of incorporating metacognitive strategies instruction into collaborative writing on writing complexity, accuracy, and fluency. *Asia Pacific Journal of Education*, *43*(4), 1071–1090. https://doi.org/10.1080/02188791.2021.1982675

Umamah, A. & Cahyono, B. Y. (2020). Indonesian university students' self-regulated writing (SRW) strategies in writing expository essays. *Indonesian Journal of Applied Linguistics, 10*(1), 25–35. https://doi.org/10.17509/ijal.v10i1.24958

Vasu, K. A., Fung, Y. M., Nimehchisalem, V., & Md Rashid, S. (2020). Self-regulated learning development in undergraduate ESL writing classrooms: Teacher feedback versus self-assessment. *RELC Journal*, *53*(3), 491–508. https://doi.org/10.1177/0033688220957782

Wang, Z. (2023). Applying self and peer assessment in Chinese tertiary education English classroom. *International Journal of Social Science and Education Research*, 6(7), 127–133. https://doi.org/10.6918/IJOSSER.202307_6(7).0018

Zhang, J., & Zhang, L. (2025). Exploring the profiles of foreign language learners' writing self-regulation: Focusing on individual differences. *Reading and Writing*, 38, 1533–1559. https://doi.org/10.1007/s11145-024-10568-x

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Cambridge: Academic Press.

Zimmerman, B. J., & Schunk, D. H. (Eds.). (2011). *Handbook of self-regulation of learning and performance* (2nd ed.). London: Routledge.

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.359 -372
DOI: https://dx.doi.org/10.24093/awej/vol16no3.22

Enhancing L2 Writing Pedagogy Through AntWordProfiler and Flesch Reading Score Analyses of Vocabulary and Readability

Princess Gerbie C. Durante
College of Professional Teacher Education
Bulacan State University, Philippines
Email: pgerbie.durante@bulsu.edu.ph

Received: 06/06/2025 Accepted:09/12/2025 Published: 09/20/2025

Abstract
Vocabulary level illustrates the depth and breadth of lexical knowledge and competence, while readability refers to how accessible the written text is to a larger and more diverse audience. This cross-sectional descriptive-correlation study aimed to identity the vocabulary level and readability of students written output and their relationship, which both have implications to writing pedagogy. Characteristics of 35 participants' vocabulary level were analyzed using AntWordProfiler, while the readability of their written output was analyzed using Flesch Reading Ease Score. Results showed that the vocabulary level of all students involved in the study is within the elementary level or Level 1. However, there are also characteristics of Level 2, 3, and 5 observed. Meanwhile, the readability of majority of the participants' written output are aligned with their current grade level – college. Correlation analysis between vocabulary level and readability rendered strong positive, significant correlation between Level 1 and readability, and in contrast, strong negative, significant correlation between Level 3 and readability. Implications were drawn from the results for the development of a proposed Integrated Lexical-Discourse Approach to Writing that may serve as guiding principles for writing instruction.

Keywords: *AntWordProfiler, discourse competence, English as second language, Flesch Reading Score, lexical competence, vocabulary competence, writing pedagogy*
How to cite: Durante, P. G.C. (2025). Enhancing L2 Writing Pedagogy Through AntWordProfiler and Flesch Reading Score Analyses of Vocabulary and Readability *Arab World English Journal, 16*(3): 359 -372. https://dx.doi.org/10.24093/awej/vol16no3.22

Introduction

To effectively convey thoughts and ideas in any language, a firm grasp of the concepts and meaning of the words in a language is necessary. The ability to meaningfully and accurately represent one's ideas involves vocabulary competence. It is particularly crucial in an academic setting where communication is essential to teaching and learning. In the English language, for instance, vocabulary competence enables learners to participate in educational discourse and discussions of their interests to become involved in the school life (Sarip et. Al, 2021). Further, published literature on vocabulary competence highlights its significance on learners' capacity to communicate clearly, coherently, and profoundly both orally and in writing (Li & Zhang, 2021; Li et al., 2025).

In the Philippines, where English is a second language (L2), it is used as the medium of instruction in most primary education courses until university. Thus, the ability to use English lexical items proficiently highly influences the capacity of students to create and produce scholarly and creative writings, participate in academic and social conversations, and deliver and engage in instruction. However, despite early exposure to the English language, many Filipino students struggle to expand their vocabulary. Various local studies have reported that many Filipino students' vocabulary levels are still within the basic and intermediate levels. For instance, a study on the macro skills of English language teachers and pre-service teachers revealed intermediate proficiency in listening, reading, and writing, emphasizing the significantly limited capacity to use appropriate vocabulary in writing (Punongbayan, 2019; Raymundo, 2023). Similarly, another study in an EFL context noted that students cannot use advanced vocabulary (Sari et al., 2022). These results illustrate the difficulty of non-native English speakers using English vocabulary in communication and academic writing effectively (Strochenko et al., 2025). This difficulty is compounded by the lack of access to reading materials, digital content, and a language-rich environment, which play an important role in vocabulary development, especially among marginalized students or those from very low-income families (Aton et al., 2024; Solitana & Generoso, 2022).

Moreover, literature has shown varying efficacy of vocabulary learning strategies (VLS). For example, a local study in the Philippines has illustrated no significant correlation between VLS and vocabulary skills (Solitana & Generoso, 2022). Meanwhile, metacognitive strategies, particularly belief-based ones, have positively influenced vocabulary acquisition (Enobio & Palma, 2025; Rodriguez, 2023; Septiani et al., 2024; Ueno & Takeuchi, 2023). Such findings suggest that more personalized or individualized teaching strategies are needed to bridge the gap in vocabulary learning.

Against the limitations presented, technology-based applications such as AntWordProfiler have proved helpful in evaluating and strengthening vocabulary competence. AntWordProfiler is a free, corpus-based tool that classifies vocabulary in student writing according to frequency bands, helping educators to assess the sophistication and coverage of students' lexical choices quickly and effectively (Anthony, 2024). The application breaks down words in a written text and categorizes them as Level 1 for elementary, Level 2 for intermediate, Level 3 for advanced, and Level 5 for rare or uncommon words. Various studies (Sari et al., 2022) have used AntWordProfiler to conduct in-depth lexical analysis of large text samples. More particularly, one study has shown that students' vocabulary level may be categorized under the elementary and intermediate, and only barely advanced (Sari et al., 2022). However, in the Philippine setting, such applications are barely utilized. Although applications such as the AntWordProfiler are freely accessible and very

useful in providing information about the students' vocabulary level and use of lexical items, there is a dearth of literature on their integration in teaching practices among Filipino teachers, particularly in teacher education courses. Thus, more research may be conducted to investigate how tools such as AntWordProfiler can be used in instructional settings to resolve vocabulary competence issues and inform vocabulary instruction.

Likewise, determining how these tools can also be used to enhance writing competence, especially in terms of readability, is necessary. Readability refers to how easily a text can be read (Anindya & Sutrisno, 2019) by its intended audience. Readability factors include sentence length, vocabulary usage, syntactic complexity, and cohesion (Avila et al., 2021). It is an essential aspect of writing, especially among pre-service English teachers who will teach the English language as an L2. Moreover, future language teachers of English need to model good vocabulary competence and writing proficiency as they will serve as a rich language resource for their students. Unfortunately, some studies reveal the limited direct instruction afforded pre-service teachers to help them measure and improve the readability of their writing (Grose et al., 2021). Although there is evidence of much instruction and challenges in grammar, paragraph development, and writing conventions for various genres (Osman et al., 2025), there is limited literature on instruction that enables pre-service English language teachers to improve the readability of their written texts. This limitation is critical as pre-service teachers are generally expected to write lesson plans, design learning materials, and develop instructional aids, as well as conduct research and reports on students' performance in the classroom.

Thus, when pre-service teachers write texts with limited readability, it can limit students' understanding, especially given that the Filipino learners are L2 speakers of English. Applications that measure the readability of text based on sentence length and the complexity of words assist writers in enhancing the readability of their writing. Microsoft Word, for instance, offers free document statistics that analyze the readability of text through the Flesch-Kincaid Grade Level. Text readability can be improved using simple and less complex words, using the active voice rather than the passive voice, using less complex and compound sentences, and using illustrations, graphs, and other visual aids to clarify information presented in the text. The identified strategies can ensure that the written text is linguistically correct and accessible to a broader set of students.

In the Philippines, however, studies on writing competence mostly centered on grammar, mechanics, and organization (Manla et al., 2024; Tangan et al., 2023), and less on the interaction between vocabulary competence and readability, especially among pre-service English language teachers. Hence, this study addresses this gap by investigating how these two important facets of writing ability are related. In particular, the study seeks to accomplish four goals: determine the vocabulary level of student participants; assess the readability of their written work; examine the relationship between vocabulary competence and the readability of written work; and offer directions for using the findings to inform the instructional design towards enhancing both vocabulary competence and writing ability in English as an L2.

Specifically, this study sought to answer the following research questions:

1. What is the vocabulary level of the pre-service English language teacher-participants?
2. What is the readability level of their written works?
3. Is there a relationship between their vocabulary level and the readability of their written works?
4. How may the results of the study be used in informing the instructional design towards enhancing both vocabulary level and writing ability in English as a second language?

By addressing these questions, this research aimed to create more efficient instruction methods and tools for improving vocabulary acquisition and writing skills in the Philippine educational setting. The results will provide important implications for understanding how vocabulary and readability testing may inform instructional practices and eventually lead to the academic achievement of Filipino students, especially those enrolled in teacher education programs.

Methodology

Research Design

The descriptive-correlation design is the most appropriate research design for the study. Specifically, a cross-sectional design was implemented to identify the vocabulary level and readability of text produced by pre-service English teacher participants. Cross-sectional design is the most useful for establishing preliminary evidence for a more advanced study in the future (Wang & Cheng, 2020), while correlation design was used to explore whether there is significant correlation between the vocabulary level and readability of written works of the participants.

Sample and Locale

The participants were selected using purposive sampling. The following criteria were taken into consideration in identifying the participants qualified to participate in the study: (1) enrolled as freshmen in the teacher education program for academic year 2024-2025 at a state university in Central Luzon, Philippines, (2) part of the group taking up the foundations of inclusive and special education course where they accomplished writing tasks, (3) did not use any artificial intelligence tool in their writing assignment. Thus, initially, there were 52 participants in the study who met the first and second criteria. However, after subjecting the writing assignments to an AI detector, 32.69 percent, or 17 of the 52 participants, were identified to have used artificial intelligence to write their writing assignment, so they were removed from the list of participants. Hence, only 35 participants were included in the study sample.

Instruments and Procedure

The data was collected using a Writing Task Assignment with precise instructions asking students to provide their insights or reflections after attending a symposium about Students with Learning Disabilities and Students who are Blind and have Hearing Impairments. The writing task was available through Microsoft Forms, and participants were given one week to finalize their writing outputs.

After the participants submitted their writing outputs, the texts were subjected to AI detection using the Quillbot application, as independent tests done using Quillbot showed 100 percent accuracy in detecting human and AI-produced text (Gewirtz, 2025). Of the 52 written outputs subjected to AI detection, 17 were tagged as AI-generated, ranging from 32 percent to 100 percent. These 17 outputs were excluded from the sample list. Afterwards, the remaining 35 texts identified as entirely written by the human participants were uploaded to the AntWordProfiler and subjected to vocabulary level analysis. The level and tokens for each text were carefully recorded and organized in a Google Sheet file. Later, the same 35 texts were transferred to a Word Document file to undergo Document Statistics identification. The Flesch Reading Ease and the Flesch Kincaid Grade level were recorded for each text and were interpreted as shown in Table 1.

Table 1. *Flesch Reading Ease Scores with Education Level and Readability Rating by Ferguson et al. (2021)*

Flesch Reading Ease Score	Education Level	Readability Rating
0-29	College graduate	Very difficult to read
30-49	College	Difficult to read
50-59	10th-12th grade	Fairly difficult to read
60-69	8th-9th grade	Average
70-79	7th grade	Fairly average
80-89	6th grade	Easy
90-100	5th grade	Very easy

The Flesch Reading Ease Score developed by Rudolph Flesch in the 1940s calculates the readability of a text based on the length of the sentence in terms of number of words and the average length of a word in terms of number of syllables. The results range from 0 to 100, with the low score indicating greater reading difficulty (Ferguson et al., 2021).

Data Analysis

Percentages and frequency counts were used to identify the vocabulary level of participants and the readability of their written output. In contrast, Pearson's r was used after checking the normality and homogeneity of the quantitative data to identify whether there is a significant relationship between vocabulary level and readability of written output.

Ethical Consideration

Participation was voluntary, and students were informed that they may withdraw their consent to be part of the study at any point during the study's conduct.

Results

The succeeding section outlines the results of the study.

Vocabulary Level

The results of the AntWordProfiler analysis indicate that the vocabulary level of all 35 participants mostly falls under Level 1 or elementary, as shown in Figure 1.

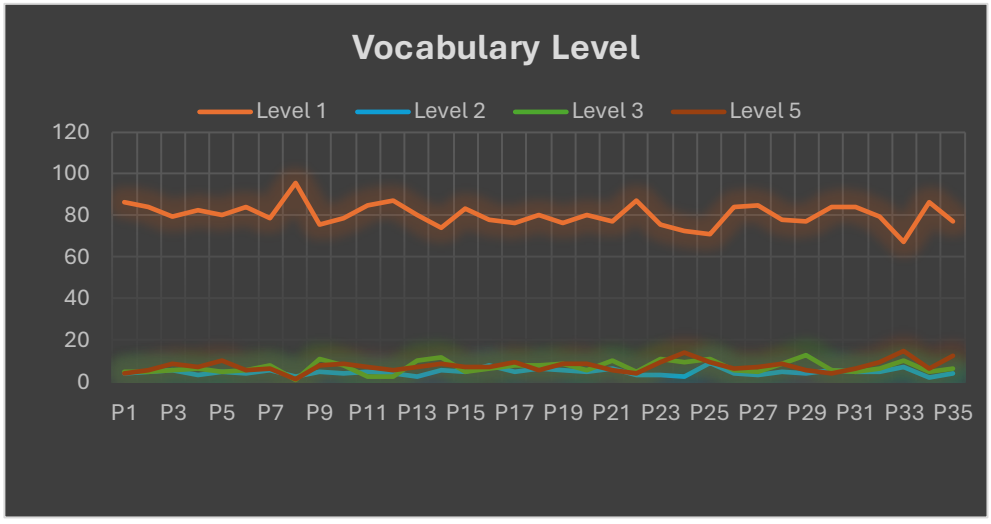


Figure 1. Vocabulary level analysis using AntWordProfiler

The orange line representing Level 1 illustrates that most lexical items participants use are simple words classified within elementary-level vocabulary. They are easy to understand and accessible to the general population. Level 2 (blue lines) and Level 3 (green line) vocabulary are not prevalently used, although there is a sprinkling of uncommon or rare words which are considered as Level 5 (maroon line) vocabulary. For instance, participant 33 has the highest vocabulary competence with a Level 5 vocabulary at 14.95 percent, Level 3 vocabulary of 10.29 percent, Level 2 vocabulary of 7.38 percent, and Level 1 vocabulary of only 67.38 percent. Meanwhile, participant 8 has the least sophisticated vocabulary: Level 1 at 95.71 percent, Level 2 at 2.41 percent, Level 3 at 1.03 percent, and Level 5 at 1.38 percent.

The following section discusses the readability of participants' written output.

Readability of Written Output

The Flesch Reading Ease analysis illustrates that 51.43 percent, or 18 participants, have a written output readability aligned with their current grade level – college, as shown in Table 2.

Table 2. *Flesch Reading Ease and Flesch Kincaid Grade Level of Written Outputs*

Flesch Reading Ease Score	Percentage distribution *(n)*	Education Level	Readability Rating
0-29	31.43 % *(11)*	College graduate	Very difficult to read
30-49	51.43% *(18)*	College	Difficult to read
50-59	11.43% *(4)*	10th-12th grade	Fairly difficult to read
60-69	2.86 % *(1)*	8th-9th grade	Average
70-79	2.86 % *(1)*	7th grade	Fairly average
80-89	0	6th grade	Easy
90-100	0	5th grade	Very easy

However, 31.43 percent or 11 participants have a Flesch Reading Ease score within 0 – 29, or very difficult to read, as the readability is analyzed to be geared more towards college graduates. Meanwhile, 17.15 percent, or 6 of the 35 participants, have a written output with a readability score below their grade level.

More specifically, the written output from participant 21 scored the most difficult readability rating of 13.4, with a grade level of 22.3, or a college graduate.

Relationship between Vocabulary Level and Readability

The succeeding discussion illustrates whether there exists a significant relationship between vocabulary level and readability of written output, as shown in Table 3.

Table 3. *Pearson r analysis results between Vocabulary Level and Readability*

	Level 1	Level 2	Level 3
FRE			
Pearson Correlation	.776**	-.445**	-.740**
Sig. (2-tailed)	.000	.007	.000
Sum of Squares and Cross-products	1988.589	-316.498	-996.999
Covariance	58.488	-9.309	-29.323
N	35	35	35

*Note: FRE = Flesch Reading Ease; **Correlation is significant at the 0.01 level (2-tailed)*

The result of the analysis between FRE and Vocabulary Level 1 shows a Pearson *r = .776*, with a p-value of *.000*. This means a strong positive correlation between FRE and Level 1 vocabulary competence. The result implies that participants who use elementary lexical items are associated with easier-to-read texts.

Meanwhile, FRE and Vocabulary Level 2 have a moderate negative correlation at Pearson *r = -.445*, statistically significant at *p = .007*. The results imply that as Level 2 lexical items in a text increase, their readability decreases, making reading more challenging.

Finally, a strong negative correlation exists between FRE and Level 3 vocabulary competence at Pearson *r = -.740*, with a p-value of *.000*, which is statistically significant. These findings suggest that the more specialized vocabulary used in a text makes reading much more difficult.

Discussion

The most striking difference between the lexical items used by participants is the use of technical jargon relevant to the writing assignment topic. The word cloud of participant 8 is focused on preparing teachers and providing educational support, with prominent words such as *knowledge, need, help, able, teachers, students, understand, future*, and *special needs.* The tone conveys aspiration, being proactive, and goal-oriented. The themes that emerge focus on the desire to gain knowledge to help students with special needs, focus on teacher's roles and their readiness, future-proofing teaching practices, and high emphasis on action words such as *apply, gain, adjust,* and *improve.*

The word cloud in Figure 2 shows the different lexical items used by the two participants.

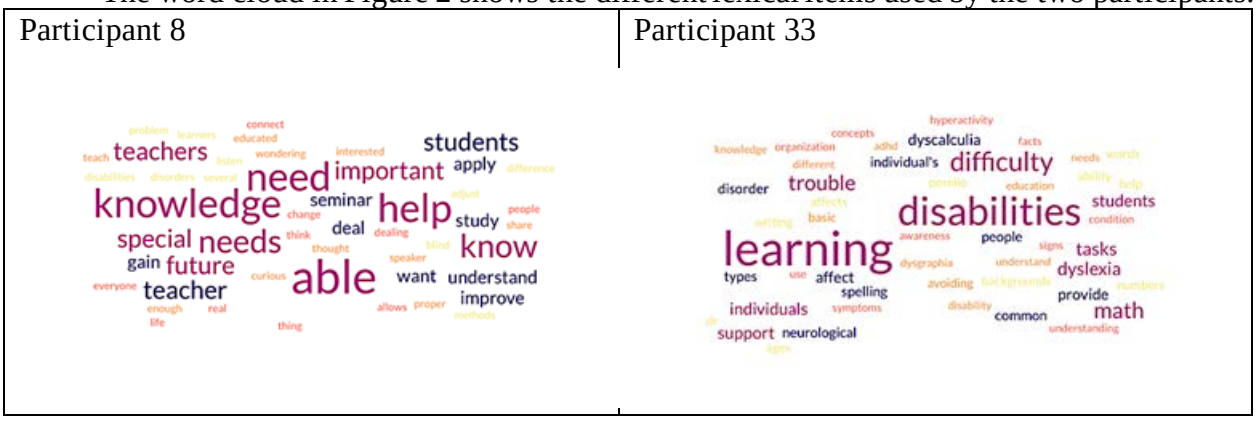


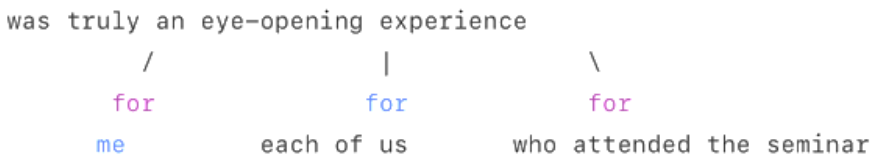
Figure 2. Word cloud comparison of lexical items used by participants

On the other hand, participant 33 is focused on understanding learning disabilities, using words such as *learning, disabilities, difficulty, trouble, dyslexia, math, support, tasks,* and *individuals* more prominently. The tone is descriptive, diagnostic, informative, and more formal than participant 8's. The themes emerging reflect a clinical or academic approach, more specifically by defining and characterizing learning disabilities, discussing specific conditions such as dyslexia, dysgraphia, and dyscalculia, iterating the impact of disabilities on learning tasks, and using lexical items that center on challenges faced by individuals. These results are similar to previous studies that students predominantly use basic vocabulary, and with limited use of advanced-level words (Sari et al., 2022), and that advanced-level vocabulary is primarily specialized words, distinct from general lexical word families (Haruna et al., 2024; Rhubido et al., 2023; Wasikarat & Soontornwipast, 2024).

The results imply that using specialized vocabulary and technical jargon specific to a particular discipline signals higher vocabulary competence. Meanwhile, using general lexical items in everyday conversation signals a lower vocabulary level. In teaching and learning, this meant that the more specialized the courses are, the higher the requirement for vocabulary competence. Acquiring specialized vocabulary may also signal acquiring content knowledge necessary to master the discipline. Likewise, the use of more general lexical items may be attributed to either lower grade levels where specialization is not yet emphasized or general education courses where students may come from varying disciplines, so the language, tone, and concepts must be accessible to a larger and more diverse audience.

Meanwhile, the sentence structure of the written output identified to be the most difficult to read illustrates the complexity of the sentence structure, as shown in Figure 3. The structure of the written output rated very difficult to read, which is characterized by several complex and compound sentences, with embedded clauses and prepositional phrases, as shown in the word tree structure analysis of sample 1. The root phrase *"was truly an eye-opening experience"* is compounded with multiple prepositional phrases and dependent clauses, showing that the experience is not limited to the writer but also to multiple participants. Moreover, the second sample from the same written output illustrates two levels of branching from the root phrase. The first branching level has two main emphases – *early diagnosis* and *constant reminder*s. In contrast, the second branching level expands into the *importance of IEPs* and ultimately iterates to whom it was aimed, the "*learners*." The written output is characterized by noun phrase coordination, packing two or more main ideas, embedded clauses that contain general truths and reinforce points, and complex modifiers that add depth and precision.

Sample 1



Sample 2

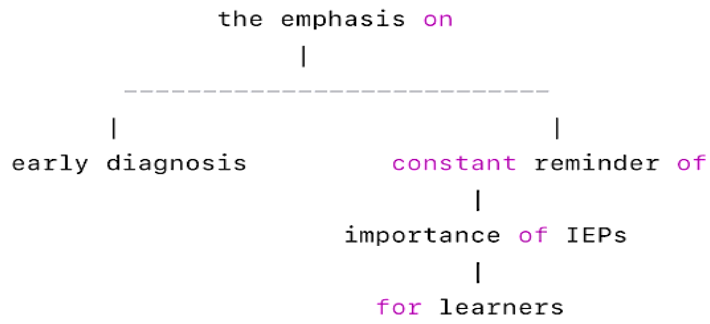


Figure 3. Sentence tree structure analysis with very difficult to read readability rating

On the other hand, the sample from the written output rated fairly average readability focuses mainly on a main verb, a core action that highlights a single idea, uses adverbs, parallel structure, and noun complements, characteristics of the basic sentence pattern, as shown in Figure Four.

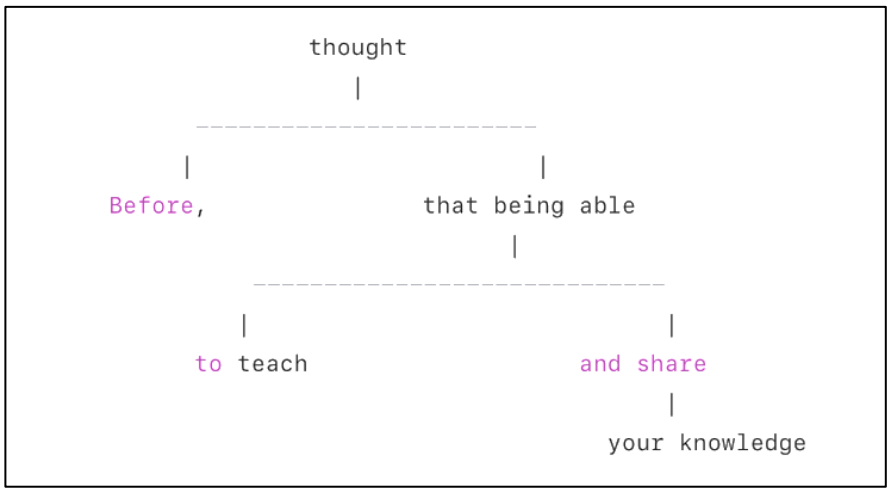


Figure 4. Sentence tree structure analysis with fairly average readability rating

The results of the analysis of sentence structure of the most difficult to read text and the relatively average readability affirm literature that increased sentence complexity – characterized by long sentences, embedded clauses, and advanced vocabulary – tends to lower readability of written output as measured by Flesh Reading Ease and Flesch-Kincaid Grade Level (Eleyan et al., 2020; Gurmani et al., 2025; Li et al., 2025; Mishra & Dexter, 2020; Sakhuja et al., 2021).

Finally, the correlation analysis illustrates a clear pattern: lexical items categorized under fundamental or generalized lexical families contribute to higher reliability. At the same time, more advanced or specialized vocabulary leads to increased text complexity. The implications for writing pedagogy are that the appropriate use of vocabulary-leveled words is necessary to align with the target readers' proficiency and thus improve comprehensibility and accessibility of the written output. These results are similar to the results of the studies by Elvriza and Nurcholis (2024), Kongcharoen et al. (2024), Lașcu (2022), and Quines (2023).

Implications of Vocabulary Level and Readability to Writing Pedagogy

Vocabulary level and readability play a significant role in developing students' writing proficiency through effective writing pedagogy. The study's results, as affirmed by recent literature, highlight the need for integrating lexical competencies with discourse competence to enhance writing instruction. The succeeding section presents a comprehensive discussion of implications and the relationship between vocabulary level, readability, and writing pedagogy towards acquiring writing competence, culminating in a proposed pedagogical model for teaching writing.

To effectively integrate vocabulary development and readability in writing instruction, the following components are proposed: Lexical Enrichment, Semantic Mapping, Discourse structuring, Integrated Writing Practice, and Reflective Feedback, as shown in Table 4.

Table 4. *Proposed Integrated Lexical-Discourse Approach to Writing (ILAW)*

Component	Objective	Strategies
Lexical Enrichment	Expand the breadth and depth of student vocabulary	⇒ Use thematic vocabulary instruction aligned with writing topics ⇒ Integrate the unlocking of difficulty practice as a pre-reading activity ⇒ Encourage extensive reading to expose students to diverse lexical terms by including diverse authors and texts in the syllabus
Semantic Mapping	Enhance understanding of word meanings and relationships	⇒ Explore the use of semantics maps to strengthen word association skills and the acquisition of specialized vocabulary. ⇒ Provide activities that use the new vocabulary encountered in context
Discourse Structuring	Organize ideas coherently and cohesively	⇒ Explicit teaching of text structures ⇒ Exposure to and analysis of elements of various text structures ⇒ Focus on coherence and cohesion in paragraph development activities
Integrated Writing Practice	Apply lexical and discourse skills in writing tasks	⇒ Design writing tasks requiring the use of specialized vocabulary learned within a specified format ⇒ Provide feedback on the initial draft and check for the integration of comments in the revisions
Reflective Feedback	Foster self-awareness and metacognition of personal writing style	⇒ Use peer review sessions to critique students' writing output ⇒ Design personalized writing tasks that allow the students to reflect on their writing practices

The principles proposed in ILAW may contribute to improving students' writing competence and deepening and broadening their vocabulary. Using ILAW as a guiding principle can equip learners with the tools to produce coherent, cohesive, and contextually appropriate written outputs that best convey their ideas, thoughts, and belief systems. More significantly, increased writing competence and deepened and broader vocabulary may empower learners to use the English language not just as tool for communication but more so as an instrument to develop critical consciousness which will help them "improve their own life conditions and take needed actions" to contribute to nation building and social justice (Durante, 2025, 214).

Conclusion and Recommendations

The study investigates the vocabulary level, readability of written works, and the correlation thereof of 35 first-year L2 language-specialization students enrolled at a state-operated teacher education college in the Philippines. The results show that the students primarily have Level 1 or elementary vocabulary competence with limited Level 2, 3, and 5 vocabulary competence; meanwhile, in terms of written output readability, most students, or 51.43 percent, have a writing competence equal to their grade level, while a few have outputs beyond their grade level. Unfortunately, the written output of 17.15 percent, or 6 of the students, was classified below their grade level. Finally, in terms of correlation, a strong positive correlation is evident between Level 1 vocabulary competence and readability of text, in contrast with the strong negative

correlation between the presence of Level 3 vocabulary competence and text readability. These findings have significant implications for writing pedagogy.

Thus, ILAW or the Integrated Lexical-Discourse Approach to Writing is proposed based on the results collected. It is comprised of five components, which serve as guiding principles that L2 language educators and pre-service L2 language teachers alike may subscribe to or use in the classroom in delivering writing instruction. The five components include lexical enrichment, semantic mapping, discourse structuring, integrated writing practice, and reflective feedback. These components iterate specific objectives and strategies that may be implemented in the classroom. It is also recommended to future researchers that the implementation of ILAW be documented as an intervention practice to test its efficacy by including baseline data and test results after the intervention.

Acknowledgement: I would like to give our heartfelt gratitude and great appreciation to all the students who willingly participated in this study. Without their participation, this work would not have been completed.

Funding: This research is not funded.
Conflicts of Interest: The author declares no conflict of interest.
Authenticity: This manuscript is an original work
Artificial Intelligence Statement: This document partially used AI (Grammarly) for readability and grammar. All ideas, arguments, analysis and conclusions are the work of the author. The author reviewed and edited the content and takes full responsibility for the content of the paper.

About the Author:
Princess Gerbie C. Durante works in the College of Professional Teacher Education, Bulacan State University as Language and Professional Education and Research Assistant Professor III. Her research interests lie English language instruction and Critical Language Pedagogy. ORCID ID: https://orcid.org/0000-0002-6692-4404

References

Anindya, W. D., & Sutrisno, A. (2019). Accuracy and readability on the translations of onomatopoeia in *The Maze Runner* novel into Indonesian language. In *Proceedings of the English Language and Literature International Conference (ELLiC), 3*, 374–379. Universitas Muhammadiyah Semarang. https://jurnal.unimus.ac.id/index.php/ELLIC/article/view/4735

Anthony, L. (2024). *AntWordProfiler* (Version 2.2.1) [Computer software]. Tokyo, Japan: Waseda University. https://www.laurenceanthony.net/software/AntWordProfiler

Aton, B. M., Canumay, J. L. C., Lapasaran, D. G., Pascual, Y. G. C., Segundino, J. L., & Moral, R. V. (2024). Vocabulary proficiency and language skills development of English majors: Various factors and significant impacts. *Texas Journal of Philology, Culture and History, 27*, 11–22. https://doi.org/10.62480/tjpch.2024.vol27.pp11-22

Avila, E. C., Lavadia, M. K. S., Sagun, R. D., & Miraña, A. E. (2021). Readability analysis of college students' written outputs using Grammarly Premium and Flesch-Kincaid tools. In *Virtual Conference on Engineering, Science and Technology (ViCEST) 2020* (*Vol.*

1933, Article 012120). *Journal of Physics: Conference Series*. IOP Publishing. https://doi.org/10.1088/1742-6596/1933/1/012120

Durante, P. G. C. (2025). Advancing Critical Literacy Through Critical Literature Pedagogy-Based Principles Of Literary Education. *Problems of Education in the 21st Century*, *83*(2), 196-217. https://doi.org/10.33225/pec/25.83.196

Eleyan, D., Othman, A., & Eleyan, A. (2020). Enhancing software comments readability using Flesch Reading Ease score. *Information, 11*(9), 430. https://doi.org/10.3390/info11090430

Elvriza, M. S., & Nurcholis, I. A. (2024). The impact of extensive reading of novels on students' English vocabulary acquisition. *Journal of Language Teaching and Research, 15*(1), 94–102. https://doi.org/10.71417/ije.v1i3.215

Enobio, I. I., & Palma, R. C. (2025). Exploring the impact of vocabulary strategies, language exposure, and beliefs on students' oral skills: A structural equation modeling approach. *Asian Journal of Language, Literature and Culture Studies*, *8*(1), 1–14. https://doi.org/10.9734/ajl2c/2025/v8i1214

Ferguson, C., Merga, M., & Winn, S. (2021). Communications in the time of a pandemic: The readability of documents for public consumption. *Australian and New Zealand Journal of Public Health, 45*(2), 116–121. https://doi.org/10.1111/1753-6405.13066

Gewirtz, D. (2025, July 14). *I found 5 AI content detectors that can correctly identify AI text 100% of the time*. *ZDNET*. https://www.zdnet.com/article/i-found-5-ai-content-detectors-that-can-correctly-identify-ai-text-100-of-the-time/

Grose, E. M., Holmes, C. P., Aravinthan, K. A., Wu, V., & Lee, J. M. (2021). Readability and quality assessment of internet-based patient education materials related to nasal septoplasty. *Journal of Otolaryngology - Head & Neck Surgery, 50*(1). https://doi.org/10.1186/s40463-021-00507-z

Gurmani, T., Abbasi, I. A., & Jatoi, Z. A. (2025). Readability and syntactic complexity in undergraduate essays on technology. *Journal of Arts and Linguistics Studies, 3*(1), 219–242. https://doi.org/10.71281/jals.v3i1.219

Haruna, H. H., Tajuddin, A. J. A., & Shamsudin, C. M. (2024). Unlocking academic vocabulary: Corpus insights from open and distance English language learning coursebooks. *3L: Language, Linguistics, Literature – The Southeast Asian Journal of English Language Studies*, *30*(4), 291–307. https://doi.org/10.17576/3L-2024-3004-20

Kongcharoen, P., Dhanarattigannon, J., & Thongnuan, T. (2024). Lexical level in English major EFL students' writing: A learner corpus study. *PASAA Journal*, *68*, 203–228. DOI: 10.58837/CHULA.PASAA.68.1.6

Lașcu, T. (2022). Development of the discourse competence through literary texts. *Journal of Language and Education*, *8*(4), 45–56. https://doi.org/10.1556/063.2022.00137

Li, C., Wang, X., & Qian, L. (2025). Exploring syntactic complexity and text readability in an ELT textbook series for Chinese English majors. *SAGE Open, 15*(1). https://doi.org/10.1177/21582440251323619

Li, X., & Zhang, H. (2021). Developmental features of lexical richness in English writings by Chinese L3 beginner learners. *Frontiers in Psychology, 12*, 752950. https://doi.org/10.3389/fpsyg.2021.752950

Manla, E. C., Esparrago-Kalidas, A. J., Napone, A. V., Delgado, N. D. P., Develos, A. M., & Castillo, D. L. (2024). Enhancing college students' English writing skills using POWER-

S & F approach. *Malaysian Online Journal of Educational Sciences, 12*(4), 1–15. https://doi.org/10.22452/mojes.vol12no4.2

Mishra, V., & Dexter, J. P. (2020). Comparison of readability of official public health information about COVID-19 on websites of international agencies and the governments of 15 countries. *JAMA Network Open, 3*(8), e2018033. https://doi.org/10.1001/jamanetworkopen.2020.18033

Osman, O.H., Al Rahbi, A. H., Binayan , C., & Abuoaf, M. (2025). EFL Students' Academic Writing Challenges at the Foundation Programme in the University of Nizwa: Challenges and Remedy. *Arab World English Journal*, *16*(1):267-284. https://dx.doi.org/10.24093/awej/vol16no1.16

Punongbayan, E. J. (2019). English proficiency of the pre-service teachers in one state university in the Philippines. *International Journal of Multidisciplinary Research and Development, 6*(6), 24–29. http://www.allsubjectjournal.com

Quines, Z. M. (2023). Impact of students' vocabulary level to their reading and writing performance. *International Journal of English Language and Linguistics Research, 11*(2), 18–32. DOI: 10.37745/ijellr.13/vol11n21832

Raymundo, J. L. (2023). Exploring the practices on macro skills integrated assessment in Philippine higher education context: Basis in designing a language training program. *International Journal of Language Education, 7*(3), 429–446. https://doi.org/10.26858/ijole.v7i3.24252

Rhubido, D., Shodiq, S., & Vidya Asteria, P. (2023). Creating Indonesian academic vocabulary by using the ant word profiler program to academic writing for BIPA learning. *KEMBARA: Jurnal Keilmuan Bahasa, Sastra, Dan Pengajarannya*, *9*(2), 623–637. https://doi.org/10.22219/kembara.v9i2.22097

Rodriguez, E. (2023). Language learning strategies and grammatical competence of the English-major students in Zamboanga del Norte HEIs. *Sprin Journal of Arts, Humanities and Social Sciences, 2*(04), 27–42. https://doi.org/10.55559/sjahss.v2i04.99

Sakhuja, M., Yelton, B., Arent, M. A., Noblet, S., Macauda, M. M., Fedrick, D., & Friedman, D. B. (2021). Words matter: An analysis of the content and readability of COVID-19 information on clinic websites. *Frontiers in Communication, 6*, 738467. https://doi.org/10.3389/fcomm.2021.738467

Sarip, N. A., Khalid, W. N. C. N., Mazelan, N. N., Khalid, N. M., Salim, N., & Hashim, H. (2021). Language learning strategies used in improving writing skills by selected Form 4 ESL learners. *International Journal of Academic Research in Business and Social Sciences, 11*(7), 110–124. https://doi.org/10.6007/ijarbss/v11-i7/10067

Sari, A. T., Basya, D., & Tjendani, E. N. (2022). Using Ant-Word Profiler app to classify vocabulary level for paragraph writing among EFL learners. *English Teaching Journal: A Journal of English Literature, Linguistics, and Education, 10*(2), 144–155. https://doi.org/10.25273/etj.v10i2.14189

Septiani, A. D., Nasrullah, R., & Faridawati, F. (2024). EFL students' metacognitive strategies in learning vocabulary. *KLASIKAL: Journal of Education, Language Teaching and Science, 6*(2), 419–431. https://doi.org/10.52208/klasikal.v6i2.1169

Solitana, J. E. A., & Generoso, J. C. (2022). Relationship between Filipino ESL learners' vocabulary learning strategies and vocabulary ability. *Silliman Journal, 60*(2). https://sillimanjournal.su.edu.ph/index.php/sj/article/view/31

Strochenko, L., Smaglii, V.,Viktoriia, C., & Yukhymets, S. (2025). Academic Writing for Postgraduates Across Disciplines: A Lexico-syntactic Perspective. *Arab World English Journal*, *16*(1):285-297. https://dx.doi.org/10.24093/awej/vol16no1.17

Tangan, A., Fugaban, A.J., Soriano, N. & Auayang, K.M. (2023). Writing Skills of Pre-Service Teachers: Basis for a Proposed Intervention Program. *Journal of Language and Linguistics in Society*, *4*(01), 13–22. https://doi.org/10.55529/jlls.41.13.22

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

Wang, X., & Cheng, Z. (2020). Cross-sectional studies: Strengths, weaknesses, and recommendations. *Chest*, *158*(1S), S65–S71. https://doi.org/10.1016/j.chest.2020.03.012

Wasikarat, P., & Soontornwipast, K. (2024). A corpus-based vocabulary analysis of first-year undergraduate economics textbooks in an international program. *LEARN Journal: Language Education and Acquisition Research Network, 17*(2), 470–486. https://doi.org/10.70730/OQYE8005

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.373-387
DOI: https://dx.doi.org/10.24093/awej/vol16no3.23

'We Divided the Questions Evenly among Ourselves to Answer them Properly': Collaborative Learning in EFL Classroom Status Quo

Nada Bin Dahmash
Department of Humanities, College of Applied Studies
King Saud University, Riyadh, Saudi Arabia
Email: naldahmash@KSU.EDU.SA

Received: 06/11/2025 Accepted:09/13/2026 Published: 09/20/2025

Abstract
Collaborative learning is a pedagogical approach employed by language teachers to enhance language proficiency and improve the learning process. Students may be asked to work collaboratively inside and outside the classroom. However, studies examining the experiences of English as a foreign language (EFL) learners working collaboratively to complete tasks in Saudi Arabia are scarce. Thus, this study aims to answer these research questions: How do students engage in collaborative learning in a vocabulary course? How do students perceive the effects of collaborative learning in a vocabulary course? Data collection involved focus groups and retrospective interviews with fifty-two students majoring in human resource management at a Saudi university. The findings revealed that students performed five primary roles: translating English to Arabic, assisting their peers, searching for information, supervising group performance, sharing ideas and knowledge, and adopting different strategies to answer assigned tasks. It also identified the resources they drew on, the distribution of tasks, how they formed groups and engaged via WhatsApp. Positive views about collaborative learning included that it provided opportunities to practice the English language, facilitated understanding of information, made studying easier, and promoted social skills. The negatives mentioned were an unwillingness to work in a group and overreliance on others. This paper recommends that teachers supervise students' interactions to ensure each member achieves their maximum potential, while advising policymakers to encourage the implementation of collaborative learning in the language classroom.

Keywords: Collaborative learning, English as a foreign language classroom, EFL students, group work, language learning, pedagogy, status quo
How to cite: Bin Dahmash, N. (2025). 'We Divided the Questions Evenly among Ourselves to Answer them Properly': Collaborative Learning in EFL Classroom Status Quo. *Arab World English Journal, 16*(3): 373-387. https://dx.doi.org/10.24093/awej/vol16no3.23

Introduction

Educators advocate collaborative learning pedagogy in higher education, noting its many benefits, such as enhanced student engagement and academic achievement (Qureshi et al., 2023). It also promotes knowledge sharing, improves communication skills, and helps learners form new friendships (Weinberger & Shonfeld, 2020). Furthermore, it develops students' self-evaluative judgement skills and teamwork skills (McKay & Sridharan, 2024). When collaborative learning occurred online, similar benefits were reported. For example, Oyarzun and Martin (2023) systematically reviewed the literature, dealing with collaborative learning online from 2012 to 2021, and observed that it provided students with opportunities to learn more effectively, acquire new skills, socialize with others and boost self-efficacy, confidence and engagement. Additional benefits included shifting task learning to become student-centered, thereby increasing autonomy and developing personalized learning.

Key factors were found to impact collaborative learning interactions, including, according to Gitterman (2008), a so-called 'hidden' or informal group that formed when students attended their first class, and which has its own communication style and interpersonal relationships but is not formally recognized by teachers. Other important factors are student's willingness to work collaboratively (Weinberger & Shonfeld, 2020), social media usage, and social presence in connection with other group members (Qureshi et al., 2023). Moreover, the students' digital literacy influences their online engagement in collaborative learning environments, and the uneven distribution of workload could influence collaborative learning interactions (McKay & Sridharan, 2024). The uneven distribution of workload is mentioned in the literature as 'free riding'. It refers to situations in which students rely on others to complete tasks and do not contribute evenly in groups, as they engage in collaborative learning, which represents a notable challenge negatively affecting collaborative learning interactions (Hei et al., 2015; Le et al., 2018).

What is Collaborative Learning?

Collaborative learning is a pedagogical approach in higher education, whereby knowledge is socially constructed through interactions between students working together performing open-ended tasks (Yang, 2023). Collaborative learning is also defined as "An environment in which students work together to solve problems and finish numerous assignments, not just by sharing knowledge but also by forming social and community partnerships" (Alzubi et al., 2024, p. 28).

According to Yang (2023), collaborative pedagogy usually involves students working in small groups on loosely structured activities, selecting and managing their own groups. Collaborative learning "creates conditions in which students can negotiate the boundaries between the knowledge communities they belong to and the one that the professor belongs to" (Bruffee, 1999, p. 144). An effective task of this type is the 'structured class discussion' (Gitterman, 2008), in which students are divided into small groups of four to six to work on a task and given clear instructions to complete it in a specific timeframe.

In Saudi Arabia, incorporating collaborative learning in the classroom is advocated as a tool to enhance English language learning skills and develop overall language proficiency (Abdelshaheed, 2024; Al-Ahdal & Alharbi, 2021; Aldossary, 2021; Alkhalaf, 2020; Alzubi et al., 2024, 2025). However, further research on collaborative learning in various EFL contexts is needed (Alzubi et al., 2024). Therefore, this study concentrates on collaborative learning in a vocabulary course to understand the perspectives of students at a Saudi university studying EFL. The following questions guide the study:

1. How do students engage in collaborative learning in a vocabulary course?
2. How do students perceive the effects of collaborative learning in a vocabulary course?

Literature Review

Cooperative Learning in EFL Contexts

The term collaborative learning is used interchangeably with cooperative learning (Alzubi et al., 2024, 2025; McWhaw et al., 2003; Yang, 2023). Therefore, research evaluating cooperative learning in EFL settings will also be discussed here.

In Saudi Arabia, the attitudes of students towards cooperative learning when developing English language at university level was examined by Khan et al. (2020) and Al-Malki et al. (2022) using quantitative measures. Khan et al. (2020) investigated the attitudes of elementary-level male EFL students towards using cooperative learning to learn grammar. Data were collected via questionnaires from 65 undergraduate students enrolled in the department as first years. The study found that cooperative learning developed students' interactions, as their comfort level when engaging with their peers was higher than when speaking with the teacher, helping weaker students to develop their motivation and confidence by increasing their opportunities to practice using English. Moreover, Al-Malki et al. (2022) explored the views of EFL Saudi female learners regarding the impact of cooperative learning on their academic and social skills. Data were collected from 100 undergraduate students majoring in computing and IT, and they reported that cooperative learning improved their social skills, enhanced their academic achievement, their leadership qualities, and their decision-making.

In Libya, Al Jawad (2023) examined the effect of cooperative learning on improving EFL university students' speaking ability using a quasi-experimental design. In total, 50 undergraduate students were divided into two groups: experimental and control, and issued a pretest and a posttest. The students in the experimental group scored higher than the control group on the posttest, denoting improved speaking ability.

These quantitative Likert scale (Khan et al., 2020; Al-Malki et al., 2022) and quasi-experimental (Al Jawad, 2023) based studies are recent and were conducted in EFL settings with university-level students whose first language was Arabic.

Collaborative Learning in EFL Contexts

Prior research has been conducted into collaborative learning to develop the English language in an EFL context among university students. For example, Heble and Mehta (2013) explored the effect of collaborative writing on the comprehension of literary terms and writing skills among undergraduate students majoring in English literature in Oman. Twenty-six female students completed questionnaires and interviews, and data were also obtained from individual essays and essays written collaboratively. Almost half the participants preferred writing essays collaboratively to writing essays individually. The students reported that their understanding of literary texts and their writing skills had improved due to collaborative writing, as it encouraged them to reread, edit, review, and share ideas.

In Korea, Chang and Windeatt (2016) investigated the process of developing collaborative learning practices among 47 low-level language undergraduate students attending an online English course over 15 weeks. Data were collected using five tools: reflective journals, questionnaires, written assignments, discussion forums, and emails. The study identified three

primary types of collaboration: reading peers' work, giving feedback, and group engagement. The participants developed their confidence, motivation, and knowledge as a consequence of reading and commenting on their peers' assignments. Additionally, they developed gradual trust and a sense of community through engagement in collaborative activities, and over time, hesitant participants became more willing to offer feedback and share their work.

In Indonesia, Listyani (2019) explored collaborative text analysis of journal articles in an academic writing course. Data collection involved reflective journals written by undergraduate students, direct observation, and interviews with two students. The students were required to work collaboratively in groups of two or three, and analyze journal articles dealing with language teaching, or the technologies associated with language teaching. The study found the majority of students enjoyed engaging in collaborative work, as it allowed them to share ideas, minimize stress, acquire new knowledge, gain a deeper understanding of content, and develop interpersonal skills.

Studies by Heble and Mehta (2013) and Listyani (2019) examined improvements in the writing skills of language students. However, whereas Heble and Mehta (2013) used mixed methods, Listyani (2019) used direct observations and interviewed just two students. In a fully online context, Chang and Windeatt (2016) explored the use of collaborative learning among students who never met their peers or teacher in real life, but did not include interviewing the students. The current study explored the use of collaborative learning during a vocabulary course delivered face to face, using focus group and retrospective interviews.

Collaborative Learning in the EFL Context in Saudi Arabia

Several scholars have previously researched collaborative learning in EFL in university contexts in Saudi Arabia (Abdelshaheed, 2024; Al-Ahdal & Alharbi, 2021; Aldossary, 2021; Alkhalaf, 2020; Alzubi et al., 2024, 2025). These studies are described in detail to help identify the focus of the current study.

Research on collaborative learning's effects on enhancing writing skills among EFL students was carried out by Alkhalaf (2020) and Aldossary (2021). For example, Alkhalaf (2020) examined the attitudes of female students towards collaborative learning to improve their writing skills. The study involved 50 students on an IELTS preparation course. Data were collected via questionnaires, and the students reported several advantages of collaborative writing: (1) planning and structuring essays, (2) enhancing the accuracy of their grammar and vocabulary, (3) developing their critical thinking, and (4) increasing their motivation and enjoyment. In Aldossary's (2021) study involving 46 male students, participants were divided into control and experimental groups and pre-tested and post-tested. The experimental group completed collaborative writing tasks over 8 weeks period and then completed a follow-up questionnaire after the intervention. The study found the experimental group scored better on the post-test and had improved their writing performance. The students in the experimental group also expressed positive attitudes about working collaboratively, having appreciated and enjoyed the opportunities to exchange their ideas and enhance their language comprehension.

Research was carried out by Al-Ahdal and Alharbi (2021) and Abdelshaheed (2024) to understand the effects of collaborative learning on enhancing the vocabulary of EFL students. In their study, Al-Ahdal and Alharbi (2021) examined the effects of collaborative learning with Mobile-Assisted Language Learning (MALL) on vocabulary retention for 160 students studying at two different universities. Their study employed an experimental design, dividing students

equally into control and experimental groups who were given a pre-test and post-test. The intervention with the experimental group involving collaborative learning with MALL lasted three months. The students in the experimental group improved their vocabulary retention and demonstrated greater motivation, engagement, and peer-interaction. Similarly, Abdelshaheed (2024) used an experimental design to explore the effect of digital collaborative learning on vocabulary learning and reading comprehension among language learners. She employed two tests and semi-structured interviews with 60 students on a reading and vocabulary-building course. Students were invited to work collaboratively with peers on structured activities, such as engaging in discussions, using Venn diagrams, and completing vocabulary quizzes online on Blackboard and Quizlet. She found the performance of the group who worked collaboratively outperformed that of the control group. It also emerged that the students valued the use of digital collaborative learning, as it provided them with support and encouragement from their classmates, and the dynamic activities assisted them in identifying their mistakes and learning at their own pace, reducing mistakes in subsequent tasks.

The effects of a collaborative learning program on students' social skills and English language learning at Najran university was examined by Alzubi et al. (2024). Their quasi-experimental design employed an experimental group and a control group and included a pre-test and post-test. They also interviewed the English teachers to acquire their perspectives regarding the effect of the intervention on the students' progress. They found that collaborative learning significantly improved students' English proficiency and social interactions. The teachers valued collaborative learning as it inspired and motivated the students to engage in groups.

More recently, Alzubi et al. (2025) examined the attitude of EFL students at Najran University towards using collaborative learning to enhance their English-speaking skills. The study required 360 students to fill in a closed-ended questionnaire. It determined the students were positive about collaborative learning to improve their speaking skills, reporting satisfaction regarding learning with their peers, enjoyment in identifying their roles during speaking tasks, and gratitude that the collaborative learning process afforded numerous speaking opportunities.

The methods in the studies reviewed were subject to some gaps that are addressed in the current study, which employed qualitative methods, in particular focus groups and retrospective interviews, to explore the features and activities students engage with when working collaboratively, both face-to-face and online. It also aims to evaluate the effects of collaborative learning from the perspective of students.

Methodology

This study utilized qualitative approaches and is a typical method for researching collaborative learning (Yang, 2023). It is also a valuable method in the field of applied linguistics, as it documents the development and beliefs of language learners (Lew et al., 2018). Data collection tools used were focus groups and retrospective interviews.

Focus groups can offer different perspectives about a situation, as they encourage participants to share their views (Brinkmann & Kvale, 2018). Retrospective interviews can help to uncover participants' thoughts and mental processes during a particular task (Dörnyei, 2007). The use of retrospective interviews to research collaborative learning in an EFL context with language students at the university level can effectively elicit vivid descriptions (Aldossary, 2021).

Participants in the focus groups were encouraged to discuss their previous and current experiences working collaboratively inside and outside class (see Appendix A). The participants

in the retrospective interviews were invited to discuss one situation, either a task where they collaborated with their colleagues inside the class or one where the collaboration took place outside the classroom. The participants were given time to visualize the task and then describe the processes involved in completing the task (see Appendix B). Focus groups and retrospective interviews were conducted via Zoom meetings, which assists in the collection of qualitative data (Archibald et al., 2019). Attendees were encouraged to share screenshots, websites, and photos to illustrate their responses.

Research Procedures

Data were collected at the end of November 2024 after the participants had attended 11 lectures for a vocabulary course, in which a collaborative learning pedagogy was implemented. The researcher was the course instructor, and asked the students to work collaboratively in-class for each lecture to perform several tasks related to the topic of the lecture and work collaboratively outside the class to present a topic related to the course concepts. The participants were encouraged to create and form their own collaborative groups comprising five to eight students, and were given a time limit to work on the in-class tasks. Focus group interview was used as the first data collection method, and retrospective interview as the second method. In the focus groups 52 students contributed to eight sessions (see Table1), and 29 students participated in the retrospective interviews.

Table 1. *The distribution of participants in each focus group session.*

Number of Sessions	Number of participants
Focus group 1	7
Focus group 2	6
Focus group 3	7
Focus group 4	6
Focus group 5	7
Focus group 6	7
Focus group 7	6
Focus group 8	6
Total	52

Participants

Purposive sampling was used (see Palys, 2008), in which the researcher invited typical participants with experience that aligned with the focus of the research (Dörnyei, 2007). Fifty-two students agreed to participate, all of whom were studying the course as a component in the fourth level of their Human Resources Management diploma at a Saudi Arabian University. The course is a vocabulary course, in which students study terminologies and concepts related to their major. The students received an information sheet about the study, the research tools and what participation would entail. The researcher explained the voluntary nature of the research, that data would be used for research purposes only, that the names of participants would be anonymized, and audio recordings would be deleted after transcribing the data. These considerations are mandated by the case study university's ethics committee regulations.

Data Analysis

Data from the focus group and retrospective interviews were transcribed in text form and uploaded to ATLAS.ti version 25, which is a computer program designed to assist qualitative data analyses. ATLAS.ti provides tools to select quotations, add codes, and retrieve the coded data.

Data were analyzed using thematic analysis (Braun & Clarke, 2006), and themes were developed from codes to answer the research questions.

Results and Discussion

Research Question (1): How Do Students Engage in Collaborative Learning in a Vocabulary Course?

To provide a clear understanding of how the students engaged with collaborative learning from their own perspective, five themes were created from the data: (1) the roles these students performed in the group, (2) the resources they used, (3) how they divided up and distributed tasks, (4) the approach they followed to form groups, and (5) the WhatsApp group they created (see Figure 1).

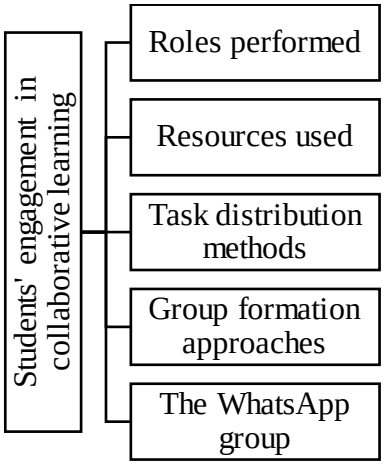


Figure 1: Engagement of the students in collaborative learning

Roles Performed

The students reported having performed different roles, some of which overlapped. These included translating from English to Arabic, assisting their peers, searching for information, supervising their group's performance, sharing ideas and knowledge, and adopting multiple strategies to provide answers for assigned tasks.

Translating from English to Arabic was one of the roles mentioned by the majority of the participants when describing their engagement with their groups. The participants translated both the unknown vocabulary and the questions. Participant 2 clarified that she translated the questions when she started working on tasks with the group. She explained, "Once we have formed the group, the first thing we do is translate questions from English to Arabic to understand them". Participant 26 explained that translating from English to Arabic is essential, as a way to begin any task, either in class or outside of the classroom. Participant 1 stated "my role was to present the functions of management, so I translated 'functions', 'organizing' and other words I did not know". It seems that translating from English to Arabic allowed the participants to function better, as they engaged in collaborative learning with the group. However, they were required only to use English during the course.

The participants reported that assisting their peers was another aspect of group work. The participants read the texts aloud to their peers, analyzed and corrected their answers, and explained information to them once they had understood the tasks. Participant 30 explained that she reads the texts aloud to her peers. She stated, "I read the questions aloud in English to help my colleagues

in the group". Participant 18 mentioned that she analyzed the answers provided by her colleagues and corrected them when they were wrong. She stated, "I help my friends in the group, and I explain the answers to them, of course, after I understand myself". Participant 40 stated, "I correct my colleagues' answers and help them when they need me to". This suggests that collaborative learning encourages the students to provide their peers with assistance. This confirms the result presented by Abdelshaheed (2024), who reported that collaborative learning when conducted online provides language learners with support from their classmates.

Searching for information is one of the roles that participants perform in the group. All the participants reported engaging in searching for information. Participant 14 explained that each member of the group searched for information. She stated, "some of us can find the information faster than others while others fail to find the information, but we all search, and this is what we mostly do as we work in the group" (P14). Participant 24 mentioned searching for information as an essential task, she stated, "if I have not searched, then I have done nothing to contribute to the group". Participants 49 and 48 explained that searching for information is the second role they perform after translating unknown words, noting that searching is the essence of collaborative work. It appears that working collaboratively develops the students research skills and their critical thinking abilities, as they evaluate and assess the information they find. This result is consistent with that presented by Alkhalaf (2020), which confirms that collaborative learning develops critical thinking among female language students.

Another role that the participants reported performing was supervising group performance. Participant 1 explained that in every group, a leader supervises the discussion and activities. She stated, "I am always the leader, and I ensure that each member is given the chance to discuss her answers". Participant 11 mentioned that supervising group activities is something she does to help ensure her group completes tasks on time. Participant 52 mentioned that she sometimes supervises ongoing work, as a leader is needed in each group for each lecture and each out-of-the class task. This result confirms findings reported in previous studies (Al-Malki et al., 2022; Listyani, 2019) that show collaborative learning enhances language student leadership quality and interpersonal skills.

In addition to the role of supervising group performance, is that of actively sharing ideas and knowledge. Participant 6 mentioned her role was to share her knowledge and experience. Participants 21 and 26 illustrated that they present their ideas in the group and express their thoughts to others. The participants considered the act of sharing ideas and knowledge to be a specific role when working collaboratively in a group. This result corresponds with the previous literature (Aldossary, 2021; Chang & Windeatt, 2016; Heble & Mehta, 2013; McKay & Sridharan, 2024), which demonstrates how collaborative learning encourages students to share their ideas and knowledge.

Adopting different strategies to provide answers to assigned tasks is amongst the roles students reported when performing in the group. For example, Participant 43 mentioned that she verbalizes what she is thinking, externally processing ideas when working in the group. Participant 20 reported that she discusses possible answers orally with her peers and her peers either support her opinion, or correct her when she is wrong. Participant 18 stated, "I write down the answers in my notebook, and my peers read from the notebook when the teacher asks someone from our group to answer". Participant 15 mentioned that they divide the answers to the tasks between them after they have finished, to participate actively when the teacher asks them to respond to tasks.

Resources Used

The resources the students reported using when working collaboratively in groups were Artificial Intelligence (AI) apps, Google app, translation apps, and notes app. The participants reported using three AI apps: ChatGPT, Poe, and Canva (see Figure 2).

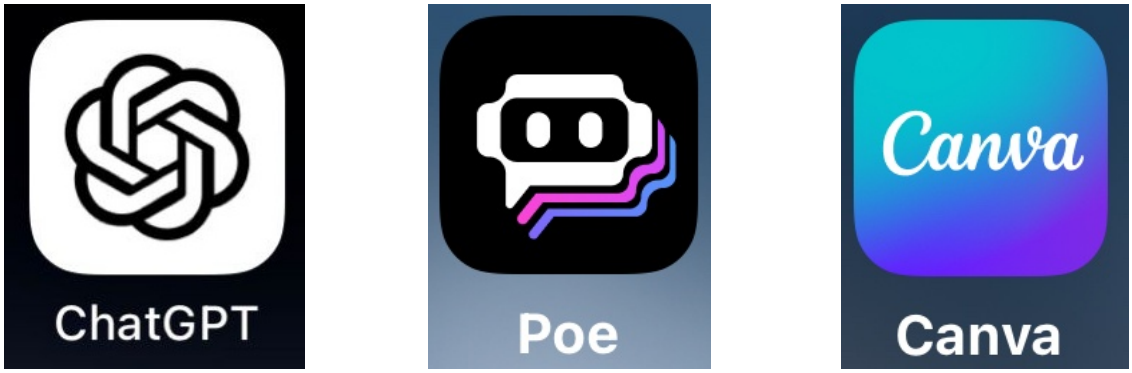


Figure 2.AI apps that were used by the participants

Most of the participants reported that they use ChatGPT to search when completing group work. Participants 10 and 12 mentioned having the ChatGPT app in their smartphone and using it to provide them with instant answers. The participants stated, "I use the ChatGPT app if answering the question is really hard and needs someone to simplify it for me and the group" (P22), "I know I should not say that! It is bad! but I use ChatGPT to check that our information is accurate" (P28). Participant 41 mentioned that she uses Poe to explain the questions in the task and to provide a simple answer. Participant 4 stated, "I used Canva because it helps me design PowerPoint presentations, using target terminology". This finding resonates with those of Alharthi (2024) and Pavlenko and Syzenko (2024), who observed that language students consider ChatGPT an essential tool for searching for information and completing language tasks. The Google app was also used by the participants to seek out information. The participants stated, "I mostly use Google to search" (P3), "In every class we search using Google. In fact, all of us search with Google" (P47), "What! Can you search without Google?!" (P50).

In addition to using AI apps and the Google app, participants reported using apps to assist them when translating from English to Arabic. Participant 50 mentioned using Google Translate to translate questions. The participants stated, "I take a photo and translate directly to understand the questions" (P2), "I copy and paste the text into Google Translate and did so several times" (P38). The participants also reported using the CollaNote app to translate from English to Arabic when working in a group. The finding relating to the use of Google Translate app are consistent with the finding of Bin Dahmash (2020), that female language students use the Google Translate app in-class as a resource for language learning.

Another resource used by participants when working collaboratively was a note app: Notability. This app assists students when writing down answers and creating charts to answer the questions in the tasks. Participant 4 mentioned that Notability helped them store and share their notes with the group via AirDrop when they have worked together. All the resources the students reported using were smartphone apps on their mobile phones.

Task Distribution Methods

The distribution of questions and items to complete tasks was decided by the leader, participants' preferences, number of questions and difficulty level. Participant 15 noted that every group has a leader not assigned by the teacher. Participants 14, 16, and 19 mentioned that the leader only divides up tasks but does not assign a member to each question, as the participants themselves choose the questions they think they can answer. Participants 36, 37, and 39 explained that the leader plans how she divides the questions, as each member must work according to her abilities and level of English. They mentioned that the member with the highest level of English takes the most challenging question and explains it to us, so that "we all learn, even those who are struggling with their English" (P37). Participant 4 illustrated that they agree with each other about how to divide up questions, and have freedom of choice, stating, "if there are many questions, we challenge ourselves to work beyond our capacity to succeed". Consistent with Oyarzun and Martin's (2023) finding, collaborative learning appears to promote personalized learning, as the participants choose the tasks they work on.

Group Formation Approaches

Strategies to form groups followed two approaches, i.e., either by the personal choice of participants or randomly. Some of the participants reported creating their own groups to work collaboratively with by selecting either peers with an outstanding academic level, or their friends. They mentioned that they chose colleagues who were proficient to improve their own performance. As participant 11 stated, "Success is contagious, and I like to work with those who are successful and learn from them". Those who chose their friends did so because they felt comfortable with them, as they encourage and motivate one another to accomplish tasks. The other participants mentioned creating groups randomly, either by choosing the nearest group, or by asking others if they want to join them as they do not have appropriate criteria. Participants 28, 29 and 30 reported that sometimes they work with those nearest to them in the class if they do not feel motivated to work, as working with friends causes more pressure as they want to please them.

The WhatsApp Group

The participants reported using online WhatsApp groups to collaborate during the course. The group was created using a smartphone application at the beginning of the term and includes all the students enrolled in the course as members; it was not supervised by the teacher. The participants shared handouts, discussed course requirements, and the resources they use to assist them in working collaboratively both in-class and out-of-class. Participants 5 and 3 stated that they created an additional WhatsApp group when they were required to work collaboratively to present concepts. They were asked to work and prepare for a slideshow outside the classroom, because they were required to present concepts to the class. This group affected what the participants did collaboratively in both modes in-class and out-of-class. The WhatsApp group resonates with Gitterman's (2008) hidden group, which influences students' interactions in collaborative learning environments.

Research Question (2): How Do Students Perceive the Effects of Collaborative Learning?

The data suggested the students had positive and negative views concerning the effects of collaborative learning, and that the positive outnumbered the negative.

The participants associated four positive effects with collaborative learning. First, it provided them with opportunities to practice the English language. The participants reported that working collaboratively provided them with opportunities to practice pronouncing business terminology and reading English texts. Participant 1 stated, "I rehearsed pronouncing revenues and expenditure with the group before the presentation". This result resonates with findings from the previous literature (Alzubi et al., 2025; Khan et al., 2020) that noted that collaborative learning provides language learners with additional opportunities to produce language.

The second effect was facilitating the understanding of information. The participants mentioned that negotiating and discussing information enabled them to learn in an enjoyable setting. Participants 11 and 13 stated:

P11: I learned about types of assets and liabilities from how we discussed them in the group. We laughed a lot as we had no assets or liabilities!

P 13: Yes, we had similar jokes about stocks and bonds. We own nothing, but I will never forget the meanings of these terms.

(Excerpt from focus group (2) translated from Arabic)

This result aligns with the finding of Listyani (2019), who reported that collaborative learning in language classes provides a more profound understanding of content. It also supports previous literature (Aldossary, 2021; Alkhalaf, 2020; Listyani, 2019) that characterizes collaborative learning as an enjoyable activity.

The third effect was preparing for examinations and making studying easier. The participants reported that working in groups reinforced knowledge and vocabulary retention, and that learning vocabulary improved their grades. The participants stated, "It improved my midterm grade, as I remembered how we had discussed some concepts in class" (P18) and "I will never forget the functions of business and I remembered them for the midterm" (P40), "it greatly improved my midterm grade" (P50). The results relating to reinforcing vocabulary retention is consistent with Al-Ahdal and Alharbi (2021), who found that collaborative learning improves vocabulary retention among English language learners at university level.

The fourth effect involved promoting social skills as teamwork, communication, and cooperation. The participants explained that working in groups encouraged them to help each other and share their knowledge. Participants stated, "it helped me develop my communication skills, as we support each other" (P1), "it motivated an environment in which we exchange ideas and help the group" (P6). This finding is consistent with that reported by Al-Malki et al. (2022), who noted that cooperative learning enhances students' social skills. This finding, as it relates to promoting teamwork, is consistent with Ojie-Ahamiojie (2024) who reported that collaborative learning develops teamwork skills.

Some negative views associated with collaborative learning were expressed by those participants who were unwilling to work in a group or accept what they viewed as overreliance on others. These participants reported that some students do not participate in the group, and benefit from the efforts of others. In addition, some rely on the most brilliant student to provide them with the answers. The participants also complained that they did not want to risk friendships or show a poor image of their group, so they did not speak with their teachers about colleagues who contributed little. Some examples of participants' comments are:

P41: Well, not all of us participate to the same degree

P44: You mean not participate at all. They might be silent all the time.

P42: When this happens, we do not tell you. We do not want to be snitches.

P46: Exactly, we do not want to embarrass ourselves in front of the others, so we do the activity with the cooperative members.

(Excerpt from focus group (7) translated from Arabic)

This finding, that not all students contribute well to group work, aligns with evidence presented by Hei et al. (2015) and Le et al. (2018), who classified students who rely on their peers in collaborative groups as free riders.

Conclusion

This paper examined how Saudi students engage in collaborative learning settings in a vocabulary course and their views regarding this approach. The roles the participants performed when engaging in collaborative learning were: translating from English to Arabic, assisting their peers, searching for information, supervising group performance, sharing ideas and knowledge, and adopting different strategies to answer assigned tasks. The resources they drew on were AI apps, Google app, translation apps, and notes app. Moreover, tasks were distributed and assigned by leaders in conjunction with participants' preferences, number of questions, and level of difficulty, and they formed groups either based on personal choice or randomly. Overall, the participants expressed more positive views about collaborative learning than negative views. Positive views included providing opportunities to practice the English language, facilitating understanding of information, preparing for examinations, making studying easier, and promoting social skills. Negative views included unwillingness to work with the group and frustrations with those who overly rely on others.

Based on the findings, several implications can guide the effective adoption of collaborative learning in language education. Teachers are encouraged to supervise the student's interactions, to ensure that each member participates, preventing 'free riders' in the group and helping each student reach their maximum potential. Teachers were advised to encourage their students to draw on AI apps and Google apps while collaborating. Policymakers could also promote the implementation of collaborative learning in language classrooms, as it offers opportunities for language practice and simplifies the learning process.

This research was limited by the gender of the participants and the data collection tools used. Future studies could include male students, which would provide a more comprehensive view of how language students engage in collaborative learning. Additionally, employing mixed methods along with both qualitative and quantitative data collection tools could lead to a deeper understanding of the impact of collaborative tasks on the learning process.

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 Grammarly to improve readability and language.

Institutional Review Board Statement: The study was approved by the Research Ethics Committee of King Saud University (reference number: KSU-HE-24-1046 and 26/11/2024).

About the Author

Nada Bin Dahmash is an associate professor at King Saud University and a member of the Saudi Linguistics Society. She has taught courses in linguistics, general English, and English for Specific Purposes. Her research focuses on second language practices, EFL learning, and the impact of technology on language learning. https://orcid.org/0000-0001-5918-5124

References

Abdelshaheed, B. (2024). Digital Collaborative Learning: Promoting Reading Comprehension and Vocabulary among Saudi University English Majors. *Arab World English Journal*, *10*(Special Issue on CALL), 118-133.

Al Jawad, A. S. H. (2023). The Effect of Using Cooperative Learning Method in Enhancing EFL Students' Performance in Speaking Skill in Libyan Universities: A Case Study of Benghazi University. *International Journal of Linguistics, Literature and Translation*, *6*(6), 64-74.

Al-Ahdal, A. A. M. H., & Alharbi, M. A. (2021). MALL in Collaborative Learning as a Vocabulary-Enhancing Tool for EFL Learners: A Study Across Two Universities in Saudi Arabia. *Sage Open*, *11*(1), 1-9.

Al-Malki, E. A. A., Gulnaz, F., Zahid Javid, C., & Chaudhry, M. Z. (2022). An Empirical Investigation of the EFL Learners Attitudes Toward the Effectiveness of Cooperative Learning. *Sage Open*, *12*(4), 1-12.

Aldossary, K. S. (2021). The Impact of Collaborative Writing on EFL Learners' Writing Development: A Longitudinal Classroom-Based Study in Saudi Arabia. *Arab World English Journal*, *12*(3), 174 -185.

Alharthi, S. M. (2024). Beyond Traditional Language Learning: EFL Student Views on ChatGPT in Saudi Arabia. *Arab World English Journal*, *10*(Special Issue on CALL), 15-35.

Alkhalaf, N. A. (2020). Saudi Female EFL Learners and Collaborative Writing: Attitudes and Challenges. *Theory and Practice in Language Studies*, *10*(9), 1118-1127.

Alzubi, 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.

Alzubi, A., Nazim, M., & Ahamad, J. (2025). Identifying English as a Foreign Language Students' Attitude to Improving Speaking Skills through Collaboration. *Journal of Education and Learning (EduLearn)*, *19*(1), 180-190.

Archibald, M. M., Ambagtsheer, R. C., Casey, M. G., & Lawless, M. (2019). Using Zoom Videoconferencing for Qualitative Data Collection: Perceptions and Experiences of Researchers and Participants. *International Journal of Qualitative Method*, *18*, 1-18.

Bin Dahmash, N. (2020). 'I Can't Live Without Google Translate': A Close Look at the Use of Google Translate App by Second Language Learners in Saudi Arabia. *Arab World English Journal*, *11*(3), 226-240.

Braun, V., & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology*, *3*(2), 77–101.

Brinkmann, S., & Kvale, S. (2018). *Doing Interviews* (2nd ed.). London: SAGE.

Bruffee, K. A. (1999). *Collaborative Learning: Higher Education, Interdependence, and the Authority of Knowledge* (2nd ed.). Baltimore: Johns Hopkins University Press.

Chang, H., & Windeatt, S. (2016). Developing Collaborative Learning Practices in an Online Language Course. *Computer Assisted Language Learning*, *29*(8), 1271-1286.

Dörnyei, Z. (2007). *Research Methods in Applied Linguistics : Quantitative, Qualitative and Mixed Methodologies*. Oxford: Oxford University Press.

Gitterman, A. (2008). Collaborative Learning and Teaching. *The Writing Center Journal 28*(2), 60-71.

Heble, A., & Mehta, S. R. (2013). Collaborative Student Writing in the Literature Classroom. *Arab World English Journal*, *4*(2), 173-184.

Hei, M. S. A. D., Strijbos, J.-W., Sjoer, E., & Admiraal, W. (2015). Collaborative Learning in Higher Education: Lecturers' Practices and Beliefs. *Research Papers in Education*, *30*(2), 232–247.

Khan, R. M. I., Mustafa, G., & Awan, A. A. (2020). Learners' Attitudes on the Infusion of Cooperative Learning in Education. *Orient Research Journal of Social Sciences*, *5*(2), 164-175.

Le, H., Janssen, J., & Wubbels, T. (2018). Collaborative Learning Practices: Teacher and Student Perceived Obstacles to Effective Student Collaboration. *Cambridge Journal of Education*, *48*(1), 103-122.

Lew, S., Yan, A. H., & Harklau, L. (2018). Qualitative Methodology. In A. Phakiti, P. De Costa, L. Plonsky, & S. Starfield (Eds.), *The Palgrave Handbook of Applied Linguistics Research Methodology* (pp. 79-101). London: Palgrave Macmillan UK.

Listyani, L. (2019). A Collaborative Peer-Consulted Text Analysis: Students' Perspectives. *Arab World English Journal*, *10*(4), 151-166.

McKay, J., & Sridharan, B. (2024). Student Perceptions of Collaborative Group Work (CGW) in Higher Education. *Studies in Higher Education*, *49*(2), 221-234.

McWhaw, K., Schnackenberg, H., Sclater, J., & Abrami, P. C. (2003). From Co-operation to Collaboration: Helping Students Become Collaborative Learners. In A. Ashman & R. Gillies (eds.), *Cooperative Learning: The Social and Intellectual Outcomes of Learning in Groups* (pp. 69-86). London: Routledge.

Ojie-Ahamiojie, G. (2024, October 16). *Using Collaborative Learning to Elevate Students' Educational Experiences*. Higher Ed Teaching & Learning. Retrieved September 14, 2025 from https://www.facultyfocus.com/articles/faculty-development/using-collaborative-learning-to-elevate-students-educational-experiences/

Oyarzun, B., & Martin, F. (2023). A Systematic Review of Research on Online Learner Collaboration from 2012-21: Collaboration Technologies, Design, Facilitation, and Outcomes. *Online Learning*, *27*(1), 71-106.

Palys, T. (2008). Purposive Sampling. In L. M. Given (Ed.), *The SAGE Encyclopedia of Qualitative Research Methods* (pp. 698-699). Thousand Oaks: SAGE Publications.

Pavlenko, O., & Syzenko, A. (2024). Using ChatGPT as a Learning Tool: A Study of Ukrainian Students' Perceptions. *Arab World English Journal* (Special Issue on ChatGPT), 252-264.

Qureshi, M. A., Khaskheli, A., Qureshi, J. A., Raza, S. A., & Yousufi, S. Q. (2023). Factors Affecting Students' Learning Performance through Collaborative Learning and Engagement. *Interactive Learning Environments*, *31*(4), 2371-2391.

Weinberger, Y., & Shonfeld, M. (2020). Students' Willingness to Practice Collaborative Learning. *Teaching Education*, *31*(2), 127-143.

Yang, X. (2023). A Historical Review of Collaborative Learning and Cooperative Learning. *TechTrends* *64*(4), 718-728.

Appendices

Appendix A: Focus Group Interview

Previous experience

1. Have you ever worked collaboratively in groups to present a topic or to answer questions as part of your English course?
2. What role do you usually perform when asked to work collaboratively in a group? Why?

Current experience:

1. How did you start answering the tasks/ questions in the group in the classroom? Describe in detail.
2. Did you choose the group members yourself? How?
3. When instructed to complete several tasks in a group, what did you do and why? Explain.
4. How did you distribute the questions/tasks in the group?
5. Did you use the Internet or any online resources when working in the group? Clarify.
6. Did working collaboratively enhance your vocabulary learning? How?
7. How did working collaboratively affect your midterm/coursework grades? Describe in detail.
8. Would you recommend continuing to implement collaborative learning in this course? Why?
9. Do you have a group related to this course that works together outside the classroom? Describe this group.
10. What is the best approach to implementing collaborative learning in English courses? Finally, is there anything you want to add that is not mentioned in the questions?

Appendix B: Retrospective Interviews

When choosing a task in the classroom:

What did you do when instructed to work collaboratively in the classroom? Describe this in detail, including how duties were distributed and how you communicated with your peers. After the time specified for the task was over, what did you do next? Describe the process you followed, e.g., reviewing the answers or discussing them with your peers.

When choosing a task outside the classroom:

When asked to present a topic from the course in front of the class, how did you prepare for the presentation with your group? Describe key steps in detail, e.g., collecting information, designing PowerPoint slides, and rehearsing the presentation.

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp. 388-406
DOI: https://dx.doi.org/10.24093/awej/vol16no3.24

Past Tense Forms in CEFR-Aligned English Textbooks Used at Malaysian Primary Schools and Their Directive Functions

Patrick Duffy Bayuong
Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia
Corresponding Author: p116380@siswa.ukm.edu.my

Harwati Hashim
Faculty of Education, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia
Email: harwati@ukm.edu.my

Melor Md. Yunus
Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia
Email: melor@ukm.edu.my

Received: 06/29/2025 Accepted:09/11/2025 Published: 09/20/2025

Abstract
Incorporating English past tense with its forms using the appropriate approach in Common European Framework of Reference (CEFR)-aligned English Textbooks at the primary level is essential for assisting Malaysian Primary School English as a Second Language (ESL) learners in understanding the past tense forms, including the distinction between regular and irregular forms. The study aimed to analyse the English past tense forms presented in CEFR-Aligned English Textbooks for Malaysian Primary Schools and to identify their directive functions in ESL learning at Malaysian Primary Schools. The study addressed a key question: how are English past tense forms presented and used in CEFR-aligned textbooks for English as a Second Language (ESL) learners? It is significant because a solid understanding of the past tense is essential for developing linguistic competence in ESL learners. The study employed a qualitative research design through document analysis, analysing four CEFR-aligned English Textbooks used in Malaysian Primary Schools. The study revealed four English past tense forms and three directive functions of these forms in ESL learning at Malaysian Primary Schools. The study, as evidenced in the textbooks, stated that the grammatical rules concerning English past tenses are not taught in isolation but are incorporated within rich contexts that primarily promote linguistic competence. The study emphasised the crucial need for transitioning ESL instruction from simply memorising the grammar rules to a more communicative and context-driven approach.

Keywords: Common European Framework of Reference - aligned English textbooks, directive functions, English past tense forms, English as a second language learners, grammar instruction, Malaysian Primary Schools
How to cite: Bayuong, P. D., Hashim, H., & Yunus, M.M. (2025). Past Tense Forms in CEFR-Aligned English Textbooks Used at Malaysian Primary Schools and Their Directive Functions *Arab World English Journal, 16*(3):388-406. https://dx.doi.org/10.24093/awej/vol16no3.24

Introduction

Mastery of English grammar, particularly English past tense forms, is a critical component of English as a Second Language (ESL) learning for Malaysian primary school learners to express, communicate, describe, or construct narratives about, depending on the contexts, the completed, ongoing, and simultaneous actions or events. Cowan (2008) stated that the tense in a verb conveys the time an action happens with the present moment of speaking. Communicating in the correct tenses competently engages ESL learners in communicating accurate details and conforming to grammatical rules (Declerck et al., 2006; Javed & Ahmad, 2013). In the context of English grammar, learning English tenses is imperative because they indicate time reference and determine how well ESL learners communicate in both verbal and written forms (Almunawaroh et al., 2021).

The current English textbooks for primary schools designed by the Ministry of Education Malaysia (MOE) adhere to the Common European Framework of Reference for Languages (CEFR). The development of these CEFR-aligned textbooks emphasised the integration of grammatical structures into meaningful contexts to enhance primary school ESL learners' understanding of the application of English past tense forms and empower their language acquisition. In addition, the grammatical rules and examples for both regular and irregular past verbs were embedded within the textbooks to achieve similar instructional objectives. Fadl (2019) argued that the constructions of the past tense materials influence learners' ability to acquire the targeted grammar items. Furthermore, these materials facilitate the learning of past tenses and enhance learners' knowledge. This study acknowledges the MOE's initiative to develop grammar learning and instruction materials, a directive move that aligns with Larsen-Freeman's Three Dimensions (2003). The dimensions – form, meaning, and use – provide a well-rounded framework for teaching grammar that effectively enhances learners' grammatical competence.

Despite these laudable initiatives and the general attention given to grammar instruction in Malaysian ESL contexts, there remains a critical need for focused research. Specifically, there is a paucity of studies examining the detailed presentation and pedagogical implications of English past tense forms within the recently implemented CEFR-aligned Malaysian primary school English textbooks. This study aimed to fill that gap by investigating the pedagogical role of these forms, particularly focusing on their directive functions. These "directive functions" refer to how the specific presentation, examples, and activities related to past tense forms within the textbooks implicitly or explicitly guide ESL learners towards understanding and correctly applying these grammatical structures, thereby directing their learning pathways.

Accordingly, this study aimed to address the following research objectives (RO):

1. RO1: To analyse the English past tense forms used in CEFR-Aligned English Textbooks for Malaysian Primary Schools.
2. RO2: To identify the directive functions of English past tense forms in ESL learning at Malaysian Primary Schools.

Based on the RO, the research questions (RQ) were formulated as follows:

1. RQ1: What English past tense forms are used in CEFR-Aligned English Textbooks for Malaysian Primary Schools?
2. RQ2: How do English past tense forms function as directives in ESL learning at Malaysian Primary Schools?

ISSN: 2229-9327

Literature Review

English Language Learning in Malaysian Primary Schools

English is a crucial subject in the Malaysian primary school curriculum, serving as an ESL (Aziz, 2007; Gill, 2002; Sidhu et al, 2018; Thirusanku & Yunus, 2014) and essential language for communication. The emphasis on English grammar learning, particularly the past tense, reflects its significance in enabling ESL learners to describe completed actions correctly and communicate effectively. Teaching grammar is important for developing ESL learners' spoken and written English competence. Nevertheless, challenges remain, including ESL learners' reliance on rote learning and lack of opportunities for contextual grammar usage, particularly in rural areas. Socio-economic and geographical factors caused disparities in English language proficiency among ESL learners in rural and urban primary schools. This has led to a shift away from traditional "chalk-and-talk" methods toward more engaging, task-based strategies, as studies have shown that language games are beneficial for enhancing learners' knowledge, competency, and motivation in learning grammar (Yaccob & Yunus, 2019). In alignment with the Industrial Revolution 4.0, these strategies increasingly involve technology. Research has further indicated that using online, gamified learning tools can lead to a significant improvement in ESL learners' grammar scores (Hashim et al., 2019). The MOE has addressed English language learning issues by developing the CEFR-aligned curriculum (MOE, 2015, 2019a), which integrates meaningful language use with grammar instructions. In addition, this curriculum prepares ESL learners to apply their knowledge beyond classroom contexts and speak English confidently.

CEFR-Aligned English Textbooks

Textbooks have been fundamental tools in primary education since the 19th century (Kovac & Alenka, 2022), serving as learners' primary sources of subject matter (Lawyer, 2024). Implementing CEFR-aligned English textbooks in Malaysian primary classrooms represents a significant and necessary shift in language instruction. The materials in the imported textbooks were adapted to the Malaysian context (Shak et al., 2021). Generally, these textbooks are used as instructional materials for teachers and learning resources for learners (Bakir & Aziz, 2022). In the grammar learning context, these textbooks were designed to integrate grammar structures, including past tense, into meaningful contexts to enhance ESL learners' acquisition and language skills. By embedding grammar instruction contextually, these materials enable ESL learners to understand the relationship between verb forms.

Beyond simply presenting the grammar rules, effective textbook design inherently guides learners through the acquisition process. The pedagogical features embedded within these materials, such as explicit instructions, illustrative examples, structured practice exercises, and communicative tasks, serve as crucial 'directives' for learners, a principle emphasised in materials development scholarship (Richards & Renandya, 2002; Tomlinson, 2011). These directives dictate how grammar is encountered, processed, and ultimately applied, reflecting contemporary approaches to language teaching (Özçelik & Kent, 2023). For instance, research highlighted that the strategic sequencing of activities and the clarity of explanations for both regular and irregular verb patterns are vital in steering the learner's cognitive engagement (Angelovska, 2017). Furthermore, the deliberate provision of scaffolding in exercises and tasks is known to facilitate learners' understanding and production of target grammatical forms (Bayuong & Arip, 2024; Green, 2023). Thus, the ways the past tense forms are presented and reinforced within these CEFR-

aligned textbooks implicitly or explicitly direct the learning pathway for primary school ESL learners.

With CEFR's adaptations, the standards are established internally, considering the needs of learning ESL in Malaysia and aligning with the national goals and aspirations (Aziz, 2017; MOE, 2015; Sani, 2018). Furthermore, CEFR is a valuable framework for Malaysia's curriculum, adaptable to the existing curriculum (MOE, 2015), addressing ESL learners' communicative needs (Chin, 2018), and enhancing the outcomes of English language education in Malaysia (Bayuong & Hashim, 2023). The CEFR-aligned approach emphasises grammar teaching through activities that promote interaction and engagement. For instance, past-tense exercises often include storytelling, role-playing, and sentence constructions that reflect authentic language use. Contextualised learning aids ESL learners in internalising grammar rules, particularly differentiating between regular and irregular verb forms. Moreover, including scaffolded exercises in the textbooks empowers ESL learners' progress and enhances their understanding of the past tense forms. Bayuong & Arip (2024) highlighted the importance of CEFR-aligned textbooks in presenting English grammar in a way suitable for ESL learners' age and enriched with context, thereby improving their comprehension and practical use of language. Effective teacher training and supplementary resources are indispensable to maximising the potential of the textbooks.

English Past Tense

Tense is defined as a linguistic concept that represents the correlation between the timing of an event and the moment of speaking (Declerck & Cappelle, 2006), a lexical element that helps to locate an event or situation in time (Chomsky, 1999, 2002; Manokaran et al., 2013), a grammatical structure that shows the timing of action about a particular moment (Greenbaum & Nelson, 2002), a verb form that indicates an event's time, condition, action, or occurrence, and a morphological representation of an event or activity within a specific temporal framework (Javed & Ahmad, 2013). Verb tenses form the essential building blocks of sentence grammar (Bukit, 2020; Fanny et al., 2022; Lista & Febriyanti, 2020), thus teaching them is crucial in enhancing language proficiency (Nijat et al., 2020). Speaking of the past tense, many linguists, according to Bardovi-Harlig (2000), argued that the English language operates with only two tenses – the present and past tenses. Like the present and future tense, past tense consists of four forms – past indefinite, past continuous, past perfect, and past perfect continuous tense, which, according to Chomsky (1999), have phase-structure rules and formulas that are universally used in the English language. The past tense and tense aspects (or forms) are difficult grammar items for ESL learners to learn (Apte, 2004; Larsen-Freeman et al., 2011; Parrott, 2000). Past studies have shown ESL learners made grammar errors in the past tense in their writing (Azlan et al., 2015; Fanny et al., 2022; Lionny & Kusumadewi, 2022; Manokaran et al., 2013; Sazali et al., 2024). Literature showed that learners tended to rely on temporal adverbials to indicate the past tense or time reference (McArthur, 2002; VanPatten, 2015). For advanced learners, they tended to depend on inflection (VanPatten, 2015). Chan (2013) asserted that Cantonese ESL learners implicitly comprehend tense. However, they struggle to associate the appropriate forms with their meaning.

Previous studies suggested the existence of a natural order or acquisition of regular and irregular past tense verb forms (Goldschneider & DeKeyser, 2001; Jia & Fuse, 2007; Salaberry, 2000). However, Ellis (2009) reported that the acquisition order might not apply in all contexts. Literature reported that learners acquire regular past tense forms by adding "-ed" to the verb root they already knew and that they have newly encountered (as in rick – ricked in Berko, 1958)

(Pinker & Ullman, 2002). In contrast, learners often memorise irregular past tense forms as fixed units (e.g., fall–fell), which could also lead to overgeneralisation errors (e.g., sit–sitted) (Pinker & Ullman, 2002; Prasada & Pinker, 1993).

In learning English tenses, studies showed that various activities in the language learning materials aid learners in automatising their existing knowledge, developing their competence (Canale & Swain, 1980), and gradually gaining their confidence (Fadl, 2019). The current CEFR-aligned English textbooks consist of engaging grammar learning activities and exercises to assist Malaysian primary ESL learners in learning and acquiring English past tenses, contextually. Including English past tense with its forms in the appropriate approach in these learning materials at the primary level is essential for assisting ESL learners in understanding the structures of past tense verbs, including the distinction between regular and irregular forms. Moreover, this enables these learners to conjugate verbs in the past tense correctly and use them effectively in sentences, supporting their language skills and grammar knowledge.

Method

This study employed a qualitative research design through document analysis. The data sources comprised four CEFR-aligned English Textbooks used in Malaysian Primary Schools: *Get Smart Plus 3 Student's Book, Get Smart Plus 4 Student's Book, English Plus 1 Second Edition Year 5 Student's Book, and Academy Star 5 Year 6 Pupil's Book*. These textbooks were selected based on authenticity, credibility, representativeness, and meaning, following Bowen's (2009) guidelines for document selection in qualitative research. In addition, the selection was further justified based on their status as core textbooks in the national curriculum for English language teaching at Malaysian primary schools, ensuring alignment with CEFR guidelines.

The data were analysed thematically. Thematic analysis allows researchers to identify trends in the data, thus highlighting emerging themes (Bowen, 2009; Fereday & Muir-Cochrane, 2006). The thematic analysis facilitated the identification of patterns and recurring themes related to the use of English past tense forms. The coding process was conducted manually to ensure a systematic and thorough review of the data. This process enabled the identification of the types and directive functions of English past tense forms in CEFR-aligned English textbooks for Malaysian primary schools, providing insights into their pedagogical applications in ESL learning contexts.

Results

The English Past Tense Forms Used in CEFR-Aligned English Textbooks for Malaysian Primary Schools

The analysis of CEFR-aligned English textbooks used in Malaysian primary schools revealed significant past tense forms. The findings, organised into distinct themes, provided valuable insights into the linguistic structures presented to ESL learners at Malaysian primary schools. Table 1 demonstrates the themes (English past tense forms) that emerged from the data analysis. This thematic analysis outlined the findings into four main themes: Simple Past Tense, Past Continuous Tense, Simple Past Tense and Past Continuous Tense, and Simple Past Tense and Present Participle. The following section delves into the specific past tense forms identified, each substantiated by specific examples extracted from the textbooks. It offers a detailed examination of their usage within the textbooks analysed in this study.

Table 1. *English past tense forms*

Themes		Evidence
F1	Simple Past Tense (SPT)	On Saturday morning, I was at the shopping mall with my mother. Extracted from Get Smart Plus 3 Student's Book, m8 The Egyptians washed the body first to make a mummy. Extracted from Get Smart Plus 4 Student's Book, m3 He loved the job and had a lot of friends and fun. Extracted from English Plus 1 Second Edition Year 5 Student's Book, u7 He said he loved sugar. Extracted from Academy Stars 5 Year 6 Pupil's Book, u9 I used to play with toys when I was little. Extracted from Academy Stars 5 Year 6 Pupil's Book, u2 The men jumped onto the beach and took two spades from the boat. Extracted from Academy Stars 5 Year 6 Pupil's Book, u3
F2	Past Continuous Tense (PCT)	Her cousin was visiting her from Australia. Extracted from Academy Stars 5 Year 6 Pupil's Book, u3
F3	Simple Past Tense and Past Continuous Tense (SPTPCT)	I was singing when suddenly all the lights went out. Extracted from Academy Stars 5 Year 6 Pupil's Book, u1 Mr Carter took all our names while we were standing outside. Extracted from Academy Stars 5 Year 6 Pupil's Book, u1 Zara said she was doing a project about food. Extracted from Academy Stars 5 Year 6 Pupil's Book, u9
F4	Simple Past Tense and Present Participle (SPTPP)	They looked down to the sea and saw a small boat arriving with three men on board. They could hear the men shouting to each other. Extracted from Academy Stars 5 Year 6 Pupil's Book, u3

Note: F – Form

F1: Simple Past Tense

The simple past tense is prominently featured in the textbooks. It is one of the core elements of the English past tense that ESL learners need to learn and master. The sentence "On Saturday morning, I was at the shopping mall with my mother" (MOE, 2018) employed the simple past tense of the verb "was", situating the action within a specific time frame and place. This structure deepens ESL learners' comprehension of the narrative's context. Additionally, this structure demonstrates how ESL learners could connect their personal experiences to grammatical structures, thus enriching engagement with the language. The sentence "The Egyptians washed the body first to make a mummy" (MOE, 2019b) expresses a completed action using the verb "washed", establishing a historical event. This structure incorporates cultural and historical knowledge, contributing to ESL learners' language acquisition and cultivating their educational experiences.

Another example, "He loved the job and had a lot of friends and fun" (MOE, 2021a) utilises the simple past tense form to demonstrate feelings and relationships, narratively. This structure offers ESL learners valuable chances to express emotions through language accurately. Furthermore, the simple past tense form was observed in the reported speech as exemplified in "He said he loved sugar" (MOE, 2021b). This structure is essential for facilitating ESL learners to practice indirect discourse, thus developing their communicative competence in both spoken and written forms.

The sentence "I used to play with toys when I was little" (MOE, 2021b) exemplifies a habitual action in the past, enabling ESL learners to convey their regular activity during their childhood. The structure implies that the action was an ongoing behaviour rather than a one-time activity. However, this activity is no longer happening as identified through the verb phrase "used to" in the simple past tense form. The sentence "The men jumped onto the beach and took two spades from the boat" (MOE, 2021b) utilises two simple past tense verbs in a context. This structure provides a detailed description of the sequential actions, enhancing ESL learners' mastery of using more than one simple past tense verb in a sentence. Additionally, this structure enhances ESL learners' use of grammatical structures and reinforces their writing skills, particularly in the narrative context.

F2: Past Continuous Tense

The past continuous tense is another form of the past tense observed in textbooks. The given example "Her cousin was visiting her from Australia" (MOE, 2021b) showcases an ongoing action "visiting" in the past, which had ended entirely. This past tense form distinguishes between a completed action (simple past tense) and a progressing action at a specific time in the past. This structure, although not providing a specific time, clearly emphasises the action that occurred in the past through the use of the auxiliary and present participle verbs in the phrase "was visiting". This phrase indicates that the cousin's visit was an ongoing and temporary event at a specific time in the past. Furthermore, it implies the state of being present of the cousin or that the cousin was physically staying with her.

F3: Simple Past Tense and Past Continuous Tense

This theme exemplifies the combination of the simple past tense and past continuous tense, showcasing specific meanings and nuances. To illustrate, the sentence "I was singing when suddenly all the lights went out" (MOE, 2021b) utilises the past continuous tense in its first clause "I was singing" and the simple past tense in its second clause "when suddenly all the lights went out". This structure conveys the suddenness and instantaneous action "went out" that interrupted a progressing background action "singing" that the subject was engaged in. This construct emphasises the simultaneous use of two past tense forms, fostering ESL learners' use of multiple verb tenses within a sentence to convey a coherent narrative flow. Similarly, the sentence "Mr Carter took all our names while we were standing outside" (MOE, 2021b) indicates the completion of an action "took" in the middle of an ongoing action "standing".

On the contrary, the sentence "Zara said she was doing a project about food" (MOE, 2021b) establishes a different context. The first clause, "Zara said," which is in the simple past tense, shows that the action of the subject speaking happened in the past. It implies that the statement or information communicated by the subject had happened entirely. The second clause, "she was doing a project about food," is in the past continuous tense, describing a progressing action that the subject was engaged in. Additionally, this sentence combines reported speech with ongoing action, aiding learners' language mastery to accurately communicate their thoughts and experiences using the past tense forms.

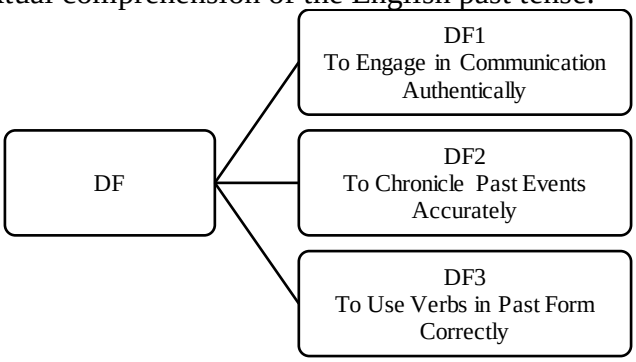
F4: Simple Past Tense and Present Participle

The theme of the simple past tense and present participle is represented by the sentences "They looked down to the sea and saw a small boat arriving with three men on board" and "They

could hear the men shouting to each other" (Academy Stars 5 Year 6 Pupil's Book). These sentences express that different past tense forms can convey coexisting actions within a single sentence, offering richness to the information. The clause "They looked down to the sea" employs the simple past tense to set the main scene using the verb "looked down" followed by "saw". The second clause "and saw a small boat arriving with three men on board" employs the present participle verb "arriving" to describe a subsequent action that concurred with the actions "looked down" and "saw". The second sentence further illustrates this theme by utilising the present participle verb "shouting" to describe a subsequent action that coexisted with the simple past action "could hear", aiding in the immediacy of the sentence.

The Directive Functions of English Past Tense Forms in ESL Learning at Malaysian Primary Schools

The analysis of CEFR-aligned English textbooks used in Malaysian primary schools highlighted the pivotal roles of the past tense forms in facilitating effective ESL learning. Figure 1 visually illustrates the themes (Directive functions of English past tense forms) that emerged from the data analysis. The themes underscored their relevance to developing ESL learners' practical and contextual comprehension of the English past tense.



Note: DF – Directive Function
Figure 1. Directive functions of English past tense forms

The past tense is a grammatical construct and a functional tool in textbooks that supports ESL learners in various communicative contexts. This study identified three directive functions of English past tense forms, essential for enhancing ESL learners' language competence. These include authentic communication (DF1), chronicling past events accurately (DF2), and correctly using verbs in past form (DF3). These functions align with the CEFR's emphasis on contextualised and meaningful language use. As a result, they ensure ESL learners can apply their knowledge in real-world contexts. English teachers can help ESL learners achieve greater linguistic accuracy and fluency by integrating these functions into their ESL instruction.

DF1: To Engage in Communication Authentically

DF1 emphasises the importance of ESL learners using the past tense to create meaningful and genuine interactions in their ESL learning. Incorporating past tense forms into their communicative activities enables ESL learners to share personal experiences, narrate past events, and respond to real-life contexts. This fosters a sense of authenticity in their language use. This strategy drives ESL learners beyond memorising grammar rules and empowers them to use the

past tense in practical and contextually relevant scenarios. This authenticity improves their confidence in English and enhances their ability to communicate effectively in diverse settings. However, the authenticity of communication may be affected. Lista and Febriyanti (2020) asserted that the interference of learners' first language (L1) in their English language production is common.

The sentence "Zara said she was doing a project about food" (MOE, 2021b) employs the simple past tense verb "said" and the past continuous verb form "was doing". This sentence demonstrates the speaker conveying reported speech and ongoing past action, highlighting the practical application of grammar in authentic communication. ESL learners can use this structure to express, share, and report details from past experiences in written and verbal forms.

On the contrary, the sentence "On Saturday morning, I was at the shopping mall with my mother" (MOE, 2018) is in the simple past tense, using the verb "was" to detail a specific event. This structure allows ESL learners to set a particular scene and provide context to their communication, enriching the authenticity of their conversations through the narration of real-life situations and personal experiences. Furthermore, this structure guides them to converse daily with people close to them, such as friends, relatives, classmates, and teachers, telling them about their activities at a specific point in the past. For instance, they can say "Yesterday, I was at Tony's house" or slightly change the structure by placing the time reference "yesterday" at the end, as in "I was at Tony's house yesterday".

DF2: To Chronicle Past Events Accurately

DF2 focuses on the use of past tense forms to narrate past events in a clear and precise manner. This function guides ESL learners to sequence actions (and verbs) and effectively describe historical, personal, or fictional events. In their ESL learning, using past tense forms enables ESL learners to construct coherent narratives and accurately communicate past experiences or events. The exercises and activities focusing on recounting past events in the CEFR-aligned textbooks help ESL learners practice using regular and irregular verbs correctly. DF2 strengthens ESL learners' grammatical proficiency and enhances their ability to convey meaningful stories, past events, or experiences in a chronological manner.

For instance, the sentence "The men jumped onto the beach and took two spades from the boat" (MOE, 2021b) illustrates sequential actions. This example not only empowers ESL learners' mastery of using more than one simple past tense verb in a sentence but also helps them to use multiple past tense verbs in narrating events cohesively. This will enable them to construct detailed and engaging accounts of past actions with grammatical precision.

Next, the sentence "I was singing when suddenly all the lights went out" (MOE, 2021b) showcases the sequencing of ongoing and completed actions in the past. In this example, ESL learners can learn to use two past forms – past continuous and simple past tense – together to narrate events, effectively and with grammatical precision. The sentence "They looked down to the sea and saw a small boat arriving with three men on board" (MOE, 2021b) demonstrates to ESL learners how they can narrate, for instance, the completed actions "They looked down to the sea and saw..." while introducing concurrent descriptions "… a small boat arriving with three men on board".

DF3: To Use Verbs in Past Form Correctly

DF3 highlights the development of ESL learners' ability to apply grammatical rules for past-tense verb conjugation accurately. The ESL learners distinguish between regular and irregular verbs and comprehend their correct forms in various contexts. This can be a difficult task for these learners for several reasons. For instance, Parrott (2000) asserted that some learners struggled with selecting the appropriate tense because their native language does not differentiate in this way. Additionally, tense marking is a significant challenge for second language (L2) learners because the tense system of their L1 differs from that of the L2 (Ambridge et al., 2014; Ayoun, 2015; Handayani & Johan, 2018; Housen & Simoens, 2016; Lista & Febriyanti, 2020; Slabakova, 2015). The struggle extends beyond simply selecting the correct tense (past, present, and future) – it also involves difficulties in applying the correct grammar rules for regular and irregular verbs in the past tense. This goes a similar justification to DF1. Therefore, using past-tense verbs, ESL learners can construct grammatically correct sentences. Additionally, DF3 aligns with the CEFR's emphasis on grammatical accuracy, ensuring ESL learners can express actions or events in the past and improve their language precision.

Expanding on the above discussion, the textbook examples demonstrate everyday contexts in which ESL learners might encounter the application of verbs in past tense forms, highlighting both regular and irregular verb forms. For instance, the sentence "The men jumped onto the beach and took two spades from the boat" (MOE, 2021b) illustrates the usage of two past-tense verbs, "jumped" and "took", in a single sentence. Additionally, this sentence demonstrates the combination of a past regular "jumped" and the irregular verb "took" in a sentence, enhancing ESL learners' usage of these types of verbs. As stated earlier in the above discussion, "… it also involves difficulties in applying the correct grammar rules for regular and irregular verbs in the past tense" – this example enables ESL learners to distinguish the different forms of past tense verbs and employ the correct grammar rule for the regular and irregular verbs in a meaningful context.

In the subsequent sentences "On Saturday morning, I was at the shopping mall with my mother" (MOE, 2018), and "Her cousin was visiting her from Australia" (MOE, 2021b), the verb "was" is used, demonstrating the utilisation of subject-verb agreement (SVA) in the past tense. This can be challenging for ESL learners when using the irregular verb "be" in its past forms. To this, a study by Manokaran et al. (2013) reported that Malaysian ESL learners (from formal schooling up to college level) made errors in the use of the past tense auxiliary "be" in their written works due to the difficulties they faced regarding this grammatical item. The verb "was" is used as the main verb in the simple past form in the first sentence, whereas it is used as an auxiliary or helping verb to the main verb "visiting" in the second sentence to express the past continuous tense. Also, the second example pinpoints the importance of ESL learners comprehending how the verb "be" in its past form, which functions as an auxiliary verb, works with the main verb "visiting" to convey continuous action in the past accurately.

In the sentence "I was singing when suddenly all the lights went out" (MOE, 2021b), ESL learners are exposed to multiple aspects of the past tense form. First, all three verbs are irregular. Secondly, the verbs "was" and "went out" are in their simple past forms, whereas the verb "singing" is in the present participle. Thirdly, the verb "was" is used as an auxiliary verb to the main verb "singing" to express an ongoing action effectively. Next, in contrast, the verb phrase "went out" is used to express a completed action within the ongoing action. The sentence "They looked down to the sea and saw a small boat arriving with three men on board" (MOE, 2021b)

employed the present participle verb "arriving" to describe a subsequent action that concurred with the completed actions "looked down" and "saw". The verbs "looked down" and "arriving" are regular, whereas the verb "saw" is irregular. The above examples underscore the significance of mastering regular and irregular past forms for grammatical accuracy.

Discussion

The results revealed a thoughtful integration of grammatical structures that catered to the learning needs of ESL learners in Malaysian primary schools. Each past tense form and the combination forms highlighted their foundational roles in ESL learners' language acquisition, fostering their grammatical mastery. This development can be understood through comprehensive frameworks of second language acquisition (Ellis, 2008) and the crucial role of language production in solidifying grammatical knowledge (Swain, 1985). While the simple past tense in the textbooks introduced ESL learners to articulate completed actions, personal experiences, past states, past habitual actions, and historical narratives (Azar & Hagen, 2009; Celce-Murcia & Larsen-Freeman, 1999), the past continuous tense allowed ESL learners to express progressive actions at a specific point in the past (Azar & Hagen, 2009; Quirk et al., 1985). The identified combination forms – the simple past tense and the past continuous tense, and the simple past tense and present participle - addressed the importance of integrating multiple past tenses within a single sentence in enhancing ESL learners' language skills and their ability to express ongoing and simultaneous actions, events, or states. These combination forms enabled ESL learners to construct coherent and engaging sentences or narratives (G`afforovna & Qizi, 2025; McCarthy & Carter, 1994).

These findings critically addressed RQ1, which investigated the representation of English past tense forms in CEFR-aligned English textbooks for Malaysian primary schools. More than merely identifying grammatical forms, this analysis highlighted how these forms are systematically embedded in meaningful, pedagogically purposeful contexts. In doing so, this study contributed to bridging the gap between prescriptive textbook content and functional language acquisition, a dimension often underexplored in prior research on ESL materials in the Malaysian context. This study's findings critically confirmed that CEFR-aligned textbooks function not only as repositories of grammatical forms but as pedagogical tools that scaffold learners' interlanguage development (Ellis, 2009). By providing structured exposure to both regular and irregular past tense forms, these textbooks operationalise the acquisition order hypothesis (Goldschneider & DeKeyser, 2001), making the abstract process of language acquisition visible within tangible classroom practices.

The thoughtful integration of grammatical structures – the past tense forms – within the textbooks suggested that effective language instruction should accompany the teaching. Additionally, including the past tense forms spotlighted the importance of contextualised learning in English language lessons. For instance, the story writing lessons in contexts – past experiences or events – emphasise the simple past tense to enhance the narrations. Similarly, English teachers can incorporate other past tense forms and their combinations, as shown by the themes, to further demonstrate diverse usage of these forms, allowing ESL learners to apply them in various contexts. Furthermore, English teachers can facilitate a deeper connection between ESL learners' understanding of the past tense forms and their learning productions (sentence writing) to ensure that the language learning is more relevant and engaging. This well-rounded approach allows ESL learners to discover the interconnection between the past tense forms to express verb tenses –

completed, ongoing, or simultaneous actions in the past. This enhances ESL learners' grammatical mastery and promotes critical thinking and creativity. In addition, themes F3 and F4 demonstrated complex grammar constructions, thus preparing ESL learners for advanced language use. The teaching, with a gradual progression from simple to compound and complex sentences, empowered ESL learners' abilities and skills to produce compelling, well-structured, and meaningful written pieces.

Critically, this study extended prior works (e.g., Bardovi-Harlig, 2000; VanPatten, 2015) by demonstrating that Malaysian ESL learners benefit from systematic exposure to grammar forms embedded in purposeful contexts, rather than decontextualised rote memorisation. Such contextualised grammar instruction aligns with CEFR principles of functional language use, contributing to ESL learners' long-term retention and transferability of skills to communicative tasks beyond the classroom (Chin, 2018; Bayuong & Arip, 2024).

The analysis of CEFR-aligned textbooks underscored the crucial role of past tense forms in ESL learning. Three themes emerged from the study: To Engage in Communication Authentically (DF1), To Chronicle Past Events Accurately (DF2), and To Use Verbs in Past Form Correctly (DF3) – empowering accurate application of grammatical rules for regular and irregular past tense forms by ESL learners. These directive functions are pivotal for developing ESL learners' practical and contextual understanding of the past tense. English teachers can benefit from these directives to enhance their grammar instruction, thus uplifting the quality of ESL teaching and learning. Bayuong and Arip (2024) and Zulkiply and Nasri (2024) provided further evidence for the effectiveness of CEFR-aligned materials in promoting a holistic understanding of language functions.

These findings addressed RQ2, which sought to examine the directive functions of English past tense forms in ESL learning at Malaysian primary schools. The three directive functions reflected how past tense forms were employed as grammatical features and essential tools for achieving communicative competence (Hassan & Gao, 2021). This approach equips learners with the necessary skills for genuine interaction (Segar & Asmawi, 2024). These findings confirmed that CEFR-aligned textbooks operationalised grammatical instruction beyond rote grammar learning, promoting purposeful and context-driven language use. What emerged from this study was not merely a descriptive account of textbook content, but a model of grammar pedagogy grounded in the CEFR's communicative competence framework. The findings challenged traditional forms-focused instruction, advocating for integrative approaches that bridge grammatical accuracy with fluency development (Canale & Swain, 1980; Neupane, 2024).

Expanding the above discussion, the analysis from this study underscored significant implications not only limited to the school level but also for higher education and broader educational domains. For instance, in higher education, integrating foundational grammatical structures, such as past tense forms, spotlights the importance of embedding contextual and functional grammar teaching across disciplines. As this continues, ESL learners can possess the linguistic competence required to excel in academic discourse and professional communication at a higher education level (Yazid et al., 2024). Next, teaching methodologies that emphasise authentic contexts can be extended to higher education to help ESL learners develop the necessary skills to handle diverse language-related contexts. Besides, aligning language modules with practical grammar applications in developing a curriculum, which reflects the CEFR's framework, is essential to ensuring relevance and progression across all levels of education (Zulkiply & Nasri, 2024). Furthermore, accentuating the importance of coherent grammar instruction holistically

bolsters the interrelationships between primary, secondary, and tertiary levels of education in cultivating lifelong language competence. Beyond immediate classroom contexts, the study advocated for curriculum planners and policymakers to embed CEFR-aligned grammar instruction systematically across educational stages. This recommendation aligns with Aziz (2017) and Sani (2018), who emphasised the need for contextualised, progressive grammar instruction to foster a national English language proficiency benchmark compatible with global standards.

By addressing RQ1, this study advanced the understanding of how CEFR-aligned English textbooks operationalise grammatical learning within communicative language teaching frameworks. This contribution is particularly significant for Malaysian ESL contexts, where prior research often prioritised general curriculum analysis over granular linguistic content. As such, these findings provided a foundation for refining textbook development, teacher training, and curriculum design in alignment with international standards and local pedagogical needs.

The study has certain limitations, particularly in its limited exploration of sentence structures and its focus solely on past tense forms. While the analysis effectively addressed the research questions, the selected examples may not fully represent the diversity of tense usage in CEFR-aligned English textbooks for Malaysian primary schools. Future studies should address these gaps by examining a broader range of sentence structures and incorporating additional verb tenses. This expanded scope would provide a more comprehensive understanding of tense usage and support more robust findings for ESL grammar pedagogy within the Malaysian context. Another notable limitation is the absence of triangulation with classroom-based empirical data, such as learner production or teacher feedback. Incorporating such data in future studies could validate the textbook content analysis with real-world learning outcomes and enable a more robust evaluation of how the identified past tense forms and directive functions (RQ2) translate into actual language use by ESL learners.

Conclusion

The thematic analysis spotlighted the various applications of past tense forms within CEFR-aligned English Textbooks for Malaysian primary schools. The sentences extracted from the textbooks evidenced that the grammatical rules concerning English past tenses are not taught in isolation but are incorporated within rich contexts that primarily promote linguistic competence.

The directive functions highlighted some pedagogical implications for ESL teaching and learning, particularly in Malaysian primary schools. The thematic analysis emphasised the crucial need for transitioning ESL instruction from simply memorising the grammar rules to a more communicative and context-driven approach. English teachers should design teaching and learning activities that encourage ESL learners to engage in authentic contexts to foster confidence and fluency and reinforce their grammatical accuracy. By emphasising all three directive functions in teaching, which should align with the CEFR framework, English teachers can help their ESL learners develop language competency.

Moreover, the findings indicated the need for continuous and interrelated grammar instruction approaches in the curriculum at all levels of education. Embedding CEFR in the curriculum ensures consistency and coherence in grammar instruction. English teachers or educators at all educational levels can prepare ESL learners for academic success and communicative competency by incorporating past-tense forms within meaningful contexts.

Funding: This research is not funded.
Acknowledgements: The authors would like to thank the Ministry of Education, Malaysia, for the permission to conduct this study.
Conflicts of Interest: The authors declare no conflict of interest.
Authenticity: This manuscript is an original work.
Artificial Intelligence Statement: AI and AI-assisted technologies were not used.

About the Authors:

Patrick Duffy Bayuong received his Bachelor of Teaching (Hons) in TESL from the Institute of Teacher Education Malaysia and a Master of Education in TESL from the National University of Malaysia. Since 2016, he has taught English as a Second Language (ESL) at the primary level in Malaysian schools. His scholarly contributions include authoring research articles and disseminating his work at academic conferences, where he has also served as a session chair. His research interests focus on ESL education, with an emphasis on primary-level pedagogy in rural contexts. ORCiD ID: https://orcid.org/0000-0002-1648-755X

Dr. Harwati Hashim is an Associate Professor and Deputy Dean of Research & Innovation at the Faculty of Education, Universiti Kebangsaan Malaysia (UKM). She specialises in Innovative Pedagogy for ESL Teaching, with a focus on Mobile-Assisted Language Learning (MALL) and language learning strategies. Recognised for integrating technology into English language instruction, she has published widely in her field and actively contributes to advancing language education. ORCiD ID: https://orcid.org/0000-0001-9709-6261

Prof Dr Melor Md Yunus is a professor of English language education at the Innovative Pedagogies Center for Teaching and Learning Innovation, Faculty of Education, Universiti Kebangsaan Malaysia (UKM). She is widely recognised for her pioneering work in integrating ICT into English as a Second Language (ESL) teaching and learning, Technology-Enhanced Language Learning (TELL) and for advancing innovative pedagogies. She was ranked among the Top 2% of Scientists in 2024 by Elsevier. ORCiD ID: https://orcid.org/0000-0001-7504-7143

References

Almunawaroh, N. F., Trilestari, K. & Riani, A. (2021). Developing an E-book with Past Tenses Materials for the Students of English Education Study Program in the Indonesian Context. *J-SHMIC: Journal of English for Academic, 8*(1): 64-74.

Ambridge, B., Pine, J. M., & Lieven, E. V. (2014). Child Language Acquisition: Why Universal Grammar Doesn't Help. *Language, 90*(3): 53-90.

Angelovska, T. (2017). Beyond Instructed L2 Grammar Acquisition: Theoretical Insights and Pedagogical Considerations About The Role of Prior Language Knowledge. *Studies in Second Language Learning and Teaching, 7*(3): 397–417. https://doi.org/10.14746/ssllt.2017.7.3.3

Apte, M. (2004). Teaching The Past Tense Verbs: A New Approach to Teaching Verbs. *ABAC Journal, 24*(3): 15-19.

Ayoun, D. (Ed.). (2015). *The Acquisition of the Present*. Amsterdam: John Benjamins Publishing Company.

Azar, B. S. & Hagen, S. A. (2009). *Understanding and Using English Grammar* (4th ed.). Pearson Education.

Aziz, H. (December 6, 2017). Textbooks Do Not Make the Teachers. *New Straits Times*. https://www.nst.com.my/opinion/columnists/2017/12/311230/textbooks-do-not-make-teacher

Aziz, N. H. A. (2007). *ESL Students' Perspectives on Language Anxiety* (Doctoral dissertation, University Putra Malaysia).

Azlan, M. A. K., Shaharuddin, W. Y. W., Berhanuddin, Q. F. & Baharuddin, K. H. (2015). Five Most Common Grammatical Errors in Narrative Essay. *In the International Conference on Islamic Education and Social Entrepreneurship 2015.* 12-14 October 2015, Langkawi, Malaysia.

Bakir, A. B. M. & Aziz, A. A. (2022). Primary School Teachers' Perceptions on The Effectiveness of CEFR Textbooks in Malaysia: A Systematic Literature Review. *International Journal of Academic Research in Progressive Education and Development, 11*(3): 1401-1415. DOI:10.6007/IJARPED/v11-i3/14800

Bardovi-Harlig, K. (2000). *Tense and Aspect in Second Language Acquisition: Form, Meaning, and Use.* Wiley-Blackwell.

Bayuong, P. D. & Arip, A. F. (2024). Uncovering English Present Tense Constructs in CEFR-Aligned English Language Textbooks for Malaysian Primary Schools. *Library Progress International*, *44*(4): 1474-1483. https://doi.org/10.48165/bapas.2024.44.2.1

Bayuong, P. D. & Hashim, H. (2023). CEFR-aligned Primary Curriculum Framework: Rural National Primary School English Language Teachers' Attitudes and Slants. *Malaysian Journal of ELT Research, 20*(1): 1-20. https://doi.org/10.52696/TTAZ3243

Berko, J. (1958). The Child's Learning of English Morphology. *WORD,* 14:2-3, 150-177. DOI:10.1080/00437956.1958.11659661

Bowen, G. (2009). Document Analysis as a Qualitative Research Method. *Qualitative Research Journal*, *9*(2): 27-40. DOI:10.3316/QRJ0902027

Bukit, H. (2020). The Error Analysis in Using Tenses Made by Students in English Teaching and Learning Process. *JETLi Journal of English Teaching and Linguistics, 1*(2): 92-101. DOI:10.55616/jetli.v1i2.21

Canale, M., & Swain, M. (1980). *Approaches to Communicative Competence.* Singapore: Seameo Regional Language Centre.

Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The Grammar Book: An ESL/EFL Teacher's Course* (2nd ed.). Heinle & Heinle.

Chan, M. (2013). *Acquisition of Be by Cantonese ESL Learners in Hong Kong and Its Pedagogical Implications*. Bern: Peter Lang.

Chin, C. (March 25th, 2018). Coming Together for the Sake of Better English. *The Star*.

Cowan, R. (2008). *The Teachers Grammar of English.* Cambridge University Press.

Chomsky, N. (1999). *Derivation by Phase.* MIT, Department of Linguistics.

Chomsky, N. (2002). *Syntactic Structures.* Walter de Gruyter.

Declerck, R., Reed, S. & Cappelle, B. (2006). *The Grammar of the English Verb Phrase. Volume 1: The Grammar of the English Tense System, A Comprehensive Analysis.*

Ellis, R. (2008). *The Study of Second Language Acquisition* (2nd ed.). Oxford: Oxford University Press.

Ellis, R. (2009). The Differential Effects of Three Types of Task Planning on The Fluency, Complexity, and Accuracy of L2 Speech. *Applied Linguistics, 30*(4), 474–509. DOI:10.1093/applin/amp042

Fadl, S. A. E. (2019). A Consideration of The Materials for Teaching Past Tenses in English Grammar Course Books. *PEOPLE: International Journal of Social Sciences, 5*(3): 297-318. DOI: https://dx.doi.org/10.20319/pijss.2019.53.297318

Fanny, V. F. R., Juhansar, Amalia, U. & Prasetyoaji, A. (2022). Investigating EFL Students' Tense Errors in Writing Narrative Text. *International Journal of Business Humanities Education and Social Sciences (IJBHES), 4*(2): 53-63. DOI:10.46923/ijbhes.v4i2.172

Fereday, J. & Muir-Cochrane, E. (2006). Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. *International Journal of Qualitative Methods, 5*(1): 80–92. https://doi.org/10.1177/160940690600500107

G`afforovna, Y. Z. & Qizi, M. R. N. (2025). The Role of Past Simple Tense and Past Continuous Tense in Teaching for EFL Students. *American Journal of Education and Learning, 3*(3): 1120-1124. https://doi.org/10.5281/zenodo.15105077

Gill, S. K. (2002). *International Communication: English Language Challenges for Malaysia.* Serdang: Universiti Putra Malaysia Press.

Goldschneider, J. M., & DeKeyser, R. M. (2001). Explaining the ''Natural Order of L2 Morpheme Acquisition'' in English: A Meta-analysis of Multiple Determinants. *Language Learning, 51(*1): 1-50. DOI:10.1111/1467-9922.00147

Green, B. (2023). Five Design Principles for Language Learning Materials Development. *ORTESOL Journal, 40*: 4-20. https://eric.ed.gov/?id=EJ1402220

Greenbaum, S. & Nelson, G. (2002). *An Introduction to English Grammar* (2nd ed.). London: Longman.

Handayani, N. D. & Johan, M. (2018). Correlation Between Problem Faced in Grammar and Writing Ability of EFL University Students. *Journal of English Education, 4*(2): 108-120.

Hashim, H., Rafiq, K. R. M. & Yunus, M. M. (2019). Improving ESL Learners' Grammar with Gamified-Learning. *Arab World English Journal (AWEJ) Special Issue on CALL* (5). 41-50. DOI: https://dx.doi.org/10.24093/awej/call5.4

Hassan, K. I. & Gao, X. (2021). Communicative Language Teaching in Malaysian ESL Context: A Qualitative Exploration into In-service Teachers' Beliefs and Practices. *International Journal of TESOL Studies, 3*(1): 58-72. https://doi.org/10.46451/ijts.2021.01.05

Housen, A., & Simoens, H. (2016). Introduction: Cognitive Perspectives on Difficulty and Complexity in L2 Acquisition. *Studies in Second Language Acquisition, 38*(2): 163-175. doi:10.1017/S0272263116000176

Javed, M. & Ahmad, A. (2013). Investigating ESL Students' Academic Performance in Tenses. *i-manager's Journal on English Language Teaching, 3*(4): 21-26. DOI:10.26634/JELT.3.4.2518

Jia, G., & Fuse, A. (2007). Acquisition of English Grammatical Morphology by Native Mandarin-Speaking Children and Adolescents: Age-Related Differences. *Journal of Speech, Language, and Hearing Research, 50*: 1280–1299.

Kovac, M. & Alenka, K. M. (2022). The Changing Role of Textbooks in Primary Education in the Digital Era: What Can We Learn from Reading Research? *CEPS Journal, 12*(2): 11-27. doi: 10.26529/cepsj.1290

Larsen-Freeman, D. (2003). *Teaching Language: From Grammar to Grammaring.* Boston: Thomson/Heinle.

Larsen-Freeman, D., Kuehn, T. & Haccius, M. (2011). Helping Students Make Appropriate English Verb Tense-Aspect Choices. *TESL Journal, 11*(4): 3-9. https://doi.org/10.1002/j.1949-3533.2002.tb00102.x

Lawyer, H. (January 9, 2024). *The Pros and Cons of Digital Textbooks vs. Traditional Textbooks.* Maranatha Baptist University. https://www.mbu.edu/blog/student-interests/pros-cons-digital-traditional-textbooks/

Lionny, G. P. & Kusumadewi, H. (2022). An Error Analysis on the Use of Simple Past Tense in Students' Recount Text Writing. *JEdu Journal of English Education, 2*(1): 32-39. DOI:10.30998/jedu.v2i1.6428

Lista, R. & Febriyanti, E. R. (2020). EFL Learners' Problems in Using Tenses: An Insight for Grammar Teaching. *IJET (Indonesian Journal of English Teaching), 9*(1): 86-95. DOI:10.15642/ijet2.2020.9.1.86-95

Manokaran, J., Ramalingam, C. & Adriana, K. (2013). A Corpus-Based Study on the Use of Past Tense Auxiliary 'Be' in Argumentative Essays of Malaysian ESL Learners. *English Language Teaching, 6*(10): 111-119.

McArthur, T. (2002). *The Oxford Guide to World English.* Oxford/New York: Oxford University Press.

McCarthy, M., & Carter, R. (1994). *Language as Discourse: Perspectives for Language Teaching.* Longman.

Ministry of Education Malaysia. (2015). *English Language Education Reform in Malaysia: The Roadmap 2015-2025.* Putrajaya: English Language Standards and Quality Council.

Ministry of Education Malaysia. (2018). *Get Smart Plus 3 Student's Book (Malaysian Edition).* Putrajaya: Ministry of Education Malaysia.

Ministry of Education Malaysia. (2019a). *English Language Management Guidebook: Primary and Secondary School.* Putrajaya: The Inspectorate, Ministry of Education Malaysia.

Ministry of Education Malaysia. (2019b). *Get Smart Plus 4 Student's Book (Malaysian Edition).* Putrajaya: Ministry of Education Malaysia.

Ministry of Education Malaysia. (2021a). *English Plus 1 Second Edition Year 5 Student's Book (Malaysian Edition).* Putrajaya: Ministry of Education Malaysia.

Ministry of Education Malaysia. (2021b). *Academy Stars 5 Year 6 Pupil's Book (Malaysian Edition).* Putrajaya: Ministry of Education Malaysia.

Neupane, B. (2024). Effectiveness of Integrated Skills Approach in Improving English Language Proficiency in Community Schools in Lamjung. *English Language Teaching Perspectives, 9*(1-2): 77-91. DOI: https://doi.org/10.3126/eltp.v9i1-2.68722

Nijat, N., Karimi, A. Q., Monib, W. K., Hisham, M. H. M. & Rahman, N. F. A. (2020). Enhancing Learners' Grammar Knowledge of Tenses and Voice Through "Tense and Voice Structured Wheel". *International Journal of Mechanical and Production Engineering Research and Development (IJMPERD), 10*(3): 14381-14396.

Özçelik, Ö. & Kent, A. K. (2023). *Designing Effective Language Learning Materials for Less Commonly Taught Languages.* Georgetown University Press.

Parrott, M. (2000). *Grammar for English Language Teachers* (2nd ed.). United Kingdom: Cambridge University Press.

Pinker, S. and M. T. Ullman (2002). The Past and Future of the Past Tense. *Trends in Cognitive Sciences 6*(11): 456–463. https://doi.org/10.1016/S1364-6613(02)01990-3

Prasada, S. and S. Pinker (1993). Generalisation of Regular and Irregular Morphological Patterns. *Language and Cognitive Processes*, *8*(1): 1–56. https://doi.org/10.1080/01690969308406948.

Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. (1985). *A Comprehensive Grammar of the English language*. Longman.

Richards, J. C., & Renandya, W. A. (Eds.). (2002). *Methodology in Language Teaching: An Anthology of Current Practice.* Cambridge University Press. https://doi.org/10.1017/CBO9780511667190

Salaberry, M. R. (2000). The Acquisition of English Past Tense in an Instructional Setting. *System, 28*(1): 135-152.

Sani, R. (February 8th, 2018). Building Capacity to Increase English Proficiency. *New Straits Times*.

Sazali, N. A. S. M., Radzi, N. A. M., Rosaidi, N. A. & Razali, R. (2024). A Corpus-Based Study of Simple Past Tense Errors in Malaysian Learners' English Writing. *International Journal of Modern Education (IJMOE), 6* (21): 359-371. DOI: 10.35631/IJMOE.621026

Segar, T. & Asmawi, A. (2024). Exploring ESL Students' Use of Multimedia in An English-Speaking Classroom. *Malaysian Online Journal of Educational Sciences, 12*(3): 61-75. https://files.eric.ed.gov/fulltext/EJ1457245.pdf

Shak, M. S. Y., Albakari, I. S. M. A., Tahir, M. H. M. & Adam, M. H. M. (2021). The Use of Imported CEFR-aligned English Language Textbooks in Malaysian Schools: Issues and Concerns. *International Journal of Academic Research in Business and Social Sciences, 11*(9): 954-963. DOI:10.6007/IJARBSS/v11-i9/10875

Sidhu, G. K., Kaur, S., & Chi, L. J. (2018). CEFR-Aligned School-Based Assessment in the Malaysian Primary ESL Classroom. *Indonesian Journal of Applied Linguistics, 8*: 452-463.

Slabakova, R. (2015). Acquiring Temporal Meanings Without Tense Morphology: The Case of L2 Mandarin Chinese. *The Modern Language Journal, 99*(2): 283-307.

Swain, M. (1985). Communicative Competence: Some Roles of Comprehensible Input and Comprehensible Output in Its Development. In Gass, S. & Madden, C. (Eds.), *Input in Second Language Acquisition* (pp. 235-253). Rowley, MA: Newbury House.

Thirusanku, J. & Yunus, M. M. (2014). Status of English in Malaysia. *Asian Social Science, 10*(1): 254-260. doi:10.5539/ass.v10n14p254.

Tomlinson, B. (Ed.). (2011). *Materials Development in Language Teaching* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/9781139042789

VanPatten, B. (2015). Input Processing in Adult Second Language Acquisition. In B. VanPatten, & J. Williams (Eds.), *Theories in Second Language Acquisition: An Introduction* (pp. 113-134). New York, NY: Routledge.

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. DOI: https://dx.doi.org/10.24093/awej/vol10no1

Yazid, N. H., Sulaiman, N. A. & Hashim, H. (2024). A Systematic Literature Review of Web-Based Learning and Digital Pedagogies in Grammar Education (2015-2024). *International Journal of Academic Research in Business and Social Sciences*, *14*(9): 2524-2542. DOI:10.6007/IJARBSS/v14-i9/22858

Zulkiply, D. I. N. & Nasri, N. M. (2024). The Use of CEFR-Aligned Textbooks among English Language Teachers: A Systematic Literature Review. *International Journal of Academic Research in Progressive Education and Development*, *13*(2): 2226–6348. DOI:10.6007/IJARPED/v13-i2/21401

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.407-425
DOI: https://dx.doi.org/10.24093/awej/vol16no3.25

The Impact of Watching English Shows on Enhancing English Language Acquisition Among Undergraduate Students

Daianne S. Gloria
Department of English, Bulacan State University
Corresponding Author: daianne.gloria@bulsu.edu.ph

Elenita P. Galvez
Department of English, Bulacan State University
Email: elenita.galvez@bulsu.edu.ph

Received:06/06/2025 Accepted: 09/14/2025 Published: 09/20/2025

Abstract
This study examines the impact of watching English-language shows on undergraduate students' English language acquisition, specifically in listening comprehension, vocabulary enrichment, and communication skills. This is guided by the central research question: To what extent does exposure to English-language shows influence undergraduate students' acquisition of English, particularly in listening comprehension, vocabulary development, and communication skills? While previous research highlights the role of audiovisual media in language learning, limited studies focus on its direct influence on undergraduate learners' proficiency and confidence. Using an explanatory sequential mixed-methods design, this study collected quantitative data from 536 undergraduate students through surveys, followed by qualitative interviews with selected participants. The findings reveal that students perceive watching English shows as a valuable supplementary tool, particularly for improving vocabulary retention, comprehension, pronunciation, and cultural awareness. The quantitative results show high mean scores in vocabulary retention (M = 4.16, SD = 0.96) and comprehension improvement (M = 4.09, SD = 0.91), reinforcing the role of media exposure in language acquisition. However, the lowest mean score (M = 3.55, SD = 0.98) on confidence in speaking suggests that passive media consumption alone is insufficient for developing spoken fluency. Qualitative insights further support these findings, highlighting improved listening skills, exposure to native-like pronunciation, and increased motivation to learn, yet also underscoring challenges in active language production. Comparison with existing literature confirms that while audiovisual media enhance language learning, they should be integrated with interactive speaking activities and structured instruction to develop comprehensive communication skills. The findings of this research contribute to the ongoing discourse on multimedia-assisted language learning. This also provides practical recommendations for English as a Second Language (ESL) learners and educators to emphasize the need for a balanced approach that combines media exposure and active language practice.

Keywords: Audiovisual media, English Language Acquisition, English as a Second Language, listening comprehension, multimedia-assisted language learning, vocabulary learning, vocabulary enrichment
How to cite: Gloria, D.S., & Galvez, E.P. (2025). The Impact of Watching English Shows on Enhancing English Language Acquisition Among Undergraduate Students. *Arab World English Journal, 16*(3): 407-425. https://dx.doi.org/10.24093/awej/vol16no3.25

Introduction

In this time of globalization and internationalization of higher education, academic and professional success is highly dependent on proficiency in English. Learning effective communication in English is an integral part of being able to have better academic and professional opportunities. An outcome that second language learners feel is an integral part of the importance of effective language acquisition strategies because English has become the world-wide lingua franca for communication in professions such as education, business, law, information technology, etc. There are many methods to learn and acquire English language use, but one of the most well-known and utilized tools for second language learners is watching English shows like movies or series episodes. This study explored the influence of watching English shows on promoting English language acquisition in undergraduate students in the Philippines.

Theories of language acquisition imply that language learning is something that humans are innately equipped to do (Chomsky, 1965). That said, Krashen (1982) conceptualizes language acquisition to include both formal instruction and informal exposure to the target language. Exposure to authentic and purposeful language input will develop listening comprehension skills, vocabulary, and pronunciation. English TV shows are a way to obtain meaningful input. Viewing films and TV shows is are authentic materials that showcase real-world contexts and natural discourse that provides an engaging language experience of immersion for English as a Second Language (L2) learners (Sherman, 2003). In addition, watching English television provides L2 learners with opportunities for authentic native-like pronunciation and exposure to idiomatic usages and cultural nuances that may also assist in transitioning between instruction and practical language use (Vanderplank, 2016).

The ease of understanding and expressive appeal of English TV shows create an accompanying language learning tool for second language learners. The underlying study proposes that L2 learners' interactive media involvement boosts motivation and attention, which progresses retention and comprehension (Danan, 2004). In the same vein, TV shows with captions and subtitles are enabling tools that help L2 learners connect spoken words with their written texts and strengthen vocabulary achievement (Pujadas & Muñoz, 2019).

However, it remains to be investigated the evidence-based effect of watching English television shows as an additional source on the language acquisition of L2 learners. This study investigates the impact of watching English television shows on how it improves undergraduates' listening comprehension and vocabulary acquisition. The findings will offer significant implications for language teachers and their students seeking quick, accessible, and efficient devices for learning the subtleties of the English language as a second language.

The Philippines, as an archipelago, is a multilingual country where English is one of the official languages and the primary medium of instruction in higher education. Many Filipino students consider English as one of the primary skills needed for academic success, career advancement, and global competitiveness. Given the popularity of English in the media, watching English shows has become a widely popular informal method of improving language proficiency among Filipinos. This policy aligns with the country's bilingual education mandate, which highlights English as a second language (ESL) for learners (Department of Education, 2016).

Given the long exposure of Filipino language learners to the English language, acquiring advanced English language proficiency is still a challenge, as they are heavily influenced by their native languages, such as Filipino and other regional languages. Despite this challenge, Filipinos are still exposed to English TV shows, and this immersive language-learning environment helps bridge that gap. This is reinforced by copious studies on the effectiveness of audiovisual media in familiarizing language learners with the realistic use of language of which comprises natural

pronunciation, colloquial expressions, and cultural contexts (Pujadas & Muñoz, 2019). The critical components of English proficiency enhance Filipino language learners' listening comprehension and vocabulary acquisition (Martin, 2019).

Moreover, the value of subtitles while viewing TV shows has highlighted the importance of a tool used by many learners of a language. These subtitles allow a language learner to make a visual connection to the spoken and written forms of English while also reinforcing grammar, syntax, and word usage (Danan, 2004). Subtitled footage is prevalent in the homes of the Philippines, and many language learners utilize media as an additional learning resource due to the lack of educational support to provide formal language teaching (Yadav, 2021). Therefore, the added use of subtitles while viewing aids language development and supports engagement and confidence motivation for language learners to transfer skills in both academic and real-world situations.

Situating this practice within the educational system of the Philippines opens a window for us to look into how watching English shows improves students' proficiency in English. Thus, we wish to delineate ways informal methods of learning could support and bolster formal methods of language education. To this end, this study aspires to enhance current literature about language acquisition by examining how specific language-related activities like watching English television shows might nurture undergraduate students' skills in listening comprehension and vocabulary development. In summary, this study aims to explore: How do English television shows influence undergraduate students' listening comprehension and vocabulary development in English?

Literature Review

Multimedia's primary function in language acquisition and learning is well documented. Research considers the effectiveness of multimedia when promoting language skills and mastery of the target language. The research has already been completed to investigate the effects of exposure to audiovisual media on listening comprehension, vocabulary knowledge, and pronunciation, as well as learners' motivation in learning the language. This section of the relevant research, which identified significant variables, indicates various ways in which multimedia can affect second language learners' English proficiency.

English Language Acquisition Through Multimedia Exposure

The impact of multimedia on the English language and second-language acquisition of learners is considerable. Research indicates that multimedia exposure can positively affect an L2 learner's listening comprehension, vocabulary development, and general language ability (Peters & Webb, 2018). Krashen's input hypothesis (1982) suggested that language acquisition takes place when learners are exposed to comprehensible input which is supported by a more recent study by Vanderplank (2016) showing that watching English language television programs increased learners' listening comprehension and supporting Webb & Rodgers (2009) argument that television series contain naturalistic and representative use of informal language and that audiovisual materials significantly contribute to language acquisition and development. However, the claim of simply engaging in passive exposure is insufficient; De Wilde et al. (2020) asserted that interactive involvement in the media contributes to and enhances the quality of exposure with the acquisition process.

Listening Comprehension and Pronunciation Development

A key aspect of language learning, such as any instance of listening comprehension and audiovisual sources, has been shown to help (Graham, 2006). Vanderplank (2016) noted that students learn to decode native speech through films or television, allowing them to process pronunciation, intonation, and rhythm. Ludke et al. (2014) showed that music instruction significantly improved pronunciation and lent credence to the idea that diverse multimodal content is beneficial to phonetic progress. That said, audiovisual resources can be supportive of listening skills as a whole, but Rosell-Aguilar (2017) reported that learners must interact with the content to reap improvements and understanding.

Vocabulary Acquisition Through Subtitles and Contextual Learning

In her contextualized research on vocabulary learning, Wang (2012) argues that subtitles are a productive way of assisting learners in vocabulary learning. Montero Perez and Peters, and Desmet (2013) researched the use of subtitles and found that learners who were exposed to vocabulary and showed vocabulary acquisition through subtitled programs had better retention than those learners who were not exposed to subtitled programs. The use of subtitles works best when you consider the different types of subtitles; bilingual subtitles support beginner learners, while monolingual (English) subtitles support intermediate to advanced learners (Dörnyei & Ryan, 2015). Rodgers and Webb (2020) further argue that exposing learners to a variety of genres, such as documentaries and dramas, allows learners to develop their lexical knowledge with contextualized language input.

Motivation and Engagement in Language Learning

Motivation is an important factor in language acquisition and learning. The use of audiovisual media increases learner engagement. Danan (2004) suggests that audiovisual media increase learner engagement. Many studies show that the students involved in extramural English do better than students who do not, because they are more motivated and greater level of English language usage, whether in speaking or writing. Sundqvist and Sylvén (2016) note that learners who adopt their extramural practices show more proficiency than those who do not. Furthermore, Krashen (2004) notes how engaging in reading and listening leads to language development and eventual retention of that language over time. On the contrary, Cobb (2007) reminds L2 learners about the importance of reinforcement—there can be no benefit from passive media exposure without engaging in the active processes of speaking or writing right afterwards.

Correspondingly, the use of multimedia has been very popular for language learners of English. It is commonly agreed upon in literature that multimedia can assist and complement developing listening comprehension, pronunciation, and vocabulary retention. However, the response to multimedia has been limited to largely Western or relatively privileged educational contexts, which raises concerns over the lack of empirically valid evidence from the educational context in the Philippines, where English language learning is a second language and where English is used as a medium of instruction. In addition, existing studies have shown that the more actively involved with multimedia, the more Helston the benefits. Ongoing research is examining ways that watching English language television shows can complement formal instruction and lead to measurable benefits for listening comprehension and vocabulary. With this as the backdrop, the research problem guiding this study originates from the lack of context-relevant evidence of English television shows as an informal and highly potent tool that contributes to second language

development in higher education institutions throughout the country.

Methods

This scholarly work makes use of the explanatory sequential mixed methods designed to investigate the impact of watching English shows on enhancing the English language acquisition among undergraduate students. The search design incorporates the following methods such as: first, a quantitative survey to gather the students' perspectives on the impact of watching English shows on learning English, and qualitative approaches to provide a comprehensive appreciation of the problem. The qualitative part will be administered through interviews with students to delve deeper into the findings.

Participants

The participants will come from 536 undergraduate students from a state university in the Philippines. They were chosen randomly based on their viewing preferences for English TV shows, and they participated in the quantitative phase. To ensure a different exemplification of viewpoints on how English TV shows affect their second language acquisition, these students were selected from a variety of educational disciplines. Responses highlight 14 undergraduate students were chosen randomly for the qualitative phase, ensuring a mix of respondents who expressed both positive and negative experiences in watching English TV shows. This group gave in-depth explanations of their distinctive learning strategies, challenges, and experiences.

Instruments

A validated research study made a survey questionnaire that was administered to 536 undergraduate students to explore their perspectives on how watching English shows contributes to their English language acquisition. The Likert-scale items were designed to assess various aspects of second language learning, such as listening comprehension, vocabulary development, pronunciation improvement, and overall motivation for learning English. The survey data were analyzed using descriptive statistics, particularly the mean and standard deviation, to identify trends and variations in student responses.

Procedures

The two-phase explanatory sequential mixed-methods approach was utilized in the data gathering. The qualitative phase was administered utilizing semi-structured interviews with undergraduate students. The students were randomly selected from the survey respondents to ensure a mix of perspectives based on their survey responses. These interviews focused on exploring the students' detailed experiences in watching English shows, including how it impacted their listening, vocabulary, and communication skills, as well as any challenges they encountered. The qualitative data were analyzed through thematic analysis to uncover recurring patterns and themes that complemented and explained the survey findings.

The sequential approach used allowed the study to integrate quantitative and qualitative data, providing a comprehensive understanding of the impact of watching English shows in enhancing English language acquisition among undergraduate students.

Results

This section presents the results to capture students' perceptions of how English shows impact their vocabulary, listening comprehension, pronunciation, speaking confidence, and overall language proficiency. The integration of statistical analysis and thematic insights offers a comprehensive understanding of how media exposure supports second-language learning in informal settings.

Table 1. *Perceived impact of watching English shows on enhancing their language proficiency*

Statement	Mean	Standard Deviation	Verbal Interpretation
1. Watching English-language shows has improved my overall English proficiency.	4.09	0.91	Agree
2. I believe that watching English shows helps me understand complex English grammar.	4.04	0.96	Agree
3. Watching English shows has enhanced my ability to speak and communicate in English.	3.98	0.97	Agree
4. My confidence in using English has increased due to exposure to English shows.	3.55	0.98	Agree
5. I have learned new vocabulary by watching English-language shows.	4.16	0.96	Agree
6. Watching English shows has improved my English pronunciation.	4.08	0.95	Agree
7. I can follow conversations in English more easily now.	4.09	0.96	Agree
8. Watching English shows has helped me understand cultural nuances in the English language.	4.06	0.97	Agree
9. I often pick up new words or phrases when watching English shows.	4.10	0.93	Agree
10. My comprehension of English texts and spoken language has improved due to watching English shows.	3.98	0.96	Agree

In Table 1, it is detailed how undergraduate students perceive watching English TV shows as facilitating specific aspects of their language skills. The mean scores of the respondents on all items reveal that they were generally in agreement that viewing English TV shows led to

improvements in their L2 skills in the context of TV. The standard deviations of the responses indicate the degree of consistency of the responses. The article claiming that viewing English TV shows has aided students in learning new words had the highest mean score (M=4.16, SD = 0.96). This implies that watching TV shows is thought to have the biggest impact on vocabulary acquisition out of all the components of language proficiency. This result is consistent with earlier research showing that incidental vocabulary learning is improved by exposure to media-rich content (Peters & Webb, 2018). Conversely, the lowest mean score (M = 3.55, SD = 0.98) pertains to the increase in confidence in using English due to exposure to English shows. While still within the "Agree" range, this lower score suggests that while watching English shows may aid in passive skills (listening and comprehension), it may be less effective in actively boosting confidence in spoken communication. The result, anchored on the study, indicates that passive exposure alone is inadequate in adopting speaking fluency without real interaction or productive language use (De Wilde et.al., 2020).

Table 2. *Perception of English as a language learning tool*

Statement	Mean	Standard Deviation	Verbal Interpretation
1.I watch English-language shows frequently (e.g., several times a week).	3.56	1.04	Agree
2. I prefer to watch English-language shows that include subtitles in my native language.	2.65	1.34	Neither Agree/Nor Disagree
3. Watching English movies or series helps me improve my language skills more than watching short videos or clips.	3.69	0.94	Agree
4.The type of English I watch (e.g., documentaries, dramas, comedies) affects how much I learn from them	4.00	0.96	Agree
5. Watching a variety of English shows (e.g., news, films, TV series) enhances my English language skills.	3.98	0.95	Agree

6. I view watching English-language shows as a fun and effective way to improve my English.	4.03	0.97	Agree
7. I believe that watching English shows is as helpful as formal ESL classes for improving language skills.	3.87	0.98	Agree
8. I find it easier to remember vocabulary or expressions when I hear them in English shows.	3.89	0.96	Agree
9 Watching English-language shows is a practical way to practice my English outside of the classroom.	3.97	0.94	Agree
10. I would recommend watching English shows as a learning tool for other ESL students.	4.01	0.96	Agree

The quantitative results in Table 2 display the students' perceptions of English-language shows as a relevant tool for language learning. The averages indicate that students from the survey mostly agree that viewing English shows a positive contribution to their language acquisition and learning, albeit with some variations in their responses. The highest average, M = 4.03, SD = 0.97, focuses on the fact that students enjoy watching English shows and that this is an effective way to improve English proficiency, which reinforces that enjoyable and fun learning experiences have the potential to improve motivation and engagement. In a similar manner, the statement that different types of English show an impact on learning outcomes has a high average (M = 4.00, SD = 0.96), which communicates that showing students different genres of English leads to improved language acquisition, which fits previous studies supporting the advantages of media diversity in language learning (Peters & Webb, 2018).

Additionally, the lowest average, M = 2.65, SD = 1.34, mirrors the choice of subtitles in one's first language, communicating a neutral stand amongst respondents. The high standard deviation for this item also generally communicates a greater variation in opinion, which could likely be attributed to some of the respondents' lower language proficiency or preferred learning style. This contradicts the findings of Wang (2012), that lower-level learners are helped more by native-language subtitles, while more advanced learners prefer English-language subtitles or no

subtitles. One other item with a relatively high standard deviation (SD = 1.04) is the frequency of viewing English shows (M = 3.56), indicating that respondents' viewing habits vary significantly, perhaps due to time availability, interest, or exposure to media in English.

Qualitative Results

Table 3. *Impact of watching English shows on the students' ability to communicate in English*

Quotes from Interview Data	Sub-Themes	Emergent Themes
"Watching English shows helps me learn by hearing how native speakers talk. It improves my listening and understanding of every language."	Listening Skills	Improved Comprehension
"Watching English shows has significantly improved my ability to communicate by exposing me to natural dialogue and cultural nuances that enhance my understanding & fluency."	Cultural Context Awareness	
"I feel empowered by the vocabulary I've heard; I learned new words in the movie or something that I watched."	Acquisition of New Words	Enhanced Vocabulary
"By watching, I get to learn expressions, phrases, and sentences in real language use."	Use of Idiomatic Expressions	
"I feel that I can now communicate to any native people that speak the language".	Increased Willingness to Speak	Boosted Confidence
"I feel confident in how to construct sentences with better grammar and communicate with the correct pronunciation of words".		

ISSN: 2229-9327

"I am very happy as I can apply the skills daily".	Positive Self-Perception	
"It feels nice because I always learn new words, and it improves my comprehension".		
"It helped me a lot to say the proper pronunciation of words".	Pronunciation and Intonation	Better Communication
"I noticed a great improvement in my verbal & non-verbal communication".	Contextual Communication	
"I get to copy the accent and the choice of words, which is different from Filipino."		

The qualitative findings provide greater insight into how watching English-language programs affects the students' communication development. The emerging themes delineated the improvements in comprehension, vocabulary acquisition, pronunciation, confidence, and communication in context. These were promising results that echo past research that discussed the importance of media exposure on second-language acquisition (Peters & Webb, 2018).

Table 4. *Suggestions for ESL learners to improve their English language skills*

Quotes from Interview Data	Sub-Themes	Emergent Themes
"Watch and watch because you will adopt it for sure, just like watching a kdrama annyeong". "I advise you to watch it according to your viewing preferences, as the more movies or television shows you watch, the more you will become accustomed to the way the English language is used".	Exposure to Natural Conversation	Watch English Shows
"If watching English shows is boring to you, it's best to start watching English-language shows that have your interests, whether it be	Genre-Specific Learning	

comedy, dramas, memes, conspiracy theories, documentaries, podcasts, and review content on things that you like. You can also start by listening to English songs". "They should enjoy what they are watching and not watch shows that are out of their preferred genre because when we are entertained, we can improve more easily". "Find genres or explore what types of topics they're interested in; it will help them to focus more on the show and understand what it is all about".		
"I suggest that they watch the English media with English subtitles so that they can easily understand and learn the correct spelling and pronunciations of the words they learn". "My little suggestion for them is to take this as an opportunity to learn and broaden their vocabulary just by watching English media". f one is having difficulties in watching and just listening to the English shows, it's best to put on English subs as well to other language shows, to further help understand and improve yourself". "Watch Filipino movies with English subtitles	Vocabulary Reinforcement Bilingual Subtitles	Use Subtitles

if you are a beginner. It will be a great help".		
"Listening to English songs will let you learn new words and idiomatic expressions". "When you listen to songs, you unconsciously learn the language as well, be it English or other languages".	Passive Learning Through Music	Listening to English Songs and Podcasts
"One great way to improve your English is by regularly watching shows, listening to podcasts, or reading articles in English, while actively trying to understand and mimic what you hear or read". "You may also listen to English podcasts, it's fun listening to them, especially if their topic is one of your interests". "My suggestions are to read books whether it is for education or as part of their leisure. Books are the best medium or tools to improve their English Proficiency. Shows or movies are the secondary medium to learn English but nothing beats the capabilities of books or written literature in terms of nurturing their vocabulary".	Thematic Podcast Selection Fiction for Vocabulary Growth	Listening to English Songs and Podcasts Read Books

"To enhance their English language skills aside from watching television, students can read books, articles, research papers, and other English-language resources. This will allow them to expand their vocabulary and practice fully understanding words better".		
"Aside from watching, you can read a book or immerse yourself in English language news and more practice in English language".	Non-Fiction for Contextual Understanding	

Table 4 displays the many different recommendations that English as a Second Language (ESL) learners can use to develop their English language skills through key subthemes and emergent themes drawn from the qualitative interviews. These themes demonstrate all the different strategies suggested by students to improve their English proficiency, which is evident in the literature regarding language acquisition.

Discussion

The perceived effects of watching English shows are to improve their language learning. The standard deviation values in the table ranged from 0.91 to 0.98, indicating a moderate range of variation in the responses of undergraduate students. The lowest standard deviation of SD = 0.91 for "Watching English-language indicates has improved my overall English proficiency" suggests that there was more agreement among respondents and, therefore, most respondents had similar perceptions across the factors. The higher standard deviation, SD=0.98 for "My confidence in using English has increased because of watching English shows," suggests more variation in how respondents perceived these effects. This variation could relate to individual differences in their experiences with language learning exposure and experiences, previous proficiency levels, and their confidence in their language use.

Other research studies have supported these findings. Webb and Rodgers (2009) found that watching television series in English increases vocabulary retention and listening comprehension. They noted that speaking requires active practice and so cannot simply be improved through passive exposure. Sundqvist and Sylvén (2016) similarly found that exposure to English language entertainment was positively correlated to vocabulary development, but does not result in improved grammar or speaking fluency without other practice. In contrast to the above studies, there are studies, including Krashen (1989), that suggest the use of comprehensible input (e.g.;

"watching English shows") can lead to language acquisition. However, the present study showed that students, while agreeing that watching English shows is beneficial to their learning, had a lower score in the factors related to confidence, indicating that a more interactive learning component is needed to supplement the learning that occurs with passive exposure.

Research in contexts within the Middle East provided similar impressions. Alghamdi (2021) investigated the informal digital learning practices of Saudi EFL learners and found that while students drew broadly on media and digital sources to enhance their vocabulary, they did not productively use language, especially speaking natively in English. Similarly, research by Alzahrani (2023) on Saudi learners using Netflix as a source of language input evidenced substantial improvements in comprehension and vocabulary through the use of visual input, but still emphasized the importance of structured speaking practice in order to improve oral fluency. These findings were evident in this study, with students finding passive engagement helpful but indicating lower confidence in speaking.

Overall, the quantitative results show that watching English shows improves various aspects of language learning, particularly vocabulary acquisition, comprehension, and pronunciation; however, the lower mean on confidence and the wider range of responses suggest that passive engagements or communications are not adequate for developing speaking fluency. The results also mirror previous studies, which suggest a need for interactive and immersive second language learning practices together with exposure to media.

The students' perception of using English as a language learning resource corroborates the existing literature noted above, which has historically based its conclusions on studies that showed watching English-language shows leads to increased vocabulary acquisition and listening skills. Webb and Rodgers (2009) consistently found that exposure to English television shows can improve word retention abilities and listening skills. Sundqvist and Sylvén (2016) also found that exposure to extramural English, such as watching television and/or movies in English, positively correlated with proficiency in the language, including vocabulary and oral skills. On the other hand, both previous studies suggested that media exposure could result in an additional resource to a formal education, rather than implying that instructional time is not supplemented. The findings in this study are in alignment with Webb and Rodgers (2009) and Sundqvist and Sylvén (2016), where watching shows in English was perceived as useful, but not more useful than taking formal ESL classes (M = 3.87, SD = 0.98).

In conclusion, the findings demonstrated that students perceive watching television shows in English positively and enjoyably; however, preferred subtitled viewing and frequency of viewing varied among students, and demonstrated that the use of English-language shows as a learning resource would ultimately depend on the learner factors. Taken together, these results suggest that while media exposure can be a helpful resource for learners, but should be augmented and used strategically with other language-learning approaches.

Impact of watching English shows on the students' ability to communicate in English

The qualitative data highlight important themes. The first theme is better understanding. Many student participants indicated that watching English shows helped them better understand and comprehend spoken language. This finding aligns with the studies of Vanderplank (2016), which showed that watching television in L2 improves listening and comprehension via authentic language input. The data suggests that cultural awareness is essential to language learning. Several participants emphasized that watching English shows provided them with cultural awareness, which is similar to the research of Sydorenko (2010). In that study, the author argued that

multimodal input (visual, auditory, and contextual) helps learners access and understand the cultural and linguistic features of the target language.

The second theme is when students indicated that they learned new words and idiomatic expressions with repeated exposure to the target language. Webb and Rodgers (2009) noted that watching English TV shows leads to incidental vocabulary learning, particularly when L2 learners engage with a variety of genres. Watching shows with subtitles was frequently mentioned as an additional learning tool, which extends Wang (2012), who found that subtitles provide extra support for vocabulary learning and retention in addition to enhancing pronunciation skills. However, students expressed conflicting opinions about the effectiveness of bilingual and monolingual subtitles, which illustrates the individual differences and personal preferences in learning.

Another important finding in the qualitative findings was the improved confidence of L2 learners who regularly watch English shows. Several participants showed a willingness to speak in English, which supported Sundqvist and Sylvén (2016) study that extracurricular exposure to English positively contributes to speaking confidence and overall fluency. Participants emphasized the improved ability to form grammatically correct sentences, with proper pronunciation, which added to Boers et al.'s (2017) argument that repeated exposure to spoken language in media develops pronunciation and intonation accuracy.

Besides watching television and movies, participants recommended listening to music and podcasts in order to develop their English proficiency. This supports Dörnyei and Ryan's (2015) assertion that engaging with English media in different modalities increases exposure and motivates learning. The findings indicated that students also recommended reading books in addition to watching English shows, which supports Krashen's (2004) importance of extensive reading in acquiring language skills.

Suggestions for ESL learners to improve their English language skills

Exposure to Natural Conversation: Watch English Shows

A significant sub-theme is "Exposure to Natural Conversation", where students highlighted the merit of watching English-language shows. The findings are aligned with research that supports exposure to authentic spoken English improves listening skills and fluency (Webb & Rodgers, 2009). Vanderplank (2016) agrees that watching television programs can help learners to process spoken language more naturally. However, it is important to recognize that while students receive valuable input by watching shows, they need to engage more actively through interaction to develop productive language skills (De Wilde et.al, 2020).

Generally Specific Learning

Second subtheme "Genre-Specific Learning" states, "language learners will view content that they find interesting to continue engaged and to have meaningful outcomes" is stated to be comparable to Peters & Webb (2018), who go on to indicate: "language learners acquire vocabulary more successfully when viewing content that they value". Similarly, Rodgers & Webb (2020) focused on the same interests that acknowledging different genres also highlight varying levels of vocabulary when learners will also select documentaries as containing greater lexical density in comparison to dramas or comedies related to their interests, where… Even though variety has an important role in motivating learners, Nation (2022) posits that repeated exposure to different genres is imperative for vocabulary development.

Vocabulary Reinforcement: Use Subtitles

An additional large subtheme is "Vocabulary Reinforcement," and indicated that many of the participants were in favor of using English subtitles as a vocabulary and understanding tool. Research shows that subtitles can help connect spoken and written forms of language, as students are learning to recognize words and spell (Wang, 2012). Dörnyei and Ryan (2015) also posit that captions support comprehension as they provide extra contextual information that helps remember new vocabulary. However, the relationship between monolingual and bilingual subtitles is still a debate that needs further research. Montero-Perez et al. (2013) have examined whether bilingual subtitles help beginner learners, and they found that unilingual English subtitles outperformed bilingual ones for the improvement of listening and vocabulary skills for intermediate and advanced learners.

Listening to English Songs and Podcasts

The next subtheme, "Listening to English Songs and Podcasts," clinches the role of music and audio content in language achievement. A number of participants recommended listening to podcasts and music to support language exposure, a strategy supported by research on auditory learning. Ludke et.al. (2014) introduced that musical input boosts pronunciation, rhythm, and intonation. In the same way, Rosell-Aguilar (2017) reported that language learners using podcasts proved improved listening comprehension and vocabulary retention. Despite these improvements, Graham (2006) summarizes that passive listening alone is insufficient; active engagement is required, such as taking notes or repeating phrases to develop the learning process.

Reading Books for Language Development

The final subtheme, "Read Books," covers the salient role of reading broadly in developing knowledge of English. Krashen (2004) views extensive reading as an important source of vocabulary learning, in disagreement that readers form better comprehension skills and grammatical accuracy if they read continually. Participants also recommended both fiction and non-fiction readings, consistent with research from Nation and Yamamoto (2012) that distinguishes that fiction books focus on narrative forms and idioms, and non-fiction books highlight academic and technical vocabulary. Nevertheless, Cobb (2007) notes that reading alone is insufficient and indicates that if readers engage in speaking and writing activities in addition to reading, they could make progress toward language acquisition more holistically.

Table 4 provides results that support existing knowledge on strategies for ESL learning by indicating that multi-media exposure, interactive experiences, and exposure to multiple source inputs are crucial to language acquisition and development. While passive exposure to English media validates vocabulary and comprehension, research indicates the best outcomes from using several strategies, like active reading, speaking practice, and focused listening. Future research should explore the long-term effects of strategies based on different levels of proficiency that lead to best practices for ESL learning.

In summary, the qualitative results indicate that consuming English content through shows is an effective and fun way to learn. The results corroborate previous studies that highlighted the importance of multimedia exposure for language acquisition and development. Conversely, the results also demonstrate that individual learning preferences are nuanced, and the most beneficial experience might be a convergence of a number of learning methodologies, such as reading, listening to podcasts, and using subtitles.

Conclusion

This study emphasizes that watching English-language shows has a considerable impact on undergraduates' English proficiency. The survey results indicate that English-language media exposure has a positive effect on vocabulary acquisition, listening comprehension, pronunciation, and cultural awareness. Also, the quantitative results indicate a high level of agreement on the effectiveness of this method, especially considering vocabulary retention and comprehension. Still, the lower scores from the survey item in relation to speaking confidence suggest passive consumption of media is inadequate for proficiency in active language production, as well. The qualitative results also support these findings because while students indicated they improved in terms of understanding natural dialogue and pronunciation, they still faced difficulties in their active production of the language. In consideration of prior research, these findings affirm the pedagogical value of audiovisual media and highlight the need to concentrate on structured language development strategies. The study overall indicates that English shows can motivate language acquisition, but to promote the acquisition of all the language skills, they should be followed with interactive and engaging activities.

Funding: Not Applicable

Acknowledgments: The authors wish to acknowledge the Bulacan State University behind this project for their support and for funding this study.

Conflicts of Interest: To ensure fair and objective decision-making, authors declared no conflict of interest in connection with manuscripts. Non-financial competing interests include a declaration of political, personal, religious, ideological, academic, and intellectual competing interests. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Authenticity: We hereby declare that this research is my original work and has not been submitted, in whole or in part, for publication or academic credit in any other institution. All sources and references used in this study have been properly acknowledged.

Artificial Intelligence Statement: This research did not involve the use of artificial intelligence (AI) tools in the generation of data, analysis, or interpretation of results. However, AI-assisted software (e.g., Grammarly, ChatGPT) was used solely for language refinement and improvement of grammar and structure in the preparation of this manuscript. The authors confirm that all intellectual contributions, analyses, and conclusions are their own.

About the Authors:

Daianne S. Gloria, Ph.D., is an Assistant Professor IV in the Department of English of the College of Arts of Letters at Bulacan State University. She is currently designated Director of the Student Program Development of the Student Policy and Program Development Office under the Office of the Vice President for Academic Affairs. Her research interests include English language education, second language teaching, educational leadership and management, inclusive education, and curriculum and instruction. ORCID: https://orcid.org/0009-0000-2266-6231

Elenita P. Galvez, Ph.D., is a Professor IV in the Department of English of the College of Arts of Letters at Bulacan State University for more than 30 years. She is a member of Professional Organizations such as the Asian Qualitative Research Association, British Council, PAPSAS, and several esteemed organizations.

Her expertise lies in identifying and overcoming barriers to effective communication, fostering a collaborative environment, and leveraging perception to improve workplace relationships. ORCID: https://orcid.org/0009-0008-9307-7651

References

Alghamdi, E. A. (2021). Informal digital learning of English vocabulary: Saudi EFL learners' attitudes and practices. *Arab World English Journal (AWEJ)*, *12*(3), 394–408. https://doi.org/10.24093/awej/vol12no3.27

Alzahrani, A. H. (2023). A prospective study for exploring Saudi EFL learners' strategic listening skills through Netflix from teachers' perspectives. *Arab World English Journal (AWEJ)*, *14*(4), 97–113. https://doi.org/10.24093/awej/vol14no4.7

Boers, F., Dang, T. C., & Strong, B. (2017). Comparing the effectiveness of phrase-focused exercises: A partial replication of Boers, Demecheleer, Coxhead, and Webb (2014). Language Teaching Research, 21(3), 362–380. https://doi.org/10.1177/1362168816651464

Chomsky, N. (1965). Aspects of the theory of syntax. MIT Press.

Cobb, T. (2007). Computing the vocabulary demands of L2 reading. Language Learning & Technology, 11(3), 38–63. https://doi.org/10125/44174

Danan, M. (2004). Captioning and subtitling: Undervalued language learning strategies. Meta: Journal des traducteurs, 49(1), 67–77. https://doi.org/10.7202/009021ar

De Wilde, V., Brysbaert, M., & Eyckmans, J. (2020). Learning English through out-of-school exposure: Which levels of language proficiency are attained and which types of input are important? Bilingualism: Language and Cognition, 23(1), 171–185. https://doi.org/10.1017/S1366728918001062

Department of Education. (2016). Bilingual education policy in the Philippines. DepEd.

Dörnyei, Z., & Ryan, S. (2015). The psychology of the language learner revisited. Routledge. https://doi.org/10.4324/9781315779553

Graham, S. (2006). Listening comprehension: The learners' perspective. System, 34(2), 165–182. https://doi.org/10.1016/j.system.2005.11.001

Krashen, S. D. (1982). Principles and practice in second language acquisition. Pergamon.

Krashen, S. D. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. The Modern Language Journal, 73(4), 440–464. https://doi.org/10.2307/326879

Krashen, S. D. (2004). The power of reading: Insights from the research (2nd ed.). Heinemann.

Ludke, K. M., Ferreira, F., & Overy, K. (2014). Singing can facilitate foreign language learning. Memory & Cognition, 42(1), 41–52. https://doi.org/10.3758/s13421-013-0342-5

Martin, I. P. (2019). English language education in the Philippines: Policies, problems, and prospects. In X. Gao (Ed.), Second handbook of English language teaching (pp. 621–637). Springer. https://doi.org/10.1007/978-3-030-02899-2_30

Montero Perez, M., Peters, E., & Desmet, P. (2013). Is less more? Effectiveness and perceived usefulness of keyword and full captions in second language learning. Computer Assisted Language Learning, 26(2), 91–113. https://doi.org/10.1080/09588221.2011.652508

Nation, P., & Yamamoto, A. (2012). Applying the four strands to language learning. International Journal of Innovation in Language Learning and Teaching, 6(2), 113–125.

https://doi.org/10.26686/wgtn.12552020.v1
Peters, E., & Webb, S. (2018). Incidental vocabulary acquisition through viewing L2 television and the role of captions. Studies in Second Language Acquisition, 40(3), 551–577. https://doi.org/10.1017/S0272263117000407
Pujadas, G., & Muñoz, C. (2019). The effect of input enhancement on learning vocabulary from audiovisual input. System, 81, 38–50. https://doi.org/10.1016/j.system.2019.01.005
Rodgers, M. P. H., & Webb, S. (2020). Vocabulary demand in television programs. Language Learning, 70(1), 165–197. https://doi.org/10.1111/lang.12372
Rosell-Aguilar, F. (2017). Podcasting in language learning. Language Learning & Technology, 21(3), 66–89. https://doi.org/10.14705/rpnet.2015.000265
Sherman, J. (2003). Using authentic video in the language classroom. Cambridge University Press.
Sundqvist, P., & Sylvén, L. K. (2016). Extramural English in teaching and learning. Springer. https://doi.org/10.1057/978-1-137-46048-6
Sydorenko, T. (2010). Modality of input and vocabulary acquisition. Language Learning & Technology, 14(2), 50–73. https://doi.org/10125/44214
Vanderplank, R. (2016). Captioned media in foreign language learning and teaching: Subtitles for the deaf and hard-of-hearing as tools for language learning. Palgrave Macmillan. https://doi.org/10.1057/9781137500458
Wang, Y. (2012). Learning L2 vocabulary with different captioning types: A study of Chinese ESL learners. *Computer Assisted Language Learning, 25*(5), 365–382. https://doi.org/10.5539/elt.v5n8p217
Webb, S., & Rodgers, M. P. H. (2009). Vocabulary demands of television programs. Language Learning, 59(2), 335–366. https://doi.org/10.1111/j.1467-9922.2009.00509.x
Yadav, M. S. (2021). Role of social media in English language learning to the adult learners. *International Journal of Linguistics, Literature & Translation, 4*(1), 25–33. https://doi.org/10.32996/ijllt.2021.4.1.25

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp. 426-440
DOI: https://dx.doi.org/10.24093/awej/vol16no3.26

Investigating the Role of AI Tool Interactions in Enhancing English Language Acquisition Among Saudi College Students

Amnah Abdulrahman Alhamam
Early Childhood Education, Department of Teacher Education
The Patton College of Education
Ohio University. Athens, Ohio 45701
USA
Email: amnhalhumama@gmail.com

Received:06/29/2025 Accepted: 08/30/3035 Published: 09/20/2025

Abstract
This study examines how Saudi college students learning English as a second language utilize artificial intelligence tools, particularly ChatGPT, to enhance language acquisition. Traditional grammar translation still dominates English instruction in Saudi Arabia, but artificial intelligence chatbots offer personalized, interactive, and engaging second language learning opportunities. While traditional grammar-translation methods dominate English instruction in Saudi Arabia, artificial intelligence chatbots provide new opportunities for personalized, interactive, and engaging second language learning. The study uses Krashen's Input Hypothesis, Vygotsky's Sociocultural Theory, and Communicative Language Teaching principles to develop a qualitative design with purposive sampling and semi-structured interviews of ten college students. Thematic analysis revealed three major findings: (1) students demonstrated limited artigificial intelligence literacy and institutional support for these tools in language learning; (2) students used ChatGPT primarily for writing assistance, brainstorming, and translation, often without critical reflection or feedback revision; and (3) attitudes toward ChatGPT ranged from optimism about its benefits to concerns about over-reliance and lack of trust. While artificial intelligence tools promise to enhance second language acquisition, the study highlights key barriers, including a lack of training, pedagogical integration, and human interaction. The findings call for curriculum reform, teacher training, and the development of culturally responsive strategies to integrate artificial intelligence effectively. This research contributes localized insights into the role of artificial intelligence in non-Western EFL contexts and offers recommendations for future policy and practice.

Keywords: Artificial intelligence, ChatGPT, English as a foreign language, language learning technologies, second language acquisition, Saudi Arabia
How to cite: Alhamam, A.A. (2025). Investigating the Role of AI Tool Interactions in Enhancing English Language Acquisition Among Saudi College Students. *Arab World English Journal, 16*(3): 426-440. https://dx.doi.org/10.24093/awej/vol16no3.26

Introduction

Artificial Intelligence (AI), specifically AI chatbots, emerged with a promise to transform Second Language (L2) learning and acquisition. Traditional language learning methods – grammar-translation and audio-lingual – emphasize structured learning, grammar rules, and vocabulary memorization through rote learning and repetitive exercises, but they lack the sustained motivation and active engagement to optimize L2 learning (Anjum et al., 2024). Additionally, most students consider L2 learning uninteresting, compromising their motivation. Fortunately, the pioneer AI chatbot, ChatGPT, is a language model that has demonstrated considerable potential in enhancing L2 acquisition (Alhazmi & Muftah, 2025). AI tools promise to address the limitations of traditional teaching methods by providing a personalized learning environment and authentic opportunities for L2 practice and acquisition.

AI chatbots, such as ChatGPT, DeepSeek, and Gemini, promise to enhance students' motivation and learning engagement in L2. These tools execute diverse tasks, including text generation, machine translation, content correction, problem-solving, summarization, grammar correction, and answering questions (Kim, 2023). The chatbots respond to users' prompts with adaptive human-like text, enhancing linguistic literacy and boosting students' motivation and engagement by offering interactive and personalized learning experiences (Asad et al., 2024). However, how learners interact with AI chatbots to complete academic tasks and practice language remains underexplored in developing nations such as Saudi Arabia.

Saudi Arabia has gradually introduced AI language learning, but traditional grammar-translation and audio-lingual techniques are still the dominant teaching methods. The conventional methods adopt rote memorization, repetitive drills, and passing engagement with static textbooks (Alqahtani, 2022), which minimizes opportunities for interaction with the spoken language and partially addresses the diverse and evolving needs of SLA learners (Vireak & Bunrosy, 2024). Additionally, most studies on AI tools in L2 learning focus on the Western context, where most students are native English speakers and have practice opportunities (Al-Seghayer, 2024). AI tools can stimulate conversational practice but should be customized to capture diverse learner needs and enhance L2 acquisition.

This study explores how Saudi Arabian college-level second language (L2) learners interact with AI tools to complete academic tasks in English and enhance learning. Specifically, the research analyzes which AI applications they use in L2 learning, their interaction, and the impact on their L2 acquisition and English proficiency. The specific objectives are:

1. To investigate students' awareness of AI tools that can enhance L2 acquisition.
2. To examine how students use AI tools to complete their academic tasks in English
3. To explore student attitudes toward AI-powered tools for language learning.

Various studies have attempted to study the impact of AI in education, particularly in the English Foreign Language (EFL) domain. AI chatbots provide engaging human conversation as a mechanism of language practice. However, empirical studies in the EFL domains are scarce due to unresolved questions about technology's implementation in education (Du & Daniel, 2024). Alqaed (2024) claims that most research on AI adoption in education emerged from 2019 as a response to the pandemic to develop strategies for adopting technology in teaching. Such a focus might limit its specialization as a second-language teaching tool. Saudi Arabia has substantially incorporated AI into its development goals; however, it lacks comprehensive studies on the challenges and opportunities related to implementing AI-based learning outcomes (Alotaibi et al., 2022). The limitation worsens due to the moral and ethical challenges of adopting AI in education

and public service. Alsaawi (2025) posits that many institutions and organizations in Saudi Arabia follow strict ethical and moral guidelines, which might hinder AI adoption as a language learning tool. Ethical concern stems from potential AI biases, especially in language learning, where instructors are concerned about the credibility and validity of the knowledge and information. This disadvantage is worse among countries where English is a second language, such as Saudi Arabia. Most current studies focus on Western L2 learning contexts, which present several opportunities to practice English because of the availability of many native speakers (Asad et al., 2024; Alhazmi & Muftah, 2025; Anjum et al., 2024). Saudi Arabia needs localized studies that account for its domestic weaknesses, such as fewer native English speakers and fewer practice opportunities. Fortunately, Saudi Arabia has made significant efforts under the Vision 2030 to integrate AI in public institutions (Kumar et al., 2025). Saudi Arabia has considerably invested in AI to improve public service and governance, providing opportunities for its inclusion in second language learning. This research aimed to identify the barriers and facilitators of AI chatbot applications in EFL. The outcomes informed educators and policymakers on the technology's practicality, challenges, and necessary interventions to optimize its strengths and minimize weaknesses.

Literature Review

AI in Language Learning

Natural Language Processing

AI is an emerging technology that promises to transform language education through interactive tools that enhance teaching and learning. The critical AI feature is Natural Language Processing (NLP) technology, which combines computational linguistics and Machine Learning (ML) models to understand, interpret, and generate human language (Asad et al., 2024). This capability enables AI to create intelligent tutoring systems, automate writing evaluation tools, and personalize learning environments (Baskara, 2023). NLP and ML are the foundation of applications that adapt to a learner's needs and provide customized feedback and support, a core factor in second language (L2) learning (Alhazmi & Muftah, 2025). AI-driven chatbots go beyond language learning to simulate human-like conversations, enabling learners to practice and develop core language skills – including writing, reading, and comprehension.

NLP-Driven Chatbots

ChatGPT has received extensive scholarly and practitioner attention among AI-driven chatbots. Asad et al. (2023) claim that ChatGPT is the pioneering technology in natural language processing that understands and generates human-like conversations. While ChatGPT embodies the real-life implementation of NLP, alternatives include Deep Seek, a Chinese AI-driven NLP developed by Liang Wenfeng, Gemini by Google, and Copilot by Microsoft. The models accomplish various NLP tasks, such as sentiment analysis, question answering, and language translation (Alhusaiyan, 2025). The chatbots have similar core functionality but differ in pre-training methodology and architecture.

Theories of L2 Acquisition

Krashen's Input Hypothesis

AI chatbots offer adaptive, interactive, and context-rich learning experiences that align with Second Language Acquisition (SLA) theories. Krashen's Input hypothesis proposes that language acquisition occurs by exposing learners to a language input higher than their current

comprehension level (i+1) (Luo, 2024). AI chatbots' adaptive inputs match this hypothesis by generating responses tailored to learners' proficiency levels. For instance, a chatbot can infer learners' level through interaction to respond with simple sentences with basic vocabulary for beginners and complex structures for advanced learners. Xia et al. (2023) claim that Krashen's input hypothesis aligns with AI-based language learning systems, such as cross-cultural language learning, which dynamically adjust to the learners' comprehension capability and the language complexity. This principle is relevant to Saudi Arabia because it focuses on individual cultural context and dismisses the generalization perception in AI-based learning. AI chatbots also provide comprehensible feedback, rephrasing sentences to correct learners' mistakes at the i+1 level (Gong, 2023). Krashen's Input hypothesis affirms that AI chatbots support L2 language acquisition through adaptive input and comprehensible feedback features.

Vygotsky's Sociocultural Theory

Vygotsky's Sociocultural Theory (SCT) emphasizes that social interactions guide and mediate a child's cognitive development, learning ability, and language acquisition (Alkhudiry, 2022). Vygotsky claims that an individual will perform better when receiving help from others to complete a task than when working alone. SCT is a culturally relevant approach to learning as it accounts for diversity and distinct cultures, such as Saudi Arabia. Vygotsky argued that understanding cognition requires going beyond individual effort to include cultural and social factors that influence behaviors, interactions, and language (Chen, 2025). A person's language development outcome depends on social interaction and relationships, which should be contextualized in a specific culture, such as Saudi Arabia. SCT enables educators to connect a community's historical and cultural profile to AI adoption as a second-language learning tool. Furthermore, it aligns with AI-chatbots integration in language learning through scaffolding interaction, where AI serves as a knowledgeable friend and guides learners through tasks such as breaking a task into smaller activities and, at the same time, offering hints and also supports collaborative learning by co-creating stories or solving problems (Sarmiento-Campos et al., 2022). However, AI lacks sociability due to the absence of cues such as tone and gestures. Despite lacking socioemotional cues, SCT theoretically approves the chatbots' relevance in L2 acquisition through scaffolded interaction and collaborative learning.

Communicative Language Teaching

Communicative Language Teaching (CLT) is a language teaching theory that prioritizes developing communication competence through meaningful interaction and real-life language use. CLT perceives language as a system for expressing meaning. Its primary function is interaction and communication (Qasserras, 2023). Like CLT, using AI chatbots in L2 learning and acquisition provides authentic communication that simulates real-world scenarios and prioritizes fluency and meaningful exchange. The technologies facilitate interactive practices, such as role-play and debates, to develop pragmatic skills and personalize conversations by tailoring content to learners' interests, thereby boosting motivation and providing instant feedback (Qasserras, 2023). Unfortunately, text-based learning often lacks the cultural immersion necessary for effective language acquisition. CLT is a framework for exploring AI chatbots' authenticity, personalization, and interactiveness in L2 acquisition.

Benefits of AI-Chatbot in L2 Acquisition

Personalized Learning

AI chatbots personalize content by dynamically adjusting language complexity and topics based on user interaction. The adaptive feedback is consistent with Krashen's Input Hypothesis and underscores the significance of comprehensible (i+1) level input (Asad et al., 2024). Novice learners receive simple instruction and basic vocabulary, while advanced learners engage with complex structures and expressions. AI chatbots tailor interaction based on user input to allow them to learn at their own pace, which addresses gaps in the traditional one-size-fits-all approach to L2 learning (Anjum, 2024). Learners can also focus on areas of interest to enhance motivation and relevance to learning. Customizing content to learners' proficiency levels and topics improves linguistic competence and increases confidence by minimizing the anxiety associated with inflexible traditional learning approaches (AbuSahyon et al., 2023). AI chatbots create a student-centered environment that supports current pedagogical demands for differentiated instruction.

Interactive Practice

AI chatbot enhances language acquisition and communication competence through interactive language practice that mimics human conversation. It deviates from traditional static exercises by engaging learners in open-ended dialogue, role play, and problem-solving activities that reflect human interaction (Ghafar et al., 2023). For example, learners can practice interview questions and responses, discuss cultural topics, or order food in a restaurant, which assists them in bridging the gap between theory and practice. Such interaction affirms SCT, which emphasizes that effective learning occurs through social collaboration and scaffolding (Asad et al., 2024). AI chatbots serve as conversational agents, providing a supportive environment for learners to experiment with language, make and correct errors, and refine their language skills. AI generates contextually appropriate responses, encouraging critical thinking and adapting language to different contexts or roles. Interactive practice provides continuous engagement to improve fluency and develop pragmatic skills that apply to real-life communication.

Immediate Feedback

AI chatbots provide instant feedback, allowing students to correct errors and refine language use in real-time. Traditional static methods delay feedback. However, AI chatbots instantaneously identify grammatical inaccuracies and lexical or syntactic errors, including explanations for rewritten portions (Ghafar et al., 2023). Immediate feedback reinforces learning, improves knowledge retention, and reduces errors. Moreover, it encourages students to request clarification and direct learning to address their inadequacies or weaknesses (Asad et al., 2024). AI chatbots can speed learning, but they should be used moderately to avoid over-reliance on automated feedback that lacks sociocultural and human interactions.

Accessibility and Convenience

AI-based teaching overcomes geographic and time barriers. The technology is available 24/7 to anyone with internet connectivity. This capability empowers users to practice language at any time and place, eliminating traditional barriers associated with rigid classroom schedules (Kim et al., 2023). A practical example is an employed learner seeking to optimize their time by learning while commuting at home, where they cannot access a native speaker. The AI convenience aligns with global trends in educational technology that emphasize flexibility and inclusivity (Anjum et

al., 2024). AI chatbots have multilingual features that allow them to switch between native and second languages, require translations, or compare linguistic structures to improve cross-linguistic awareness. AI-conversational agents are also scalable, providing similar opportunities to institutional and individual learners (Asad et al., 2024). Nevertheless, concerns about the digital divide, availability of internet connection, or access to internet-enabled devices affect equitable technology adoption and learning.

Limitations of AI-Chatbot in L2 Acquisition

Lack of Human Interaction

AI chatbots adopt a conversational style but lack essential elements of human interactions, such as sarcasm, idioms, and cultural references, which are vital in advanced language proficiency. This limitation implies that the technology generates grammatically correct responses but cannot interpret context-dependent details necessary for pragmatic and sociolinguistic skills (Asad et al., 2024). The lack of cultural and emotional intelligence slows knowledge transfer into the real world. Zaghlool and Khasawneh (2023) argue that language mastery extends beyond syntax and vocabulary to encompass tone, body language, and cultural norms. These features are absent in IA-generated texts. The limitation can lead to superficial L2 acquisition, where learners excel in structured tasks but are weak in real-life scenarios. Human interaction with native speakers remains irreplaceable for modeling non-verbal communication and cultural nuances, which helps learners adapt their language use to diverse social contexts (Zaghlool & Khasawneh, 2023). Moreover, AI can lead to prompt dependency, which manifests as a pattern of obsessive thoughts, impaired control, and craving for AI-generated content more than learning. Although AI dependency is not a disorder, it raises concerns about institutional responsibility, academic integrity, and student autonomy, which are critical factors in learning (Revesai, 2025). Therefore, despite AI chatbots offering valuable real-time practice, the technology partially substitutes for the depth of human-led learning.

Risk of Over-Reliance

The availability and accessibility of AI chatbots lead to over-dependence, reducing the motivation for real-life conversations needed to develop pragmatic and interpersonal skills. Zaghlool and Khasawneh (2023) claim that learners who rely solely on AI chatbots for language practice become accustomed to controlled, error-tolerant interactions, making them ill-prepared to deal with the unpredictability of human dialogue. For instance, interpreting non-verbal cues or adjusting language politeness levels can be learned from face-to-face interaction with native speakers. Additionally, Zaghlool and Khasawneh (2023) claim that successful language acquisition depends on dialogue interaction marked by learners negotiating meaning and receiving spontaneous feedback, processes that AI chatbots cannot genuinely replicate. Furthermore, AI solutions lack inherent responsibility, as they are just tools, raising accuracy, compliance, and authentication concerns. AI hallucination occurs when an AI system generates false or misleading information and presents it as facts due to patterns in its training data, insufficient or biased data, or an inability to fill gaps in its knowledge (Sun et al., 2024). AI chatbot passive learning limits knowledge generalization, forcing instructors to complement AI-based learning with human-to-human interaction opportunities to promote holistic skill development.

Quality of Feedback

AI advancement has failed to address inconsistent, inaccurate, and inappropriate feedback that reinforces errors and confuses learners. The tools generate plausible-sounding but sometimes incorrect explanations, including misapplied grammar rules or outdated cultural information. Asad et al. (2024) claim that AI tools might provide conflicting advice on verb or tense usage to concretize errors. The immediate automated feedback system may focus on surface-level corrections, such as spelling, rather than addressing deeper linguistic and pragmatic issues, like coherence. AI chatbot-generated responses may lack pedagogical intentionality because it does not systematically track learners' progress or tailor feedback to individual developmental needs (Luo et al., 2024). The randomness in the content can frustrate learners and erode their trust in the tool's reliability. Educators emphasize the need for human oversight to validate AI-generated content and address gaps in contextual understanding, ensuring that learners receive accurate and pedagogically sound guidance.

Methodology

This study adopted the qualitative method to explore how Saudi students use AI chatbots to complete English language tasks. It evaluated awareness, usage, and attitudes toward AI-powered tools for language learning. The qualitative design was appropriate because it gathered comprehensive data from a small sample in their natural setting to understand the usage and impact of ChatGPT on language learning in a non-Western academic setting. In contrast, the qualitative method would only provide empirical findings without a deeper analysis of the unique cultural and religious factors shaping language education.

Participants

The study's targeted population was college students in Saudi Arabian colleges taking English as a second language (L2). It adopted purposive sampling because the recruitment depended on specific criteria to ensure sample diversity. The screening questionnaire inquired about age, first language, English proficiency level, and experience using ChatGPT. The selection criteria were: participants should be (1) aged 18-25 years; (2) Arabic as a first language; (3) English as a second language; (4) English proficiency levels intermediate to advanced determined through self-reported proficiency or university placement exams; (5) have prior experience using ChatGPT for academic English tasks, such as writing, editing, brainstorming, summarizing, or translating. The recruitment process began with inviting all L2 students through emails and social media (WhatsApp). Students who accepted the invitation completed the screening questionnaire (Appendix A), which was used to select twenty participants based on the eligibility criteria (twelve females and eight males). The participants completed semi-structured interviews through Zoom or Microsoft Teams.

Data Collection

Data collection employed a semi-structured interview to gather comprehensive contextual data from the five purposively sampled participants. The research questions were developed from literature reviews to address the knowledge gap. Also, a pilot study was performed with two colleagues taking English as a second language to review the questions and ensure they were clear and understandable. The final interview sheet had eight questions, consisting of opening (2), central (4), and closing (2) questions as listed in Appendix B. The interview process began with

all ten participants (six males and four females) receiving an email requesting that they select the best date, time, and channel (Zoom or FaceTime) for the interview. Every interviewee was informed of the research objectives, that the responses would be anonymous, and that any collected data would be secure. The interviewer informed the participants of their rights before asking them to sign an informed consent form. The interview lasted about 15 to 20 minutes, collecting audio and written data for analysis and interpretation.

Data Analysis

Data collected was analyzed using qualitative thematic analysis to explore the usage and impact of ChatGPT in L2 learning and acquisition. The analysis process adopts the Squires (2023) framework, consisting of six phases: data familiarization through reading and re-reading interview scripts, initial code generation, theme development, defining and refining, and report generation. It used NVivo to generate the codes, but it lacked inter-coder reliability due to the research scope and chosen methodology. The thematic analysis allows an in-depth understanding of how Saudi Arabian college-level students engage with ChatGPT in L2 learning.

Findings

Demographic Characteristics

The purposive sample of five college-level L2 students comprised twelve males and eight females from both science and arts disciplines and all years of study, as summarized in Table 1.

Table 1. *Demographic characteristics of the sample (N = 20)*

Demographics	Characteristics	Frequency (n)	Percentage (%)
Gender	Male	12	60
	Female	8	40
University Major	Computer Science &Information Technology	4	20
	Electrical and Mechanical Engineering	4	20
	Business Administration	4	20
	Islamic Studies and Sharia Law	4	20
	Medical Sciences	4	20
Year of Study	Freshman	4	20
	Sophomore	4	20
	Junior	4	20
	Senior	8	40

Thematic Analysis

Theme 1: Limited AI Literacy for L2 Learning

Table 2 summarizes the direct quotes from the first objective: to investigate students' awareness of AI tools that can enhance L2 acquisition. It shows the participants' limited awareness and integration of ChatGPT in L2 learning. Most respondents cited a lack of formal guidance in applying ChatGPT and claimed they were uncertain about the technology's future. One of the respondents stated, "*I tried ChatGPT once, but I did not know how to ask the right questions to get useful answers*," and another added, "*I did not know how to start using it for language practice. No one taught us*." Additionally, the curriculum still emphasizes textbooks and grammar drills, providing limited opportunities for the students to practice using AI tools. A respondent cited the

absence of formal adoption by stating, "*Our classes focus on textbooks, so AI feels unnecessary. Teachers do not encourage it."* Such statements highlight institutional resistance towards AI in L2 learning, which poses further challenges or barriers to students. These findings answer the research question and prove limited awareness, use, and integration of AI tools in L2 language curricula.

Table 2. *Supporting quotes for the theme awareness of AI tools for L2 learning*

Supporting Interview Quotes for Awareness of AI Tools for L2 Learning
• I tried ChatGPT once, but could not ask the right questions to get valuable answers. • I did not know how to start using it for language practice. No one taught us. • Our classes prioritize textbooks, so AI feels unnecessary. • I use AI twice weekly because our classes focus on textbooks and grammar drills. The AI chatbot feels unfamiliar and unnecessary. • The most common task is writing assistance, like translating thoughts from Arabic to English.

Theme 2: AI Tools Usage Patterns for L2 Learning

Table 3 summarizes interview responses about the second research objective: to examine how students use AI tools to complete their academic tasks in English. The interview responses suggest the varied utility of ChatGPT in daily writing tasks. One respondent indicated a sporadic engagement by claiming, "*I use ChatGPT daily to write assignments,*" while another showed a lack of familiarity in the statement, "*I tried ChatGPT once but did not know how to ask the right questions.*" Most students use ChatGPT for specific tasks, such as brainstorming or translating Arabic to English texts, but their usage lacks reflective engagement. These findings suggest diverse applications of ChatGPT, though most usage is superficial and lacks significant linguistic development.

Table 3. *Supporting quotes for the theme AI tools usage pattern*

Supporting Interview Quotes for AI Tools Usage Patterns
• I use ChatGPT daily to write assignments, but do not revise the output. • I use ChatGPT to brainstorm essay ideas, which helps me structure my thoughts better. • The most common task is writing assistance, like translating thoughts from Arabic to English. • I feel confident using it for brainstorming, but I am unsure if it improves my skills. • I tried ChatGPT once, but I did not know how to ask the right questions to get valuable answers.

Theme 3: Attitudes Toward AI Tools for L2 Learning

Table 4 summarizes direct quotes from the interview about the third research objective: to explore student attitudes toward AI-powered tools for language learning. The interviewees had mixed attitudes toward using ChatGPT for L2 learning. Some students were optimistic about the technology. For instance, one of the participants claimed, "*My writing improved significantly,*" while another criticized AI, stating, "*Overusing ChatGPT made me lazy.*" These comments also highlighted a distrust of AI-generated content, with some students preferring human feedback. Despite the recognized benefits of AI, most students questioned the pedagogical value of ChatGPT

in language learning, pointing out a gap between the use of AI chatbots and skill internalization. AI could improve language learning, but its features do not sufficiently provide the necessary support for quality academic work.

Table 4. *Supporting quotes for the theme attitudes toward AI tools for L2 learning*

Supporting Interview Quotes for Attitudes Toward AI Tools for L2 Learning
• My writing improved when I asked ChatGPT to explain and practice grammar rules. • ChatGPT helped me refine my essay vocabulary usage, which boosted my test scores. • It saves time, but I am not internalizing the language. I just copied the output. • Overusing ChatGPT for translations made me lazy. My speaking skills did not improve. • I do not trust ChatGPT's feedback. What if it is wrong? I prefer human corrections.

Discussion

Low Awareness and Institutional Barriers

AI chatbots are novel technologies whose application in education will face mixed reactions. Saudi Arabian students exemplify a group with limited awareness of AI's significance in learning, a situation common among EFL contexts where educational systems are similarly rooted in traditional, teacher-centered methodologies. The slow response probably stems from emphasizing traditional rote and memorization methods in the L2 education system. This challenge in Saudi Arabia reflects Vygotsky's Sociocultural Theory, which claims that a lack of institutional support in adopting a new technology discourages the structural assistance needed to facilitate the process (Alkhudiry, 2022). Furthermore, the problem is evident in the student's struggle to formulate the correct prompts. Low AI literacy is a knowledge gap widened by institutional practices that prefer textbooks to AI integration in learning (Al-Seghayer, 2024). Such challenges are common in EFL content and highlight the need to address learning disparities. Global trends advocate for AI integration in education, but institutional inertia in Saudi Arabia's English language classes slows AI adoption in learning.

Usage Patterns and Pedagogical Impact

Artificial intelligence has several capabilities and functionalities that address diverse needs. The varied usage patterns among Saudi Arabian students reflect AI's potential in language learning. The application of AI among L2 learners aligns with Krashen's Input Hypothesis, which states that adaptive input at the i+1 level increases the speed of language acquisition (Luo, 2024). Such applications benefit student learning and development; however, the passive use of AI tools, such as copying output, overlooks cognitive development and undermines deeper learning. This problem is common among countries where English is a second language and is a cautionary issue for classroom management. While AI can enhance learning, it can be a drawback for students who do not see its pedagogical value in L2 learning. Previous studies have also highlighted the need for cautionary application of AI in learning. For instance, Ghafar et al. (2023) claim AI chatbots provide a 24/7 interactive space for language practices, but Qasserras (2023) warns that unstructured use undermines the need for meaningful and context-rich interaction for effective L2 learning, as postulated in Communicative Language Teaching (CLT) theory. The research further suggests that learners using ChatGPT for rote translation of text from Arabic to English overlook CLT's emphasis on pragmatic skills such as negotiating meaning, which led to claims that AI does not improve speaking skills. AI applications among L2 learners in Saudi Arabia remain limited

and fail to effectively develop pragmatic language skills. Instructors in the EFL context need to establish a pedagogical framework to avoid using AI as a passive learning strategy and failing to develop essential communication skills. Educators and policymakers should review the technology to ensure that language teaching in Saudi Arabia technologically matches international trends.

Attitudes and Ethical Considerations

Every technological transformation elicits mixed reactions depending on its usability and people's readiness for change. Similarly, Saudi Arabia's college-level L2 learners exhibit mixed attitudes towards AI tools' benefits in personalized feedback and accessibility. The findings are consistent with Luo et al.'s (2024) claim that AI can improve grammatical performance and mastery of a second language, but it might produce inconsistent responses that lack pedagogical value and encourage overreliance on machine-generated texts. These mixed attitudes and concerns are shared across EFL countries. AI usage has also raised trust concerns, as some student prefer human assistance and feedback in their assignments. Students expressed concerns about the trustworthiness of AI-generated content and preferred human feedback. Therefore, AI tools can assist L2 students, but they require regulations to balance the instructor and chatbot's roles to ensure student language acquisition is reflected in theory and practice.

Conclusion

This research confirms that AI is a valuable teaching tool, particularly for students learning a second language. It highlights the potential and challenges of integrating AI into second language classes. The most critical conclusion is the need for a blend of technology, pedagogy, and culture to improve the use of AI tools in L2 learning. Although AI chatbots offer interactive and adaptive feedback for language practices, their effectiveness depends on structured integration into the current curricula. Saudi Arabia's textbook-centric teaching methods stifle the collaborative and interactive environment of AI. Optimizing AI usage in classrooms demands multistakeholder collaboration to address infrastructure barriers, curriculum content, pedagogical design, and AI literacy. AI is a valuable tool, but it should be used alongside traditional learning theory that accounts for culture and socio-emotional learning activities in language acquisition.

At the same time, making ChatGPTs and other AI tools work well in second language education requires bigger changes in how language is taught. It's not enough to introduce ChatGPT to students; it must be carefully connected to teaching methods focusing on communication, culture, and active learning. Schools and universities in Saudi Arabia should focus on improving and updating old curricula, training teachers and students to use ChatGPT effectively and ethically, and making sure lessons stay culturally appropriate. It is also important to note that AI should support, rather than replace, real conversations, critical thinking, and social interaction. Future research should continue to explore how to combine AI's benefits with traditional human-centered learning to help students become stronger and more confident language users. Future studies can also be implemented on a larger number of participants with different learning levels, as this study was limited to only 10 college-level students.

Funding: This paper 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: AI was moderately used to rephrase some setnces for clarity; all content was reviewed and verified by the author.

About the Author:

Amnah Alhamam is a Ph.D. student in Curriculum and Instruction with a focus on Early Childhood Education at Ohio University, Athens. Her research interests include early childhood language acquisition, literacy development, and the integration of artificial intelligence in education. She has professional experience in curriculum design and teacher professional development in Saudi Arabia. Her current work explores instructional practices that support children's language and social interaction within the Saudi educational context. ORCID: https://orcid.org/0009-0007-8196-1274

References

AbuSahyon, A. S. A. E., Alshorman, O., & Al-Absi, B. (2023). Investigating the impact of AI-driven chatbots on the acquisition of English as a foreign language among Saudi undergraduate students. *International Journal*, *10*(2), 3075-3088. https://doi.org/10.15379/ijmst.v10i2.3049

Alhazmi, A. A., & Muftah, M. (2025). Evaluating ChatGPT's reliability in second language acquisition (SLA): Insights on language skills and technology's role. *Electronic Journal of e-Learning*, *23*(1), 96-112. https://doi.org/10.34190/ejel.23.1.3762

Alhusaiyan, E. (2025). A systematic review of current trends in artificial intelligence in foreign language learning. *Saudi Journal of Language Studies*, *5*(1), 1-16. https://doi.org/10.1108/SJLS-07-2024-0039

Alkhudiry, R., 2022. The contribution of Vygotsky's Sociocultural Theory in mediating L2 knowledge co-construction. *Theory & Practice in Language Studies (TPLS), 12*(10), 2117 – 2123. https://doi.org/10.17507/tpls.1210.19

Alotaibi, N. S., & Alshehri, A. H. (2022). Prospers and obstacles in using artificial intelligence in Saudi Arabia's higher education institutions—The potential of AI-based learning outcomes. *Sustainability*, *15*(13), 10723. https://doi.org/10.3390/su151310723

Alqahtani, A. (2022). Review and analysis of theories underlying grammar teaching methodologies. *Arab World English Journal*, *13*(4), 80-91. https://doi.org/10.24093/awej/vol13no4.6

Alqaed, M. A. (2024). AI in English language learning: Saudi learners' perspectives and usage. *Advanced Education, 12*(25), 125-142. https://orcid.org/0000-0003-3012-2028

Al-Seghayer, K. (2024). Voices across the sands: A historical journey of English language teaching in Saudi Arabia. *British Journal of English Linguistics*, *12*(3), 24 – 45. https://doi.org/10.37745/bjel.2013/vol12n32445

Alsaawi, A. (2025). Challenges and opportunities associated with AI chatbots in language learning from the perspective of users in Saudi Arabia. *International Journal of Learning, Teaching and Educational Research, 24*(5), 400-415. https://doi.org/10.26803/ijlter.24.5.21

Anjum, F., Raheem, B. R., & Ghafar, Z. N. (2024). The impact of ChatGPT on enhancing students' motivation and learning engagement in second language acquisition: Insights from students. *Journal of E-learning Research, 3*(2), 1-11. https://doi.org/10.33422/jelr.v3i2.679

Asad, M., Faruqui, A., Zehra, S., Arif, M., Batool, W., & Siddiqui, M. A. (2024). The impact of AI tools (ChatGPT) on second language learning. *Policy Research Journal, 2*(4), 805-813. https://orcid.org/0009-0004-7273-5118

Baskara, R., 2023. Exploring the implications of ChatGPT for language learning in higher education. *Indonesian Journal of English Language Teaching and Applied Linguistics*, *7*(2), pp.343-358. https://eric.ed.gov/?id=EJ1391490

Chen, B. (2025). Using Vygotsky's sociocultural theory to explore ethnic cultural representation in Taiwanese preschool children's play. *Frontiers in Education*, *10*, 1 – 10. https://doi.org/10.3389/feduc.2025.1569322

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*. https://doi.org/10.1016/j.caeai.2024.100230

Ghafar, Z.N., Salh, H.F., Abdulrahim, M.A., Farxha, S.S., Arf, S.F., and Rahim, R.I., 2023. The role of artificial intelligence technology on English language learning: A literature review. *Canadian Journal of Language and Literature Studies*, *3*(2), 17-31. https://doi.org/10.53103/cjlls.v3i2.87

Gong, J., 2023, February. The concept, content, and implication of Krashen's input hypothesis. In *2022, the 4th International Conference on Literature, Art and Human Development (ICLAHD 2022)* (pp. 1208-1213). Atlantis Press. https://doi.org/10.2991/978-2-494069-97-8_154

Kim, S., Shim, J., & Shim, J. (2023). A study on the utilization of OpenAI ChatGPT as a second language learning tool. *Journal of Multimedia Information System*, *10*(1), 79-88. https://doi.org/10.33851/JMIS.2023.10.1.79

Kumar, R., Singh, A., Ahmed Kassar, A. S., Humaida, M. I., Joshi, S., & Sharma, M. (2025). Adoption challenges to artificial intelligence literacy in public healthcare: An evidence-based study in Saudi Arabia. *Frontiers in Public Health*, *13*, 1558772. https://doi.org/10.3389/fpubh.2025.1558772

Luo, Z. 2024. A review of Krashen's Input Theory. *Journal of Education, Humanities and Social Sciences*, *26*(1), 130-135. https://doi.org/10.54097/3fnf5786

Qasserras, L., 2023. Systematic review of communicative language teaching (CLT) in language education: A balanced perspective. *European Journal of Education and Pedagogy*, *4*(6), 17-23. https://doi.org/10.24018/ejedu.2023.4.6.763

Revesai, Z. (2025). Generative AI dependency: The emerging academic crisis and its impact on student performance—A case study of a university in Zimbabwe. *Cogent Education*, *12*(1), 1 – 33. https://doi.org/10.1080/2331186X.2025.2549787

Sarmiento-Campos, N.V., Lázaro-Guillermo, J.C., Silvera-Alarcón, E.N., Cuellar-Quispe, S., Huamán-Romaní, Y.L., Apaza, O.A. and Sorkheh, A., 2022. A look at Vygotsky's sociocultural theory (SCT): The effectiveness of scaffolding method on EFL learners' speaking achievement. *Education Research International*, *2022*(1). https://doi.org/10.1155/2022/3514892

Squires, V. (2023). Thematic analysis. In *Varieties of qualitative research methods: Selected contextual perspectives* (pp. 463-468). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-04394-9_72

Sun, Y., Sheng, D., Zhou, Z., & Wu, Y. (2024). AI hallucination: Towards a comprehensive classification of distorted information in artificial intelligence-generated content.

Humanities and Social Sciences Communications, *11*(1), 1-14. https://doi.org/10.1057/s41599-024-03811-x

Xia, Y., Shin, S., & Kim, J. (2023). Cross-cultural intelligent language learning system (CILS): Leveraging AI to facilitate language learning strategies in cross-cultural communication. *Applied Sciences*, *14*(13),1 – 27. https://doi.org/10.3390/app14135651

Zaghlool, Z.D. and Khasawneh, M.A.S., 2023. Incorporating the impacts and limitations of AI-driven feedback, evaluation, and real-time conversation tools in foreign language learning. *Migration Letters, 20*(7), 1071-1083.https://doi.org/10.59670/ml.v20i7.4863

Appendices

Appendix A

Screening Questionnaire

Section 1: Demographics

1. Age: ____
2. Gender: Male ☐ Female ☐ Prefer not to say ☐
3. Native Language (L1): ____
4. Second Language (L2): ____
5. University Major: ____
6. Current Year of Study: ☐ Freshman ☐ Sophomore ☐ Junior ☐ Senior

Section 2: English Proficiency Level

- How would you rate your English proficiency? Beginner ☐. Intermediate ☐. Advanced ☐. Near-native ☐.

7. Have you ever taken an English proficiency test (e.g., IELTS, TOEFL, STEP1)? ☐ Yes ☐No. If yes, what was your score? ____

Section 3: ChatGPT Usage

- How often do you use ChatGPT for academic purposes? ☐Daily. ☐Weekly. ☒Occasionally. ☐Never (End of survey – Thank you for your time)

10. What academic tasks do you use ChatGPT for? (Check all that apply)
    - Vocabulary Building ☐
    - Conversation Practice ☐
    - Writing Assistance ☐
    - Grammar Explanations ☐
    - Reading & Comprehension Support ☐
    - Other (please specify): ____
11. How confident are you in using ChatGPT for academic English tasks?
    - Not confident at all ☐
    - Somewhat confident ☐
    - Confident ☐
    - Very confident ☐

[1] STEP is the English Proficiency test (Standardized Test of English Proficiency) used in Saudi Arabia.

Appendix B
Semi-Structured Interview Guide

Opening Questions:

1. Can you describe how you first started using ChatGPT for academic tasks?
2. How often do you use ChatGPT, and for what specific academic purposes?

Main Questions:

1. In what ways do you feel ChatGPT has helped improve your English skills (e.g., writing, vocabulary, grammar)?
2. Have you encountered any challenges or limitations when using ChatGPT? If so, what are they?
3. Do you think using ChatGPT has made you more dependent on AI for language tasks? Why or why not?
4. Have you noticed any changes in writing or academic performance since using ChatGPT?

Closing Questions:

1. Would you recommend ChatGPT to other students for academic writing? Why or why not?
2. Is there anything else you want to share about your experience with ChatGPT in learning English?

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp. 441- 460
DOI: https://dx.doi.org/10.24093/awej/vol16no3.27

Literary Vocabulary of Charlotte Bingham's Novels: Stylistic and Functional Aspects

Yuliia Haidenko
Department of English for Humanities, Faculty of Linguistics, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine
Corresponding author: gaydenko.julia@ukr.net

Oksana Serheieva
Department of English for Humanities, Faculty of Linguistics, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine
Email: serheieva.oksana@lll.kpi.ua

Olesia Sribna
Department of English for Humanities, Faculty of Linguistics, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine
Email: olesia.lubianova@gmail.com

Olha Hrabar
Department of English for Humanities, Faculty of Linguistics, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine
Email: takichico@gmail.com

Received: 06/07/2025 Accepted:08/02/2025 Published:09/20/2025

Abstract

The study focuses on the question of how literary vocabulary in Charlotte Bingham's novels performs stylistic and functional roles in the structure of the literary text. The aim of the work was to identify and classify literary lexical units, determine their stylistic and functional colorings, and analyze their influence on the overall stylistics of the works. The main material was a corpus of 10 novels with a volume of 5125 pages. The methodology included a continuous sample for selecting novels. In addition, the stylistic classification according to Morohovsky et al. and the functional classification according to Brandes were used. The results show that Bingham's works present all types of literary vocabulary according to Morohovsky et al., which have a national, social, temporal, or professional coloring; neologisms are an exception. Functional colorings directly determined the stylistic functions of vocabulary in Bingham's prose. Further research should be directed at analyzing the colloquial vocabulary of her works, comparing literary and colloquial vocabulary within one work, identifying the connection between the choice of vocabulary and genre-thematic features, as well as studying the dynamics of the lexical composition in the author's work during different stages of her career. Such approaches will allow creating a more complete picture of the writer's individual style and a deeper understanding of the patterns of functioning of her linguistic means.

Keywords: Charlotte Bingham's novels, functional coloring, lexical units, literary vocabulary, stylistic coloring, stylistic functions, vocabulary strata

How to cite: Yuliia Haidenko, Y., Serheieva. O., Sribna, O., & Hrabar, O. (2025). Literary Vocabulary of Charlotte Bingham's Novels: Stylistic and Functional Aspects. *Arab World English Journal, 16*(3): 441- 460. https://dx.doi.org/10.24093/awej/vol16no3.27

Introduction

A literary text serves as a kind of reflection of the author's vision of reality and a means of communication between the author and the reader, containing linguistic units formed by the writer's creative plan (Culler, 2011; Matsevko-Bekerska, 2019; Plesník, 2022; Raj, 2012). The growing interest in research closely related to the human factor has led to the establishment of an anthropocentric paradigm in the analysis of a literary text (Kauza, 2020; Korol, 2014). Supporters of this approach view the author as a linguistic personality who consciously selects a system of linguistic means to express his or her individuality and worldview (Muslat, 2012; Ramazan, 2022). Within the framework of anthropocentric analysis, a literary text appears as the result of creative activity, which cannot be fully understood without taking into account its connection with the personality of the creator (Xu & Li, 2018). Stylistically marked vocabulary acquires key importance as a means of conveying the author's style and artistic intention (Raj, 2012; Wang, 2025). Despite extensive scholarly attention given to the analysis of stylistically colored vocabulary in English-language literature, the work of individual authors remains almost unexplored in the linguistic aspect. In particular, the novels of Charlotte Bingham – a modern English writer, a member of the Writers' Guild of Great Britain and laureate of the Romantic Novel of the Year award (1995) – have not yet become the object of special linguistic analysis. Her prose is distinguished by its masterful reproduction of historical eras, mostly the mid-twentieth century, with careful attention to cultural and social details. Subtle descriptions of interpersonal relationships, social hierarchy, and gender roles characterize her works. The combination of a historically authentic atmosphere and a socially sensitive perspective makes the analysis of Bingham's vocabulary particularly interesting: her choice of words not only forms an authentic background, but also encodes additional meanings related to the psychology of characters, the dynamics of plot relationships, and cultural trends. It is these features that create a promising and relevant basis for studying the stylistic and functional potential of literary vocabulary within the anthropocentric paradigm.

Therefore, the aim of the article is to analyze the stylistic and functional aspects of literary vocabulary in Charlotte Bingham's novels.

The study corresponds to modern trends in linguistics, oriented towards the anthropocentric paradigm and the understanding of the writer as an "artist of the word," who forms an individual approach to linguistic material, in particular to stylistically marked vocabulary. The analysis of Bingham's work allows not only to supplement the idea of her author's style, but also to expand the scientific vision of how vocabulary can simultaneously perform aesthetic, semantic, and cultural functions. In addition, the work fills a gap in the study of English fiction, which combines a historical atmosphere with a deep social subtext.

To achieve the goal, the following tasks are planned:

1. To define the concept of "literary vocabulary" within the framework of anthropocentric analysis.
2. To describe the groups (strata) and functional colorings of literary vocabulary in Charlotte Bingham's novels.
3. To identify and analyze the stylistic functions performed by literary lexical units in her works.

The current study answers the following research questions:

1. What literary lexical units are present in Charlotte Bingham's novels and how many of them are there?
2. What functional colorings does literary vocabulary create in her novels and which ones predominate?
3. What stylistic functions do literary lexical units perform in Bingham's works?

Literature Review

Stylistic Strata of Vocabulary

The concept of vocabulary is complex and controversial, as it has different interpretations and areas of use. In the simplest sense, vocabulary should be defined as the set of words that make up a language, or as a set of words that a particular person owns and uses (Encyclopedia Britannica, n.d.).

Despite differences in terminology, most researchers agree that the lexical composition of any language can be systematized into groups (strata) based on common properties (Ćarkić, 2008). Many scholars have considered the problem of lexical stratification, but their approaches differ in the level of detail and methodological emphasis. In particular, Morokhovskyi et al. (1991) proposed one of the most detailed stylistic classifications, focusing on subtle differences in stylistic connotation. Instead, Brandes (1990) developed a more functionally oriented model, which became the basis of this study. Other authors, for example, Babenko (2014), Kukharenko (2014), Oniani (2020), and Yefremova et al. (2019), supported a three-layer division of vocabulary into literary, neutral, and colloquial layers; however, they distinguished them mainly by terminology and level of detail.

The role of vocabulary in literary writing has also been widely analyzed. In particular, Khajieva & Kendjaeva (2019), Kishchenko & Striy (2022), and Malii & Opanasenko (2022) identified the features of the use of vocabulary as markers of the individual style of the writer. Researchers have suggested that the predominance of certain lexical strata in the author's texts reflects his creative individuality and sociocultural position (Riaz, 2021; Sydorenko, 2018). At the same time, the vocabulary of Charlotte Bingham's novels has not yet been the subject of a separate linguistic analysis, which determines the scientific novelty of this study. This work uses the Brandes (1990) classification, which distinguishes five groups of vocabulary: literary, standard literary, colloquial, vernacular, and vulgar. This typology is based on the broader three-layer model of Kukharenko (2014), which distinguishes literary, neutral, and colloquial layers. As Nikolov (2016) notes, the common property that unites words within each layer is called aspect: "the literary layer has a pronounced bookish character and relative stability; the colloquial layer is characterized by a lively oral character and is more changeable; the neutral layer is universal and most stable, because it is used in any communicative situation" (Nikolov, 2016, p. 119). In turn, Kukharenko (2014) clarified that the difference between the layers is determined by the presence or absence of stylistic connotation and the sphere of use. Neutral words do not have such a connotation and are appropriate in any communication. Literary and colloquial words, on the contrary, carry stylistic shades: the former satisfy the needs of official, scientific, and poetic discourse, while the latter are used in informal everyday speech.

Functional Stratification of Literary Vocabulary

In the system of stylistic lexicology, the stylistic coloring of a lexical unit means its general stylistic affiliation within the three-layer stratification of the vocabulary – for example, the "bookish" character in the literary layer or the “spoken” character in the colloquial layer (Mao et al., 2024; Kukharenko, 2014). For the literary layer, stylistic coloring is determined by the elevated, official, or specialized character of the word, which distinguishes it from neutral and colloquial vocabulary.

Functional coloring, in contrast to stylistic, is a narrower semantic and pragmatic shade that arises within a broader stylistic category. It reflects additional semantic or contextual associations – territorial, temporal, professional, national, or social – that are determined by the origin of the word, the history of its use, or the scope of its application (Brandes, 1990; Halliday, 1964; Sydorenko, 2020). Moreover, as Ćarkić (2008) noted, functional coloring is formed under the influence of geographical, temporal and social factors that accompany the emergence and functioning of a lexical unit. Thus, if

stylistic coloring determines the place of a word in a particular stylistic layer, then functional coloring specifies its semantic and cultural shades in a specific communicative situation.

According to Brandes (1990), literary vocabulary is functionally divided into three main groups. The first group includes foreignisms, barbarisms, and exoticisms, that is, words and phrases whose lexical meaning indicates the origin of the referent or the nationality or social status of the speaker. The second group consists of archaisms, historicisms, and neologisms, i.e., words and phrases whose meaning "reflects the timeframe of their coinage and functioning" (Karachun et al., 2020). The third group includes terms with nomenclature distinguished as a separate subgroup. This group comprises words and phrases whose lexical meaning indicates a person's profession. The boundaries between the functional colorings of some literary vocabulary groups are blurred. This is because some literary lexical units (for example, foreignisms) can create several functional colorings simultaneously but bring out only one depending on the communicative situation and the speaker's intention (Brandes, 1990).

Although nomenclature is often perceived as a purely technical or administrative vocabulary, in the Brandes/Morokhovskyi classifications, it belongs to literary vocabulary due to its bookish, codified character and its main functioning in formal, specialized, and documentary contexts. In scholarly texts, such vocabulary can perform stylistic functions, creating credibility in the depiction of a professional sphere, emphasizing the profession of a character, or giving a passage an elevated tone.

Thus, within the framework of modern classifications of Morokhovsky et al. (1991) and Brandes (1990), nomenclature vocabulary is attributed to literature, since its attribution is determined not by origin, but by function in the literary text. Despite the technical or terminological source, in prose it is used to create the effect of documentary, accuracy of description, and intellectual richness of the narrative. In such a space, nomenclature acquires a "book" character: it takes the reader beyond the boundaries of everyday speech, raises the register, and often performs a figurative or symbolic role. This brings it closer to other groups of literary vocabulary in terms of stylistic and functional coloring. In addition, according to the anthropocentric paradigm, the choice and method of integrating such vocabulary into the text reflect the author's conceptualization of the world and conscious language strategies that form the narrative voice and influence the reader's perception of the work.

Functional Coloring of Literary Vocabulary: Categories and Roles

Brandes (1990) explained foreign vocabulary, i.e., foreignisms, barbarisms, and exoticisms, as non-assimilated or partially assimilated language units used to elevate the language of a specific text fragment (words that we do not fully understand have a certain charm) and supply it with a proper functional national or social coloring.

Foreignisms are words and expressions used primarily in literary speech and reproduced as foreign language interspersions. They do not undergo grammatical or phonetic transformations that violate the established norms of the donor language. Foreignisms do not enter into stylistic opposition with the native words of the language where they appear, as they remain non-assimilated and do not belong to its word stock and expressive means (Morokhovskyi et al., 1991). In literary texts, foreignisms usually perform several stylistic functions: they help realistically represent the speech of a foreign character or reflect a character's high social status or education level (Nikolov, 2018).

Barbarisms are words and expressions borrowed from other languages but only partially assimilated. They retain their foreign appearance and/or sound, so they are perceived as originating from another language. Holovach (2016) rightly stated that barbarisms have almost exact synonyms in the borrowing language but are deliberately chosen by fiction authors for specific stylistic purposes, for example, to indicate characters' nationality, culture, or education.

According to Boltayevna's (2022) definition, exoticisms are "words borrowed from foreign languages that are included in non-equivalent vocabulary and denote realities characteristic of foreign peoples and countries" (p. 7). In contrast to barbarisms, they do not have synonyms within the borrowing language. As their meaning is strictly nominative, exoticisms form a morphologically homogeneous class of words limited to nouns. They designate specific realities of the donor language and, as a result, do not belong to the lexical and stylistic paradigms and expressive means of the borrowing language. However, fiction authors use them to depict certain aspects of foreign cultures, lifestyles, clothing, and customs (Holovach, 2016).

Besides, literary vocabulary also consists of archaisms, historicisms, and neologisms used to display functional temporal coloring and create the effect of elevated language. Pearson Publishing (n.d.) defines an archaism as "a word or phrase that is no longer used." Linguists treat archaisms as words that have fallen out of use but are still understood by native speakers (Kukharenko, 2014; Morokhovskyi et al., 1991). Based on the classification by Yefremova et al. (2019), archaisms in this study are divided into lexical and grammatical. Lexical archaisms are either obsolete words for existing objects or archaic meanings of the words from the active vocabulary. Grammatical archaisms are outdated grammatical forms ousted by newer ones. In literary texts, archaisms are used for several stylistic reasons: to create an atmosphere of antiquity or to produce a humorous effect.

Lexical units related to archaisms are historicisms. In modern linguistics, historicisms are explained as words designating historical phenomena that are no longer in use. Like archaisms, historicisms appear in literary texts to draw the reader's attention to a particular historical era and to render the text historically authentic (Nikolov, 2018).

Lexical units whose stylistic effect is opposite to that of archaisms and historicisms are neologisms. Linguists treat neologisms as words distinguished primarily by their novelty (Babenko, 2014; Yefremova et al., 2019). Llopart-Saumell & Cañete-González (2023) define a neologism as "a lexical unit that is characterized by not being previously recorded in a given language" (p. 3). Yefremova et al. (2019) add that a neologism is a word recently created or an existing word used in a new way while still preserving its sense of novelty. According to Babenko (2014), neologisms fall into lexical and stylistic categories. Lexical neologisms are coined to denote new objects, concepts, and realities. Stylistic neologisms are new, alternative names for existing phenomena. Lexical neologisms are treated as new words of the language as they are designed to pass into its lexical system and become legitimate vocabulary units. In contrast, stylistic neologisms emerge to suit a specific communicative situation, so their wide distribution and incorporation into the vocabulary are not expected (Llopart-Saumell & Cañete-González, 2023). Understandably, stylistic neologisms are context-dependent, and one can adequately interpret their meaning only by considering all the factors of a communicative situation that originated them. Researchers noted that speakers coin stylistic neologisms for a special communicative situation based on existing lexical units or word-formation models. This means that a stylistic neologism is an asset of speech predefined by the language system (Karachun et al., 2020; Turchak, 2022). Thus, in literary texts, lexical neologisms describe the signified from a temporal perspective. Stylistic neologisms help express the author's subjective attitude to specific people or phenomena and create a sense of humor or playfulness (Nikolov, 2018; Gao & Huang, 2025; Musa et al., 2023).

Another group of literary lexical units comprises terms that imbue text fragments with functional professional coloring. Sheremeta (2021) defined a term as "a word or a phrase that expresses a specific concept of any branch of science, technology, art, public life" (p. 143). On the other hand, Kukharenko (2014) stated terms as the "words denoting objects, processes, phenomena of science, humanities, technique" (p. 16). Terms are a defining feature of scientific style, so they are primarily used in scientific

works, technical texts, etc. However, they are often found in contexts outside their typical usage, specifically literary texts, where they perform stylistic functions (Bilenko, 2019; Larsson et al, 2025). In this case, the stylistic purpose of terms may vary. Fiction authors mostly use terms to render a text fragment more scientific and informative. In addition, terms are used to create "the true-to-life atmosphere of the narration," describe details of technical progress, or portray characters' interests, occupations, preferences, etc. (Nikolov, 2018, p. 120).

Within the group of terms, nomenclature forms a separate subgroup. Nomenclature is "a set of names used in a specific branch of learning or activity, as in biology for plants and animals" (HarperCollins, n.d.). In contrast to terms that denote abstract concepts and categories, nomenclature is a set of names of specific objects in a particular field of science, technology, art, etc. Like terms, fiction writers use nomenclature to describe characters' hobbies and occupations.

Thus, as can be seen from the analysis of the literature, previous studies focused mainly on the stylistic role of everyday, poetic, and archaic vocabulary in literary texts. In contrast, nomenclatural vocabulary was often considered only within the framework of technical and terminological studies, without taking into account its artistic potential. At the same time, the main gap is the lack of a systematic analysis of nomenclatural vocabulary within the framework of a holistic stylistic and functional classification of literary speech.

This study will fill this gap and use an approach that integrates functional and stylistic typology based on the analysis of the corpus of various novels by Charlotte Bingham. Such an approach will make it possible to identify the role of the author in the use of nomenclature to form an artistic effect. This will also make it possible to clarify its place in the structure of literary vocabulary. The significance of the work lies in expanding ideas about the stylistic resources of a literary text and deepening the anthropocentric approach, which considers lexical choice as an indicator of a writer's style.

Method

At the lexical level, Charlotte Bingham's novels feature stylistically colored vocabulary, including literary, colloquial, vernacular, and vulgar. This article focuses on literary lexical units that constitute the largest group in the discussed literary works (see Figure 1).

This study used several methods described in Selivanova (n.d.). This allowed us to analyze the literary vocabulary of Charlotte Bingham's novels. The method of analysis and synthesis was used to collect, systematize, and summarize linguistic data on literary vocabulary and other linguistic phenomena related to the topic of the study. The continuous sampling method was used to isolate all stylistically colored lexical units from the selected novels, which ensured the completeness and representativeness of the corpus. The descriptive method was used to record each unit and describe its relevant characteristics.

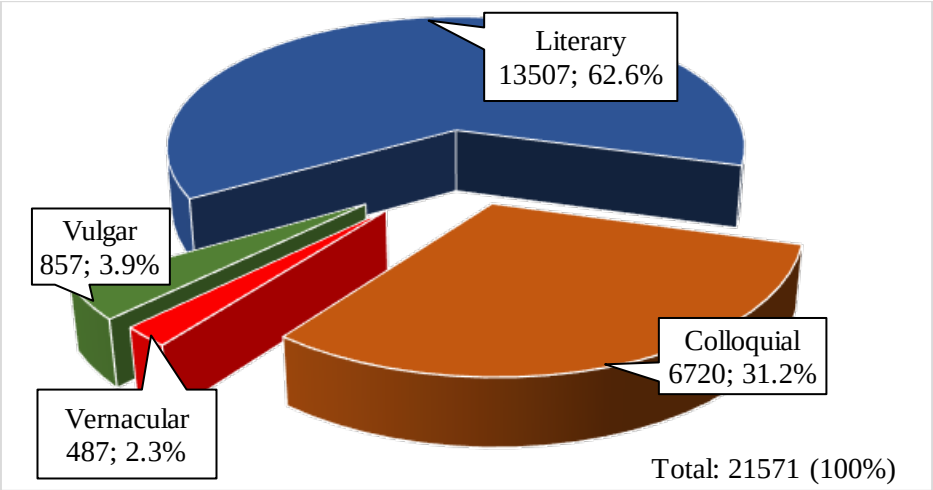


Figure 1. Stylistically colored vocabulary in Charlotte Bingham's novels: distribution

Operationalization of categories

To ensure clarity and consistency of the classification, the study adopted the Brandes (1990) scheme with refinements by Morokhovskyi et al. (1991), Kukharenko (2014), and Nikolov (2018). For each category, explicit inclusion and exclusion criteria were formulated:

Foreignisms – should contain completely unassimilated lexemes that retain the original spelling/phonetics and are not integrated into the morphological system of the recipient language; identified through etymological analysis and orthographic/phonological markers.

Barbarisms – included partially assimilated borrowings that retain foreign features but are not fully integrated into the recipient language; determined by comparing their morphology/phonetics with standard specific equivalents.

Exoticisms – non-equivalent borrowings to denote culturally specific realities, mainly nouns without functional synonyms in the recipient language; they are determined through semantic analysis and establishing a cultural referent.

Archaisms – words that have fallen out of use but remain understandable to native speakers. These categories are divided into lexical archaisms (obsolete words or meanings) and grammatical archaisms (obsolete grammatical forms).

Historicisms – lexemes that define fundamental historical realities that have entirely disappeared; they differ from archaisms in the absence of modern equivalents.

Neologisms – new words or meanings; they are divided into stylistic neologisms (created by the author for artistic effect) and lexical neologisms (new units that have entered general use).

Terms and nomenclature – industry-specific lexemes that function in a formal or professional register.

Each isolated unit was coded by two researchers who used a common code guide with definitions, examples, and decision-making rules. In controversial cases – for example, when distinguishing exoticisms from barbarisms or when defining stylistic vs. lexical neologisms – categorization was carried out by consensus based on clear operational criteria. At the same time, to check consistency, Cohen's kappa coefficient was calculated on a random sample of 15% of the data, which was 0.87. This indicated a high level of reliability.

Sample and Procedure

A study of the corpus of 10 novels by Charlotte Bingham, written in different periods of her work, with a total volume of 5125 pages: In Sunshine or Shadow (Bingham, 1991), Stardust (Bingham, 1993), Change of Heart (Bingham, 1994), Grand Affair (Bingham, 1997), The Kissing Garden (Bingham, 1999), Distant Music (Bingham, 2002), Magic Hour (Bingham, 2005), Out of the Blue (Bingham, 2006), The White Marriage (Bingham, 2007) and The Daisy Club (Bingham, 2009). The total volume of pages is 5125.

For this, a continuous sampling method was used. This means that the corpus of 5125 pages was manually reviewed, page by page, without gaps, in order to identify and record every instance of literary vocabulary. The selection was not random or partial – the process covered all texts in full.

The identification of literary vocabulary was carried out based on the classifications of Brandes (1990) and Morokhovskyi et al. (1991). Before the coding began, an operational list of categories was prepared with clear definitions and distinctive features of each subgroup (distinguishing foreignisms, barbarisms, and exoticisms; lexical vs. stylistic neologisms; lexical vs. grammatical archaisms).

At the same time, the attribution of a lexical unit to literary vocabulary was not based only on coincidence with the static dictionary. Each potential example was contextually evaluated within the novel to check whether it meets the defining and functional criteria of literary vocabulary — in particular,

whether bookish has a stylistic coloring and belongs to the literary layer according to the three-layer model of Kukharenko (2014). To confirm the classification, contextual features were taken into account: narrative style, speaker characteristics, and communicative situation. As a result of the procedure, 13507 units of literary vocabulary were recorded, which were entered into the database with an indication of the category, functional coloring, stylistic function, page number, and contextual notes. This made it possible to conduct both quantitative and qualitative analysis.

Procedures

The study of the literary vocabulary of Charlotte Bingham's novels was carried out in five successive stages. In particular, at the first stage, based on analysis and synthesis, the definition of stylistically colored and neutral vocabulary was clarified, and relevant classifications of the former were selected. At this stage, the concept of literary vocabulary was defined and its main groups, functional colorings and stylistic functions were described, relying on Brandes (1990), Morokhovskyi et al. (1991) and Kukharenko (2014).

At the next stage, a register of the literary vocabulary of Charlotte Bingham's novels was created, which was formed using the continuous sampling method and the descriptive method. The entire corpus of 5125 pages (10 novels) was reviewed manually, page by page, without gaps. As a result, 13507 units of literary vocabulary were identified, which were entered into the database with the page number, category, functional coloring, and contextual notes.

The third stage was componential. At this stage, componential analysis was applied to all 13507 units. For each lexeme, integral and differential semes were determined, and paradigmatic semantic relations (synonymic, antonymic, hyponymic) were established. The processing was carried out manually, and the results were arranged in tables for further comparison.

The fourth stage involved quantitative analysis, which covered the entire sample. The frequency and percentage distribution of categories were calculated in Microsoft Excel, and statistical testing was carried out in IBM SPSS Statistics. Graphs and charts were created in Microsoft Excel.

At the same time, an etymological analysis was also carried out. This method was used for all units using authoritative dictionaries (Oxford English Dictionary, Cambridge English Dictionary). This made it possible to establish the origin of the word, the type of borrowing, and the level of assimilation. Also, the method of word-formation analysis was applied to all units to determine the morphological structure, derivation patterns, and productivity of word-formation types. In addition, stylistic analysis was carried out for all units in order to determine the stylistic and functional coloring, as well as stylistic functions in the context of novels.

No partial sampling was used at any stage – all 13507 units of literary vocabulary were analyzed in full. Examples with unclear affiliation (for example, the distinction between exoticisms and barbarisms, or the definition of stylistic vs. lexical neologisms) were classified based on pre-established operational criteria after discussion between coders.

Results

The literary vocabulary of Charlotte Bingham's novels covers all lexical units defined in stylistic classifications. The largest category is nomenclature (9916 units; 73.5%), which reflects the author's frequent appeal to formalized, codified vocabulary from various professional spheres. For example, in "The White Marriage," the expression "registry office" is used, which functions in an official-administrative context. This group also includes words such as "probate", "curriculum vitae", "parish register", "quarterly report", "medical register", which not only perform a nominative function, but also

convey the social environment and professional status of the characters.

The second most frequent group is barbarisms (1714 units; 12.7%). This group expresses partially assimilated borrowings that retain the features of the source language. In particular, the work "In Sunshine or Shadow" features a "soirée", which lends the episode an atmosphere of sophistication. Also, found are "café", "château", "fiancé", "menu", "façade", which are used to denote objects and phenomena closely related to the culture of the donor country.

Archaisms (1042 units; 7.7%) in Bingham's works are both lexical and grammatical forms that refer to past eras. For example, in her works, the word "carriage" creates a sense of historical authenticity. Other examples: "drawing room", "attire", "courting", "shall not", "therewith", "hitherto", "perchance", allow to enhance the effect of ancient speech and correspond to the chosen period style.

At the same time, foreignisms (301 units; 2.2%) are represented by completely unassimilated foreign language units in their original form. In Stardust, "joie de vivre" is used, in Grand Affair, "faux pas", in Out of the Blue, "à la carte", and in The Kissing Garden, "bon voyage". Such usage gives the text a cosmopolitan flavor and accurately reproduces the speech of the characters.

The group Terms (209 units; 1.5%) includes specialized lexemes from narrowly professional fields: "tenancy agreement" (Out of the Blue), "by-laws", "statute", "contractual obligation", "trust fund". At the same time, historicisms (116 units; 0.9%) are words denoting objects or phenomena that have fallen out of use. In Grand Affair, "spats" is mentioned, an element of men's clothing of the early 20th century. Other examples concern the use of such words as "crinoline", "telegraph office", "gas lamp", "penny-farthing", and "powder horn".

Less common is the group of neologisms (116 units; 0.9%). They involved the use of new words in the general language. In the work "The Daisy Club", there is "designerwear", which refers to the modern fashion industry. The terms "microflat", "supermodeldom", "downsizing", "lifestyle guru", "glamping" are also distinguished, which reflect current social, economic and cultural trends.

The rarest group is exoticisms (93 units; 0.6%), which denote culturally specific realities without exact English equivalents. For example, in Magic Hour we use "kimono", in Stardust – "sari", in Distant Music – "fiesta", in Out of the Blue – "siesta", in Grand Affair – "poncho". Thus, the corpus of Charlotte Bingham's novels is characterized by a wide range of literary vocabulary (See Figure 2).

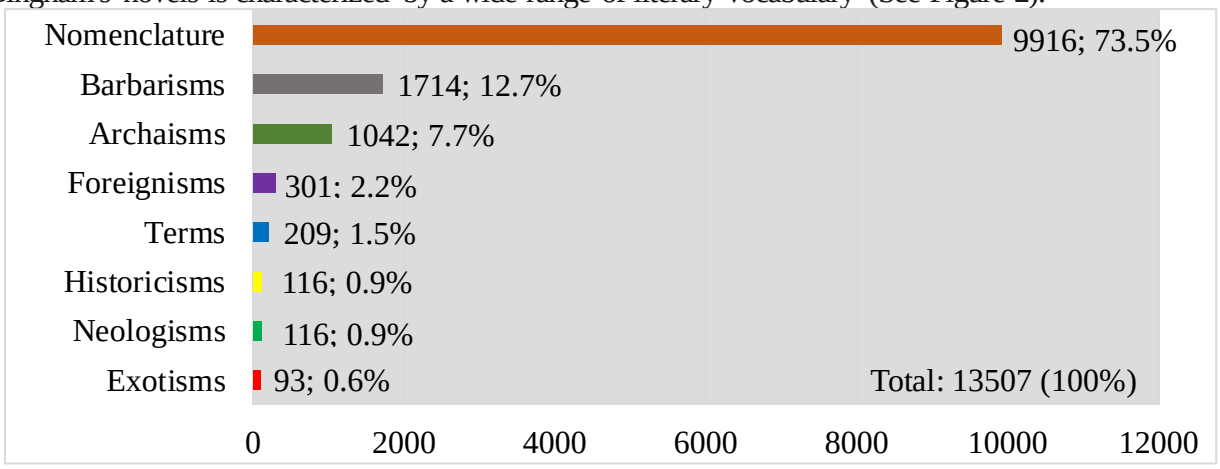


Figure 2. Literary vocabulary in Charlotte Bingham's novels: distribution

At the same time, it should be noted that neologisms, especially lexical neologisms, are extremely rare in the corpus (only 116 units; 0.9%). This may be an important factor that caused the lack of a statistically significant effect of their functional coloring on the stylistic functions they perform. Thus, this aspect should be considered as a potential limitation of the study and a direction for further testing

on a broader or more diverse material.

Research findings reveal that based on its functional coloring, the literary vocabulary of Charlotte Bingham's novels falls into three groups: 1) literary vocabulary of functional national and social coloring, 2) literary vocabulary of functional temporal coloring, and 3) literary vocabulary of functional professional coloring. Figure 3 shows the proportion of each group.

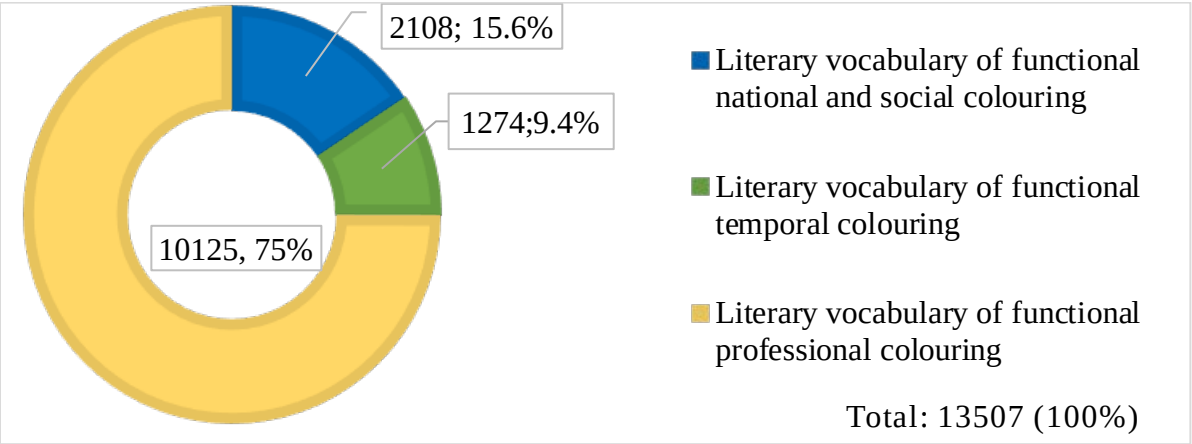


Figure 3. Literary vocabulary of functional coloring in Charlotte Bingham's novels: distribution

Figures 4-6 illustrate the distribution of word classes within each literary vocabulary group that renders functional coloring. Barbarisms are the most prevalent in the literary vocabulary of functional national and social coloring (1714 units; 81.3%). Archaisms constitute the majority of the literary vocabulary of functional temporal coloring (1042 units; 81.8%). Nomenclature (9916 units; 97.9%) dominates the literary vocabulary of functional professional coloring.

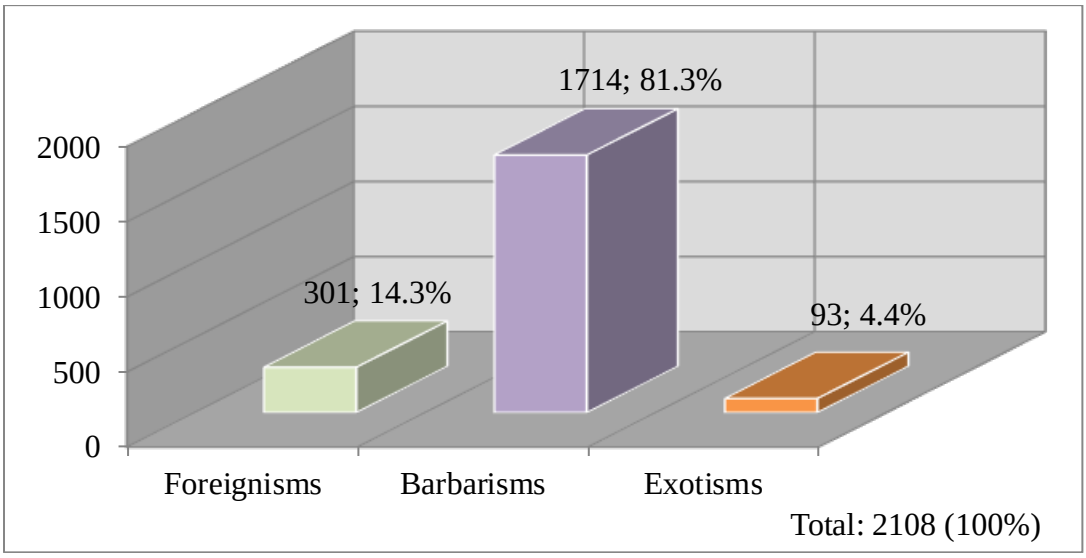


Figure 4. Literary vocabulary of functional national and social coloring in Charlotte Bingham's novels: distribution

ISSN: 2229-9327

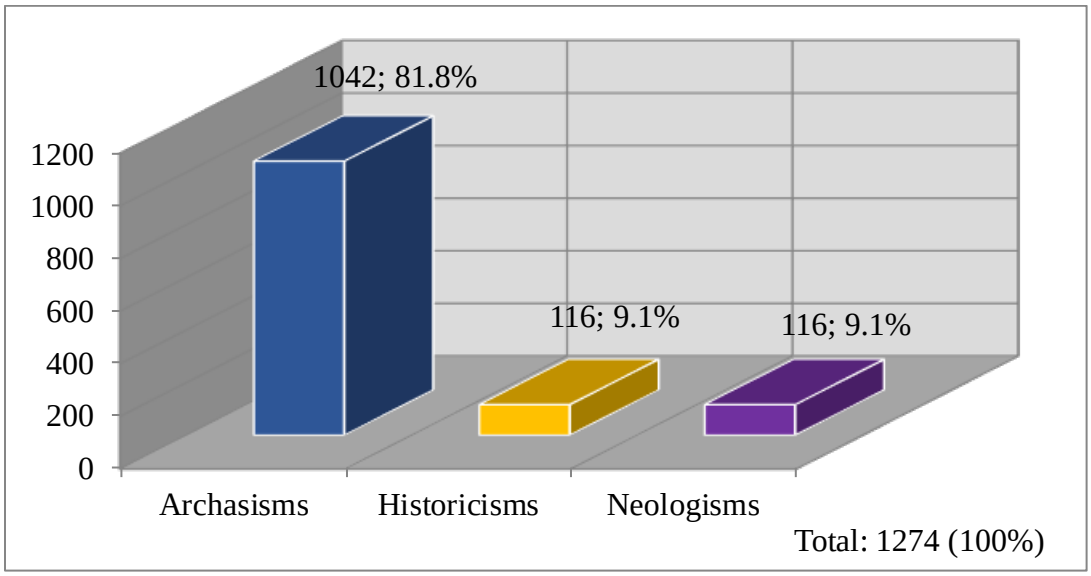


Figure 5. Literary vocabulary of functional temporal coloring in Charlotte Bingham's novels: distribution

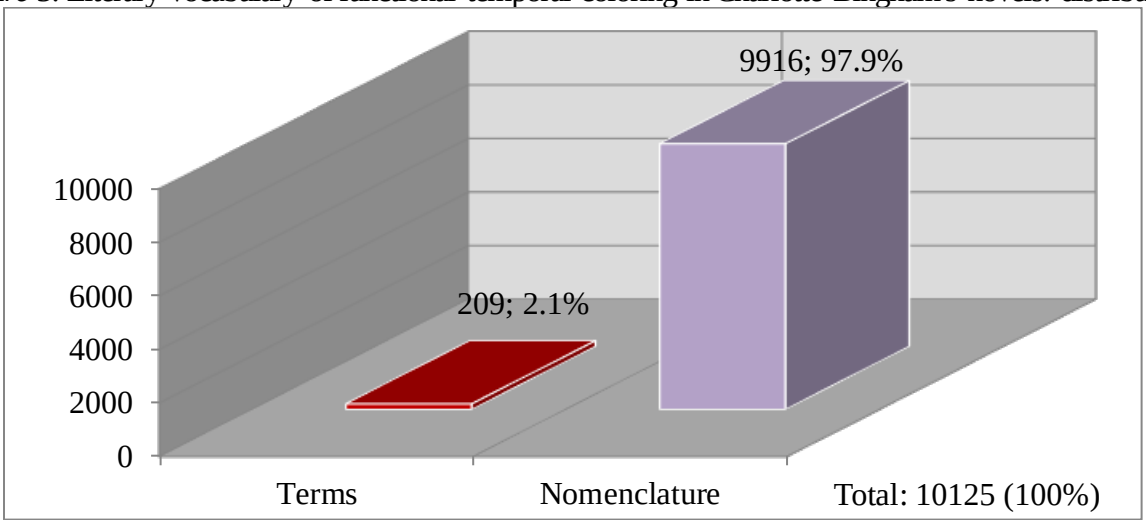


Figure 6. Literary vocabulary of functional professional coloring in Charlotte Bingham's novels: distribution

Table 1 reveals the stylistic functions of literary vocabulary in Charlotte Bingham's novels. The study results demonstrate that these stylistic functions are influenced by, but not limited to, the functional coloring of lexical units. The exception is neologisms whose functional coloring does not affect their stylistic functions.

Table 1. *Stylistic functions of literary vocabulary in Charlotte Bingham's novels*

Lexical units	Stylistic functions
Literary vocabulary of functional national and social coloring	
Foreignisms	1. Creating the effect of elevated language.
	2. Designating the characters' high social status and education level.
	3. Hiding the truth.
	4. Designating the characters' nationality or referents' origin.
Barbarisms	1. Designating the characters' nationality or the referents' origin.
	2. Expressing the characters' emotional and evaluative attitudes.
Exoticisms	1. Designating the referents' origin.

	2. Spatial orientation.
Literary vocabulary of functional temporal coloring	
Archaisms	1. Creating an atmosphere of antiquity.
	2. Expressing the characters' emotional and evaluative attitudes.
	3. Creating the effect of elevated language.
	4. Creating a humorous effect.
Historicisms	1. Rendering a text fragment historically authentic.
	2. Expressing the characters' emotional and evaluative attitudes.
Neologisms	1. Expressing the characters' emotional and evaluative attitudes.
	2. Creating a humorous effect.
Literary vocabulary of functional professional coloring	
Terms	1. Enhancing the authenticity of the narration.
	2. Creating a humorous effect.
	3. Designating the characters' occupation or profession.
Nomenclature	1. Designating the characters' occupation or profession.

Discussion

Concerning the first question proposed by the researchers, Charlotte Bingham's novels contain all groups of literary vocabulary. Nomenclature (9916 units; 73.5%), barbarisms (1714 units; 12.7%), and archaisms (1042 units; 7.7%) turned out to be the most numerous. Foreignisms (301 units; 2.2%) and terms (209 units; 1.5%) are less frequent. Historicisms (116 units; 0.9%), neologisms (116 units; 0.9%), and exoticisms (93 units; 0.6%) are rare.

Thus, this division showed a characteristic feature of Bingham's authorial style. The predominance of nomenclature and professional coloring indicated a constant desire to reproduce in detail material culture and specialized areas – from fashion and home crafts to military aviation and music. Such accuracy is typical of historical and social works, in which the effect of immersion depends mainly on the detailed naming of objects, tools, and professions. It also reflects the writer's anthropocentric orientation. In particular, she actively uses technical and professional vocabulary in her works. She thus fixes the characters in specific social roles and environments, making their personal and cultural identity more tangible for the reader.

Concerning the second question, the literary vocabulary of Charlotte Bingham's novels creates functional colorings such as national, social, temporal, and professional. Foreignisms, barbarisms, and exoticisms contribute to the functional national coloring of Charlotte Bingham's novels. Among these, only one subgroup – foreignisms – creates functional social coloring of the analyzed texts. Archaisms, historicisms, and neologisms add to the functional temporal coloring of Charlotte Bingham's novels, while terms and nomenclature form the functional professional coloring of the author's imaginative writing (Bingham, 1991; Bingham, 1993; Bingham, 1995).

The study's findings show that in Charlotte Bingham's novels, the group of literary vocabulary with functional professional coloring prevails (10125 units; 75%). Literary vocabulary with functional national and social coloring is less common (2108 units; 15.6%). The group of literary vocabulary with functional temporal coloring is the smallest in number (1274 units; 9.4%).

Concerning the third question, the literary vocabulary of Charlotte Bingham's novels performs several stylistic functions. These functions are related to but extend beyond the functional coloring of literary lexical units. Notably, in the analyzed texts, the stylistic functions of neologisms are independent of their functional coloring.

Foreignisms, barbarisms, and exoticisms represent the first group – literary vocabulary of functional national and social coloring – and perform several stylistic functions in Charlotte Bingham's novels.

In the analyzed texts, foreignisms appear as French lexemes. In the dialogues of fictional characters, they fulfill several purposes. Firstly, foreignisms create the effect of elevated language, indicating the characters' high social status and education level. For example, *'I adore old General Whiskers, but dear darling Louisa is sometimes too toffee-faced for words. Comme toute la haute bourgeoisie. Is that what is making you anxious?'* (Bingham, 1999, p. 40). Secondly, foreignisms may be associated with argot (a secret language of a closed social group invented to conceal the meaning of their conversations from outsiders) since the characters use them to hide the truth, as evidenced by the author's comments. For example, *'I am happy to hear it. Quatre ou cinq?' This was their code for meeting up at tea time* (Bingham, 2007, p. 88). Thirdly, foreignisms perform the characterological function, indicating the nationality of the speaking character. For example, *'My dear Ottilie, if you think that le tout Paris-le gratin as we call it in France – if you think that we know many, many people, I do assure you that we do not.'* (Bigham, 2002, p. 118).

There are rare instances of foreign proper names in Charlotte Bingham's novels. These are French, Italian, or German names of individuals or products that perform the characterological function, indicating the nationality or origin of the signified: *A La Fuite Des Heures* (Bingham, 2002, p. 277), *Verdi's Aida* (Bingham, 2002, p. 246), *Mein Kampf* (Bingham, 1999, p. 312).

The barbarisms of Charlotte Bingham's novels validate Holovach's (2016) idea that these are lexical units with near-exact synonyms in the borrowing language. For example, *Diane Smith-Werner was not the sort of woman who believed in kismet*. (Bingham, 1994, p. 42) (kismet = fate); *Finally, one morning, Amelia walked off into the garden, secateurs at the ready*. (Bingham, 1999, p. 273) (secateurs = pruning shears).

In the novels by Charlotte Bingham, this group is represented mainly by barbarisms from French: *protégée* (Bingham, 1994, p. 298), *cul-de-sac* (Bingham, 2002, p. 130). Barbarisms from other languages are less frequent. They are lexical units borrowed from Italian: *fiasco* (Bingham, 2002, p. 255), Latin: *paraphernalia* (Bingham, 2005, p. 349), Spanish: *vamoose* (Bingham, 2007, p. 348), Arabic: *kismet* (Bingham, 1994, p. 42), as well as Hindi: *pukka* (Bingham, 2007, p. 416).

Researchers (Holovach, 2016; Khajieva & Kendjaeva, 2019; Hossain, 2024) note that the primary stylistic function of barbarisms in the literary text is to indicate a character's origin, nationality, cultural or educational status. This is quite understandable since, due to incomplete assimilation, barbarisms are perceived as foreignisms. However, in Charlotte Bingham's novels, they are primarily used to express the characters' emotional and evaluative attitudes: *'A tall chap. Quite tall, dark-haired, tanned, handsome, dressed to the tens, absolutely pukka, not a whisker out of place.'* (Bingham, 2007, p. 416); *She plays the violin,' Freddie whispers again. 'She is a virtuoso.'* (Bingham, 1994, p. 421).

The exoticisms of Charlotte Bingham's novels are nominative lexical units marked by a strong connection with the donor language, as they designate realities specific to the daily life of its native speakers (Boltayevna, 2022; Holovach, 2016). In the analyzed texts, exoticisms indicate the foreign origin of the referents: 1) Arabic: *kasbahs* (Bingham, 2007, p. 280); 2) Greek: *pyramid* (Bingham, 2007, p. 248); 3) Italian: *spaghetti bolognese* (Bingham, 2002, p. 363); 4) Indian: *maharajahs* (Bingham, 2002, p. 342); 5) Spanish: *chili con carne* (Bingham, 1994, p. 91); 6) Latin: *toga* (Bingham, 2002, p. 527); 6) German: *dachshunds* (Bingham, 2009, p. 152); 7) Turkish: *turban* (Bingham, 2002, p. 107); 8) Scottish: *tam-o'-shanters* (Bingham, 2002, p. 459).

The empirical material shows that Charlotte Bingham uses exoticisms to display one of the chronotope (time-space) elements, particularly, as a spatial-characteristic method: *'When the basement was flooded in the winter, where was Nephew Sam? In Morocco, painting in kasbahs, smoking a hookah, and acquiring a taste for wearing a fez.'* (Bingham, 2007, p. 280).

The second group – literary vocabulary of functional temporal coloring – includes archaisms, historicisms, and neologisms whose stylistic functions in the analyzed texts vary.

The archaisms of Charlotte Bingham's novels are categorized as either lexical (archaic meanings of words still in active use or obsolete words for existing objects) or grammatical (outdated grammatical forms) as outlined by the classification of Yefremova et al. (2019).

Empirical material uncovered that the archaic meanings of the words from active vocabulary are context-dependent: *'You always did have the last word, didn't you, you rascal.'* (Bingham, 2002, p. 86) (rascal – [old-fashioned] a dishonest person (Cambridge Dictionary (n.d.))). *It was a blooming miracle that it was Father who had suggested that she have a motor car.* (Bingham, 2009, p. 122) (blooming – [old-fashioned] bloody (Cambridge Dictionary (n.d.)).

In Charlotte Bingham's novels, obsolete words are represented mainly by nouns: *curmudgeon* (Bingham, 2007, p. 115), *damsels* (Bingham, 2007, p. 164). Obsolete words functioning as other parts of speech in the writer's works are less frequent. They are verbs: *slay* (Bingham, 2007, p. 123), adjectives: *jolly* (Bingham, 1994, p. 205), or adverbs: *gaily* (Bingham, 1993, p. 239). The group of obsolete words also features outdated barbarisms and slang words where nouns dominate, and verbs are present to a limited extent: *mannequin* (Bingham, 1999, p. 71), *guvnor* (Bingham, 1993, p. 298), *dig* (Bingham, 1993, p. 523).

In the analyzed texts, obsolete words serve various stylistic functions. They are used to 1) create an atmosphere of antiquity and express the characters' emotional and evaluative attitudes: *'I should not have done that,' George muttered. – 'Of course, you should,' George's immediate neighbor assured him. 'Bounder asked for it.'* (Bingham, 1999, p. 222); 2) produce the effect of elevated language: *As he cantered faster and faster along the bridge to find Pippa, he laughed out loud as he saw himself, lance in hand, on a strong white charger, off to do battle, to joust for his lady's hand* (Bingham, 1993, p. 167).

Obsolete grammatical forms occur only sparingly in the researched works. The author uses them to 1) create the effect of elevated language: *'If we love in the other world as we do in this, I shall love thee to eternity.'* (Bingham, 1994, p. 499); 2) render specific text fragments humorous: *'All this summer heat must be making everyone crotchety, particularly husbands kicking their heels in the country with no one but their wives to talk to, and all the mistresses away picking up tans in the South of France. And not just tans, methinks.'* (Bingham, 2007, p. 273) (methinks – old use and humorous for *I think* (Cambridge Dictionary (n.d.)).

The share of historicisms in Charlotte Bingham's novels is modest. Two groups of lexical units contribute to the historical stylization of the analyzed texts and guide the reader through the timeline of events. The first group includes the names of objects or phenomena that have disappeared (Nikolov, 2018): *broadsword* (Bingham, 2002, p. 374), *guillotine* (Bingham, 2009, p. 412). The second group consists of the lexical units whose meaning refers to the contextually significant historical era. These include: 1) names of historical events or realities: *the Normandy landings* (Bingham, 2007, p. 351), *the Dig for Victory campaign* (Bingham, 2009, p. 224); 2) borrowed historicisms: *Communists* (Bingham, 1999, p. 190), *Gestapo* (Bingham, 2002, p. 117). The latter can express the character's emotional and evaluative attitudes: *'George —. He called you a Fascist.' – 'He called me a lot worse than that.'* (Bingham, 1999, p. 344); *'If you can rise above things, you will always be graceful and humane, you will*

never become a Nazi, or a lost cause, both of which have caused so much trouble in this world.' (Bingham, 2009, p. 450).

Few lexical units represent neologisms in Charlotte Bingham's novels. Following Babenko's (2014) classification, our research focused on lexical and stylistic neologisms. The quantitative analysis revealed no lexical neologisms; however, several stylistic neologisms appear in the studied texts. Stylistic neologisms are elements of speech, as they are created for a particular communicative situation. Nevertheless, their form and meaning are predefined by the language system since they follow its semantic, word-formation, and grammatical patterns (Turchak, 2022). Stylistic neologisms of the studied literary works are the elements of Charlotte Bingham's speech, represented by various word-formation structures: 1) affixal: *un-Cliffie-like* (Bingham, 2002, p. 292); 2) compound: *Gent's Relish* (Bingham, 2002, p. 52); 3) derivative: *an Uncle Randy-ism* (Bingham, 2007, p. 414) (word composition and suffixation: Uncle + Randy + ism); 4) conversive: *dear boying* (Bingham, 2002, p. 528) (conversion: $dear_{Adj}$ $boy_{N} \rightarrow$ dear boy_{V}); 5) telescoped: *blame-aholic* (Bingham, 1994, p. 491) (telescoping: blame + alcoholic); 6) abbreviated: *G.A.* (Bingham, 2009, p. 70) (abbreviation of the character's proper name *Guy Athlone*).

In Charlotte Bingham's novels, stylistic neologisms are more emphatic than words from an active literary vocabulary. This is because they result from the author's need to designate new concepts and are, therefore, fresh in form and context-dependent. The author uses stylistic neologisms to 1) convey the emotional and evaluative attitudes: *Clifton stood backstage at the little amateur theatre looking strangely un-Cliffie-like.* (Bingham, 2002, p. 292); 2) create a humorous effect: *Sam frowned. He knew that the reference to 'she' must mean Sunny Chantry, and that being so, he must now tread with the feet of the sanctified, be as stalky as he had ever been. He reached back in time to an Uncle Randy-ism* (Bingham, 2007, p. 414).

Therefore, the rarity of neologisms in the corpus, in particular lexical ones, to some extent explains why their functional coloring does not affect the stylistic functions they perform in the studied texts. Presumably, this feature is due to Bingham's orientation to the traditional lexicon, consistent with the frequent historical and period decorations of her novels, where anachronistic innovations could violate the authenticity of the narrative. Instead, the author introduces stylistic neologisms only when they have a clear communicative or humorous purpose, and not to expand the vocabulary of the language.

The third group – literary vocabulary of functional professional coloring – includes terms and nomenclature with the least diverse stylistic functions in Charlotte Bingham's novels.

The author uses terms for several stylistic purposes: 1) to enhance the authenticity of the narration (Nikolov, 2018): *Ottilie had quickly and promptly set about dealing with Mrs Tomber before the wretched woman could set about effecting any more 'shrinkage,' as it was known in the trade.* (Bingham, 2002, p. 283); 2) to produce a humorous effect: *She had never before realized that the definition of the theory of gravitation had been the result of some reprehensible inertia on the part of Isaac Newton.* (Bingham, 2007, p. 189); 3) to mark the character's occupation or profession: *By dusk, Freddie had sketched out the entire concerto, the opening allegro, the andante cantabile, and the concluding movement, which he designed to be composed as another allegro, this one andante grazioso – cadenza – allegro ma non troppo.* (Bingham, 1994, p. 87).

The nomenclature of Charlotte Bingham's novels comprises a set of names used in a specific area of human expertise or profession. These are the names of: 1) cars: *Hillman 6* (Bingham, 1999, p. 21); 2) car parts: *inlet over exhaust* (Bingham, 2007, p. 26); 3) military aircraft: *Sopwith Tiger* (Bingham, 1999, p. 272); 4) aircraft details: *elevators* (Bingham, 1999, p. 272); 5) military organizations: *VAD* (abbreviation of *Voluntary Aid Detachment*) (Bingham, 2009, p. 199); 6) fabrics: *Abraham's organza*

(Bingham, 2007, p. 141); 7) sewing machines: *Singer sewing machine* (Bingham, 2007, p. 228); 8) movie cameras: *newsreel cameras* (Bingham, 1993, p. 239); 9) musical instruments: *oboe* (Bingham, 1994, p. 93); 10) muscles: *calf muscles* (Bingham, 2007, p. 42); 11) medicines: *thiamine* (Bingham, 1993, p. 310).

The research findings show that Charlotte Bingham uses nomenclature primarily to depict the character's profession or occupation. In the novel "Change of Heart" (Bingham, 1994), for instance, this stylistic function is performed by the names of musical instruments. Thus, the high proportion of vocabulary with a professional connotation indicates that Bingham's artistic world is based on rich social and professional contexts. The use of specialized vocabulary not only clearly outlines the place of the characters in certain professional or cultural spheres but also serves as a means of thematic disclosure. Regardless of whether these are descriptions of social life, military service, or musical skill, professional words are actively used in these works. Such integration of professional terms corresponds to her anthropocentric narrative strategy, in which the life experience, career, and hobbies of the characters become the core of plot development and emotional involvement of the reader.

Limitations

Despite the large size of the corpus (5125 pages, 13507 units of literary vocabulary), the results should be interpreted taking into account several limitations. As can be seen from the analysis of the results, individual categories are rare, in particular, neologisms (0.9%). This, in turn, limited the possibility of drawing statistically significant conclusions about the influence of their functional coloring on stylistic functions. Secondly, the classification of individual units, particularly the distinction between barbarisms, foreignisms, and exoticisms, or the definition of types of archaisms, sometimes presented borderline cases. The resolution of these cases depended on the agreed-upon decisions of the researchers. Finally, it should be recognized that the analysis covered only novels by one author and one language environment; therefore, generalizing the conclusions to other genres, eras, or authors requires additional research. It is also worth noting that the lack of automated morphological and stylistic markup tools has led to a completely manual nature of the selection and classification of units, which, despite verification measures, may contain a human factor.

Conclusion

The novels by Charlotte Bingham feature an extensive palette of stylistically colored lexical units. Among these, the most common are literary, used primarily in formal communicative situations. All groups of literary vocabulary appear in the analyzed novels; however, nomenclature, barbarisms, and archaisms are the most frequent. The analysis reveals that the stylistic functions of Charlotte Bingham's literary vocabulary are related to but extend beyond the functional coloring of lexical units, and in some cases, bear no relation to the functional coloring. The above-mentioned features of the analyzed literary vocabulary reveal the linguistic choice and creative individuality of the contemporary English novelist Charlotte Bingham. The quantitative and qualitative analysis showed that Bingham mainly uses literary vocabulary – especially nomenclature and barbarisms – to precisely construct socially and professionally marked contexts, applying functional colorings. Instead, the author leaves archaisms, historicisms, and neologisms to achieve specific tonal and expressive effects.

Several functional colorings of literary vocabulary were revealed in the studied works, including national, social, temporal, and professional, with the professional dominating. It is important that the stylistic functions established in Bingham's prose are related to functional coloring, but are not limited to it; in several cases, they do not coincide with it at all, fulfilling broader narrative, emotional, or

characterological tasks. The totality of the obtained results made it possible to define Charlotte Bingham as a linguistic personality who makes conscious, contextually formed lexical choices, forming the stylistic individuality of her works.

Further research should be expanded to a detailed study of colloquial vocabulary in Bingham's novels. This, in turn, will allow us to more fully imagine her stylistic repertoire and show how the author combines literary, neutral, and colloquial layers of vocabulary in the formation of a narrative voice and the reproduction of the interaction of formal and informal registers in the social world of her works of art.

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: AI and AI-assisted technologies were not used.

About the Authors:

Yuliia Haidenko has a Ph.D. in Philology and is an associate professor at the Department of English for Humanities of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute." Her research interests include the semantic, stylistic, and pragmatic potential of the author's language and the author's speech in belles-lettres. ORCID: https://orcid.org/0000-0002-4063-525X.

Oksana Serheieva is a lecturer at the Department of English for Humanities of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute." Her research interests include the lexical-semantic and stylistic aspects of color names in English idioms. ORCID: orcid.org/0000-0002-7396-5511.

Olesia Sribna is a CELTA-certified ESP teacher, educator and researcher with a focus on digital technologies at the Department of English for Humanities of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute." Her research interests include online teaching and integration of AI and technology-enhanced learning in the ESP classroom. ORCID: https://orcid.org/0000-0002-2226-0300.

Olha Hrabar is a lecturer at the Department of English for Humanities of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute." Her research interests include lexical, grammatical, and stylistic peculiarities of prose and poetry translation. ORCID: https://orcid.org/0000-0001-5752-396X.

References

Babenko, O. (2014). *A book of practice in the stylistics of the English language.* Kyiv: NUBiP.
Bilenko, T. (2019). Terms functioning in the works of Ukrainian poets of the 19th and 20th centuries. *International Humanitarian University Herald, 42*(1), 10–13.
Bingham, C. (1991). *In Sunshine or Shadow.* London: Bantam Books.
Bingham, C. (1993). *Stardust.* London: Bantam Books
Bingham, C. (1994). *Change of Heart.* London: Bantam Books.

Bingham, C. (1997). *Grand Affair*. London: Bantam Books.
Bingham, C. (1999). *The Kissing Garden*. London: Bantam Books.
Bingham, C. (2002). *Distant Music*. London: Bantam Books.
Bingham, C. (2005). *Magic Hour*. London: Bantam Books.
Bingham, C. (2006). *Out of the Blue*. London: Bantam Books.
Bingham, C. (2007). *The White Marriage*. London: Bantam Books.
Bingham, C. (2009). *The Daisy Club*. London: Bantam Books.
Boltayevna, D. (2022). Barbarisms and exotisms are a special type of non-equivalent vocabulary. *Central Asian Journal of Theoretical and Applied Science, 3*(12), 7–8. https://cajotas.centralasianstudies.org/index.php/CAJOTAS/article/view/1016
Brandes, M. (1990). *Stylistics of the German language* (2nd ed). Moscow: Vysshaya shkola.
Cambridge University Press (n.d.). *Cambridge Dictionary*. https://dictionary.cambridge.org
Ćarkić, M. (2008). *Semantics and Style. Stylistyka, 17*, 5–16. https://czasopisma.uni.opole.pl/index.php/s/article/view/3609
Culler, J. (2011). *Literary theory: a concise introduction* (2nd ed). Oxford: Oxford University Press. https://doi.org/10.1093/actrade/9780199691340.001.0001
Encyclopedia Britannica, Inc. (n.d.). Vocabulary. In *The Britannica Dictionary*. https://www.britannica.com/dictionary/vocabul%5Cary
Gao, Q., & Huang, J. (2025). Design and Implementation of Classical Literature Sentiment Analysis System Based on Ensemble Learning and Graph Neural Network. *International Journal of Cognitive Computing in Engineering*. https://doi.org/10.1016/j.ijcce.2025.05.004
HarperCollins (n.d.). Nomenclature. In *Collins Dictionary*. https://www.collinsdictionary.com/dictionary/english/nomenclature
Holovach, T. (2016). *Exoticisms as markers of intercultural communication (based on modern English media texts)* (Doctoral Dissertation). Lviv: Ivan Franko. National University of *Lviv*.
Hossain, K. I. (2024). Literature-Based Language Learning: Challenges and Opportunities for English Learners. *Ampersand*, 100201. https://doi.org/10.1016/j.amper.2024.100201
Karachun, Y., Haidenko, Y., & Borkovska, I. (2020). Compound term-nouns in electrical engineering texts: structural, semantic, and functional peculiarities. *Arab World English Journal*: *Special Issue on English in Ukrainian Context*, 142–160. https://dx.doi.org/10.24093/awej/elt3.13
Kauza, I. (2020). *Thought in thought as the form of inner speech in Anglophone artistic prosaic texts (based on Margaret Drabble's novels)* (Doctoral Dissertation). Lviv: Lviv Polytechnic National University.
Khajieva, F. & Kendjaeva, G. *(2019). Study* of *stylistic lexicology. TEST Engineering and Management, 81, 4162–*4169. http://testmagzine.biz/index.php/testmagzine/article/view/567/507
Kishchenko, A. & Striy, L. (2022). Lexical means of expression of the category of expressiveness in modern Ukrainian prose. *Odesa National University Herald, 2(26), 42–53.*
Korol, L. (2014). Semantic centers of anthropocentrism in artistic work. *Humanitarian Education in Technical Universities, 30*, 156–168.
Kukharenko, V. (2014). *A book of practice in stylistics.* Vinnytsya: Nova Knyha.
Larsson, T., Kim, T., & Egbert, J. (2025). Introducing and comparing two techniques for key lexical bundles analysis. *Research Methods in Applied Linguistics*, *4*(3), 100245. https://doi.org/10.1016/j.rmal.2025.100245
Llopart-Saumell, E. & Cañete-González, P. (2023). Are Stylistic Neologisms More Neological? A Corpus-Based Study of Lexical Innovations of Women and Men. *Languages, 8*(3), 1–17. https://doi.org/10.3390/languages8030175

Malii, A. & Opanasenko, O. (2022). Vocabulary designating emotions in a literary text (based on the German and English languages). *Transcarpathian Philological Studies, 25*(1), 132–136.

Mao, Y., Liu, Q., & Zhang, Y. (2024). Sentiment analysis methods, applications, and challenges: A systematic literature review. *Journal of King Saud University – Computer and Information Sciences*, 102048. https://doi.org/10.1016/j.jksuci.2024.102048

Matsevko-Bekerska, L. (2019). The text – the reader – the literary writing: a cognitive and narratological paradigm. *Scientific Journal of Polonia University, 34,* 36–46. https://doi.org/10.23856/3404

Morokhovskyi, O., Vorobiova, O., Lykhosherst, N., & Tymoshenko, Z. (1991). Stylistics of the English language. Kyiv: Vyshcha Shkola.

Musa, F., Othman, Z., Md Yunus, M., & Rashid, T. (2023). Challenges and Strategies in Workplace English Communication: Undergraduate Accountant Internship. *Arab World English Journal*, *14*(4), 283–299. https://doi.org/10.24093/awej/vol14no4.17

Muslat, A. W. M. (2012). *Translation and the characteristics of literary text. Arab World English Journal, 3*(1), 42-49. https://awej.org/images/AllIssues/Volume3/Volume3Number1March2012/3.pdf

Nikolov, A. (2018). Stylistic classification of the English vocabulary. *Trakya University Journal of Faculty of Letters (TUJFL), 6*(11), 115–122. https://dergipark.org.tr/tr/download/article-file/592538

Oniani, T. (2020). Stylistic classification of English vocabulary. *International Journal of Science and Research, 9*(2), 1548–1552. https://www.ijsr.net/archive/v9i2/SR20221100932.pdf

Pearson Publishing (n.d.). Archaism. In *Longman Dictionary of Contemporary English.* https://www.ldoceonline.com/dictionary/archaism

Plesník, L. (2022). Some comments on the relationship between text and reality. *Ars Aeterna, 14*, 1–11. https://doi.org/10.2478/aa-2022-0007

Raj, P. (2012). Author and text: reading Michel Foucault's "What is an author?" *The Criterion: An International Journal of English, 3,* 1–11. https://www.researchgate.net/publication/234028790_Author_and_Text_Reading_Michel_Foucault's_What_is_an_Author

Ramazan, F. (2022). Methodological assumptions of revealing Agatha Christie's idiostyle pattern. *Scientific journal of the Lviv State University of Life Safety "Philological Periodical of Lviv," 12,* 155–161. https://doi.org/10.32447/2663-340X-2022-12.25

Riaz, M. (2021). Literary and colloquial English: formality differences between the novel Pride and Prejudice and its screen adaptation. *Journal of Research in Humanities, 57,* 103–123. https://www.researchgate.net/publication/358817382_Literary_and_Colloquial_English_Formality_Differences_between_the_Novel_Pride_and_Prejudice_and_its_Adaptation

Selivanova, O. (n.d.). *The main methods of modern linguistics of text.* https://www.scribd.com/document/582100584/the-main-methods-of-modern-linguistics-of-text?language_settings_changed=English

Sheremeta, L. (2021). Concepts of terminology, terminology studies, the term system, and the term. *Current Issues of the Humanities, 35*(6), 140-145. https://doi.org/10.24919/2308-4863/35-6-21

Sydorenko, I. (2018). The notion of idiostyle in linguistic studies of literary texts. *Advanced Linguistics, 1,* 15–20. https://doi.org/10.20535/2617-5339.2018.1.148751

Sydorenko, I. (2020). Lexical variety of Arthur Miller's dramatic texts. *European Socio-Legal and Humanitarian Studies, 1,* 83–92. https://ehs-journal.cunbm.utcluj.ro/wp-content/uploads/2021/06/ESLHL-2020-11.pdf

Turchak, O. (2022). Verbal neologisms of individual authors in the newspaper language of the late 20th century: semantics and word formation. *Alfred Nobel University Journal of Philology, 1,* 153–160.

Xu, M. & Li, Z. (2018). Author, text, and reader: death of the author and meaning of the text. *Journal of Literature and Art Studies, 8*(12), 1669-1673. doi: 10.17265/2159-5836/2018.12.006

Wang, H. (2025). Multimodal Digital Storytelling as Literacy Learning and Moral Cultivation Practices for University EFL Students. *Arab World English Journal*, *16*(1), 366–383. https://doi.org/10.24093/awej/vol16no1.23

Yefremova, N., Pavliuk, A., Sheludchenko, S., & Boichuk, V. (2019). *English lexicology*. Lutsk: Lesya Ukrainka Eastern European National University.

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp.461-471
DOI: https://dx.doi.org/10.24093/awej/vol16no3.28

Learning Self-Efficacy, Regulation, and Resilience among English-Majored Students in the Vietnamese Context

Do Minh Hung
Foreign Languages Faculty, Dong Thap University Cao Lanh City, Vietnam
Email: dmhung@dthu.edu.vn

Received: 06/23/2025 Accepted:09/13/2025 Published: 09/24/2025

Abstract
Self-efficacy, regulation, and resilience are crucial in the general learning process across disciplines and settings. These three factors were explored in past research worldwide, but their correlations have yet to be investigated sufficiently among English-majored students in the current Vietnam context. Thus, the present study aimed to tap into these three factors by seeking answers to the guiding questions of what target students' levels of learning self-efficacy, regulation, and resilience were, and what possible correlations between the three factors were. For data collection, 400 English-majored students took part in a questionnaire survey. These students included four groups of first-year students, sophomores, juniors, and seniors from a university in Vietnam. Then, metric data analyses show that students' learning self-efficacy and regulation were at the upper-average level, while their resilience was at the high level. In addition, learning self-efficacy appeared to have positive correlations with the other two factors. However, learning self-efficacy levels found in the present study did not significantly differ among the four groups of first-year students, sophomores, juniors, and seniors. These findings provide more insights into the three factors, pedagogical implications, and further inquiries into the topic under discussion.
Keywords: English-majored students, learning self-efficacy, regulation, resilience, Vietnamese context

How to cite: Do, M. H. (2025). Learning Self-Efficacy, Regulation, and Resilience among English-Majored Students in the Vietnamese Context. *Arab World English Journal, 16*(3):461-471. https://dx.doi.org/10.24093/awej/vol16no3.28

Introduction

It is generally recognized that self-efficacy is worth researching in educational areas because it impacts learners' thoughts, emotions, and intentional behaviors from the outset and during the learning journey. A decent understanding of learning self-efficacy and its building blocks will provide educators and teachers with channels to create and develop learner self-efficacy conducive to their learning journey. Thus, self-efficacy has been investigated in the educational setting over the past decades (Akanni & Oduaran, 2018; Bandura, 1986; Lestari et al., 2020; Lin et al., 2025; Luo et al., 2023; Ma et al., 2024; Maier et al., 2021; Mahmud et al., 2024; Nie et al., 2012; Xu et al., 2022; Wang et al., 2013; Xu & Wang, 2025). Self-efficacy is correlated with normative behaviors, learning manipulations, and goal-related expectations at designated levels (Bong & Skaalvik, 2003). It is essentially a driving force for learners to self-regulate learning activities, emotional states, and resilient efforts for desired learning outcomes (Bandura, 2012; Van-Zyl et al., 2022). In other words, learner self-efficacy is believed to be a potent predictor of academic engagement, regulations, and resilience. That means these three dimensions (engagement, regulations, and resilience) together are among the essential mediators for positive learning performance and achievement (Genç et al., 2016; Do, 2022; Luo et al., 2023; Mohebbi, 2025; Zimmerman et al., 1992). Past investigations indicate that students of high self-efficacy likely stay on their learning pursuits rather than giving up despite facing various stresses, challenges, and setbacks by being self-motivated and making greater endeavors for learning success attainment (Bandura, 2012; Chu et al., 2024; Luo et al., 2023; Zimmerman et al., 1992). It is also noted that strong efficacy fosters students to confidently set higher academic goals and pool more affordable resources for goal attainment (Bong & Skaalvik, 2003; Gong & Xu, 2024; Mohebbi, 2025; Waddington, 2023).

Still, few empirical studies on learning self-efficacy have been done in Vietnam's contexts, especially the link between learning self-efficacy, regulation, and resilience in higher education. Moreover, the impact of academic years (i.e., freshmen, sophomores, juniors, and seniors at college) on student efficacy has yet to be sufficiently explored in previous relevant studies. With that in mind, the present study examined the learning self-efficacy perceived by college students majoring in English as a foreign language (EFL). To its end, the present study was directed by four target questions:

1. At what level of learning self-efficacy do English majors self-perceive?
2. How do these students display their learning regulation and resilience?
3. Does the learning self-efficacy of students have correlations with their learning regulation and resilience?
4. Is learning self-efficacy impacted by their academic years at college?

Literature Review

What is Self-efficacy?

Bandura (1986) defined self-efficacy as a person's evaluative belief in their capability to perform a given task. It denotes personal self-appraisal based on reliable skills and abilities already earned. Thereby, people will determine if they can take up the informed task (such as jumping across a 2-meter ravine, running 5 km long within 30 minutes, or finishing a Master's degree program in one year). More importantly, perceived self-efficacy reflects individual forethoughts, expectations, and commitments to task completion and its values. That means if positive or strong enough, self-efficacy will stimulate people's motivations to exercise resources optimally and take

up a strategic course of action. Accordingly, their self-regulation and sufficient resilience are also pushed up until the task is achieved as desired. On the other hand, individuals of low efficacy will seek causal attributions so that they would not get involved in the informed task (Bong & Skaalvik, 2003). Such avoidance attitudes resulting from low efficacy will likely impede human determination, effort, dedication, and positive performances (Waddington, 2023). Thus, self-efficacy influences human cognitions, affect, and executive functioning (Bandura, 1986).

How Can Self-efficacy Be Developed?

People develop self-efficacy in four significant ways (Bandura, 2012). First of all, *mastery experience* is the fundamental source from which one comes to acquire the nature of something by getting involved in doing it several times, again and again. Obtaining success allows for self-efficacy to be promoted. Thus, when learning something such as a second/foreign language, learners need to gain knowledge about and practice the target language used in different ways to familiarize themselves with this language in experiential communication (speaking and writing in particular). That repetitive process enables learners to build and reinforce confidence in the target-language environments. The second way for self-efficacy development is via *social modeling (*also known as *vicarious learning)*, which occurs when people observe others in similar performances or hear of others' relevant stories of ultimate achievements through consistently strategic actions and perseverant efforts. These inspiring stories likely motivate people to feel more confident in themselves based on the common reasoning that “others like me can do well, so can I”. As a result, it is advisable to integrate relevant exemplars in the classroom setting so that students can learn from others' valid experiences, especially how others take up self-regulations and face challenges and setbacks with solid resilience until the ends are obtained.

Social impact via *verbal persuasion* is assigned as the third source for confidence building. Positive encouragement and support from others around, such as relatives and friends, is believed to drive people to take further steps in their perceived capabilities. According to Bong and Shaalvick (2003), verbal persuasion is considered "most effective when people who convey the efficacy information are viewed as knowledgeable and credible and when the information is viewed realistically" (p.6). Thus, expert-qualified teachers of high efficacy beliefs appear to be the most available channel of verbal persuasion for learner efficacy growth. Teachers are advised to grant learners salient and timely feedback constructively so that learners likely understand what they should do for improvement and recognize that improvement is within reach (Waddington, 2023). Such teachers will also tend to encourage learners to take learning self-regulations and learning-assisted challenges for personal potential to be explored (Schunk & DiBenedetto, 2016). This approach will lead learners to recognize that mastery is achievable when there is ample investment in cognition, affect, and appropriate choices of necessary execution and regulations. The fourth way to enhance self-confidence (and thus an increase in self-efficacy) is the *state of personal physiology and emotion*. Physical and affective well-being is closely related to someone's perceived efficacy, especially in challenging and stressful situations. This inner strength empowers one to reduce anxieties or fears of inappropriate performances (such as speaking English poorly in front of the classroom due to concerns about making several pronunciation or grammatical mistakes).

Studies on English Learning Self-efficacy

On learning English as a foreign language, past worldwide investigations have repeatedly recorded the pivotal role of self-efficacy in the learning process and achievement (Chen, 2022; Genç et al., 2016; Lin et al., 2025; Polrak et al., 2023; Truong & Luu, 2024; Truong & Wang, 2019; Zhang et al., 2023). A pertinent study by Lin et al. (2025) among English as a foreign language (EFL) learners from China showed a positive correlation between participants' English-speaking self-efficacy and their English learning motivation. Similarly, the study by Torres et al. (2023) among EFL undergraduates from Peru showed a positive correlation between academic self-efficacy and students' performances in English reading, writing, and oral communication skills. On this line of previous research involving EFL students, other studies such as Genç et al. (2016) in Turkey, Huang (2024) in Taiwan, Li et al. (2023) in China, and Polrak et al. (2023) in Thailand also recorded the participants' English performances associated with their academic self-efficacy. Additionally, it is recognized that attentionally locating low self-efficacy among EFL students will pay initial ways to prevent negative self-beliefs easily navigated to burnout (Waddington, 2023), i.e., giving up the educational pursuits because they somehow come to believe that they lack a decent aptitude for foreign languages (Genç et al., 2016). Thereby, ongoing investigations on learning efficacy should be generally conducted to leverage learner motivations and efforts, especially among struggling students, to stay firmly on track consistently until the desired goals are reached.

Methodology

Background and Participants

The present study was set in the current education context of Vietnam, where the demand for English learning and mastery is increasing, especially in tertiary education across the country (Do, 2024; Do et al., 2025; Ha et al., 2025; Truong & Luu, 2024). At the time of this study conducted (early 2025), the participants were 400 undergraduates (122 males and 278 females) majoring in EFL. They were learning at a university, but they came from several provinces in Vietnam. They included freshmen (122, 30.5%), sophomores (93, 23.25%), juniors (107, 26.75%), and seniors (78, 19.5%). Their ages ranged between 19 and 22 years old.

Questionnaire Building for Data Collection

In line with past research (Bandhu et al., 2024; Bong & Skaalvik, 2003; Chen, 2022; Chu et al., 2024; Do, 2022; Do et al., 2025; Gong & Xu, 2024; Mohebbi, 2025; Zimmerman et al., 1992), this study argues that learning regulations and resilience are two robust factors linking learning self-efficacy and academic achievement. In other words, learning regulation-resilience is a patient pathway from learning self-efficacy to academic achievement. This dictates that there likely exist specific positive correlations between the assumed regulation-resilience and learning efficacy. As such, to test these assumed correlations in a quantitative research survey, the researcher of this study designed a structured questionnaire scale based on the existing literature addressed earlier, with particular reference to Vietnam's current EFL education context. The principle of a clear, parsimonious scale was carefull considered. Thus, the questionnaire encompassed three subscales: one measuring learning self-efficacy, the second measuring learning regulation, and the third measuring resilience. Each subscale was formatted by a 5-point range (i.e., *Really Unsure* = 1 point, *Unsure* = 2 points, *Almost Sure* = 3 points, *Sure* = 4 points, *Really Sure* = 5 points). It is

assigned in this study that scores between 1 and < 3.00 pts are rated as a low level; from 3.00 to <4.00 is average, while the range between 4.00 and 5.00 is a high level.

First, the initial questionnaire was double-checked independently by three EFL experts to ensure content validity. Then, it was piloted among 100 students from the target population to secure the questionnaire's reliability. Thereby, drawing from the feedback and making several adjustments to wording, item numbers, and ordering, the final questionnaire scale comprised 17 items as follows:

Table 1. *The questionnaire scale*

Subscales	17 Items	Exemplar items
Learning Self-Efficacy (LSE)	5	I strongly believe in my ability to learn English well. I can do well on all the English tasks assigned by instructors.
Learning Regulation (LR)	6	I can make detailed plans to reach my learning goals. I constantly monitor my learning progress to make necessary modifications.
Resilience (RE)	6	I will make greater efforts in the face of learning challenges. I will keep learning until the desired outcomes are achieved despite setbacks or failures.

Collecting and Processing the Data

Having been permitted by the university authorities, the researcher made appointments to visit several EFL classes at this university. In each class, details on the purpose of the survey were provided. Additionally, information about the questionnaire contents was clarified. It was emphasized that participation in the survey was optional, and the results would be reported anonymously. This meant that no personal information about the participants was made known in any case. Then, the students received and completed the questionnaire within approximately 15 minutes. All completed questionnaires were collected on the spot.

After that, the researcher examined every single returned questionnaire to make sure that they had sufficient information for further analysis. Then, for processing these quantitative data on the responses to the overall scale and subscales, the popular software SPSS 26 (Statistical Package for Social Sciences) was used. The results are reported in the next section.

Results

Questionnaire Scale Reliability and Validity Values

First, a reliability analysis test was run by SPSS. The statistical result for the total scale is Cronbach's Alpha = 0.888. Meanwhile, the three subscales of LSE (learning Self-efficacy), LR (learning regulation), and RE (resilience) are 0.730, 0.864, and 0.803, respectively. These values are above the threshold level of 0.700 (Hair et al., 2019); thus, the total scale and subscales are appropriate for further analysis and result interpretations.

Then, for validity, the confirmatory factor analysis (SPSS) was applied. The value indicated by the Kaiser-Meyer-Olkin Measure of Sampling is 0.889 adequacy and Bartlett's sphericity test has a significance level of <.000. Additionally, as seen in Table 2 of factor analysis, the results show that 17 items are loaded convergently on their estimated factors only. All loading values are good, above 0.500 (Hair et al., 2019/2020). These together ensure the validity of the questionnaire scale used in the present study.

Table 2. *Factor loadings*

17 Items	LR	RE	LSE
LR1	.835		
LR2	.823		
LR3	.750		
LR4	.639		
LR5	.607		
LR6	.592		
RE1		.766	
RE2		.732	
RE3		.691	
RE4		.628	
RE5		.558	
RE1		.551	
LSE1			.737
LSE2			.692
LSE3			.647
LSE4			.609
LSE5			.577

Students' Levels of Learning Self-efficacy, Regulation, and Resilience

Table 3 displays the results of 400 participants involved in the present study. As addressed above, The participants responded to each item by rating their score levels on LSE, LR, and RE factors via the questionnaire scale. The descriptive statistics are as follows:

Table 3. *Descriptive statistics (N=400 participants)*

Subscales	Minimum	Maximum	Mean	Std. Deviation
LSE	1.40	5.00	3.78	0.56
LR	1.00	5.00	3.59	0.78
RE	2.00	5.00	4.38	0.54

It, therefore, shows that among the three factors, RE tops the list with the mean score of 4.38 (out of 5 points); i.e., standing out at a high level with the smallest Std. deviation: 0.54. Meanwhile, LSE and LR are rated 3.78 and 3.59, respectively. As a result, the scores for these two factors are just at the upper-average level (in the range of 3.00 < 4.00 points).

Correlations between Learning Self-efficacy, Regulation, and Resilience

To identify the assumed correlations between learning self-efficacy, regulation, and resilience among the participants involved in the present study, the Pearson correlation test was conducted. Table 4 below presents the results:

Table 4. *Correlation values between LSE, LR, and RE*

		LSE	LR
LR	Pearson Correlation	.443**	
	Sig. (2-tailed)	.000	
RE	Pearson Correlation	.488**	.582**
	Sig. (2-tailed)	.000	.000

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

Thus, it shows clearly that learning self-efficacy (LSE) has significantly positive correlations with both learning regulation (LR) and resilience (RE) at the values of 0.443** and 0.488**, respectively. This indicates that when LR and RE increase, there is a robustly corresponding increase in the LSE. In addition, LR and RE are also positively correlated at a high level of 0.582**.

This result has yet to be reported in the past (to our best knowledge), although some relevant works like Chu et al. (2024) and Duan et al. (2024) did examine RE.

The Impact of Academic Years' Experience on Learning Self-efficacy

Estimating this impact, the researcher repeatedly ran the Independent-Samples T-tests among the four groups of freshmen (I), sophomores (II), juniors (III), and seniors (IV). In other words, score means of learning self-efficacy (LSE) among the four groups are cross-compared respectively to determine whether the estimated differences are statistically significant. Table 5 shows the statistical results.

Table 5. *The score means on the LSE in comparison*

Pairs	Means	Std. Deviation	Sig. (2-tailed)
IV	3.82	.582	0.657
III	3.78	.620	
IV	3.82	.582	0.438
II	3.79	.438	
IV	3.82	.582	0.296
I	3.76	.563	
III	3.78	.620	0.176
II	3.79	.438	
III	3.78	.620	0.107
I	3.76	.563	
II	3.79	.438	0.721
I	3.76	.563	

Table 5 records that all score means across the four groups are under 4 points, ranging from 3.82 (IV) by seniors to 3.76 (I) by the first-year students. In between are sophomores (M=3.79) and juniors (M=3.78). The Std. deviations within groups are not large, all under 0.630. So, the variances in learning self-efficacy across the groups and within each group appear small. That is why all Sig. (2-tailed) values are greater than 0.100, which indicates no statistically significant difference across the four groups in learning self-efficacy (Hair et al., 2019).

Discussion and Implications

First, as seen in Table 3, this finding of EFL students' upper-average level of learning self-efficacy is particularly in line with past studies like Genç et al. (2016), Lestari et al. (2020), Lin et al. (2025), Truong and Luu (2024), and Wang et al. (2013). In other words, there is still room for students across countries and contexts to cultivate their efficacy in English learning (and language skills of listening, speaking, reading, and writing in particular) for improved academic performances. However, the result of a high level of resilience self-rated by participants in this study somehow differs from the finding by Duan et al. (2024), who reported a moderate-to-high level of English learning resilience among high schoolers from China. This is partly because college English-majored students could have gained more mastery experience in dealing with difficulties and challenges in English learning than their high school counterparts.

Secondly, the findings laid out in Table 4 underscore that the three factors of learning self-efficacy, learning regulation, and resilience jointly play their roles in the process of learning a foreign/second language in general terms. As a result, it opens more doors for both teachers and students to cultivate learning self-efficacy, which is believed to be the safeguard against learning burnout and the predictor of academic performance (Bandura, 1986; Lin et al, 2025; Luo et al.,

2023; Mohebbi, 2025; Waddington, 2023). Accordingly, it makes sense that teachers are advised not only to instruct learners to exercise their learning autonomy consistently but also to guide them to overcome challenges, stresses, and setbacks. Under the ample teacher's social persuasion and others, learners should be positively aware that unfavorable learning conditions are just like the practical tests for their resilience and efficacy, rather than an inevitable halt to the learning process underway. Furthermore, the three factors of LSE, LR, and RE are believed to benefit learners not just in general learning and EFL particularly but across various aspects of life because they are globally transferable skills.

Thirdly, statistical analyses presented in Table 5 suggest that academic years' experience did not significantly impact learning self-efficacy among EFL majors, i.e., their learning-efficacy level did not change remarkably from the first to final year during the undergraduate learning journey. This finding is somehow aligned with the result reported by Lin et al. (2025), which indicated that there was no significant difference in English-speaking self-efficacy among freshmen, sophomores, and juniors. One possible explanation for this point is that when moving up to the next learning years (i.e., higher levels), students mostly have to deal with more demanding coursework and assignments (Truong & Luu, 2024; Waddington, 2023). Thus, strategies to increase student learning self-efficacy associated with learning regulation and resilience are essential to EFL majors regardless of their academic years of experience. The ultimate goal is to assist students, especially the struggling ones, to confidently sustain the learning journey until they reach the final end.

Conclusion

The present study has three primary goals. First, it assessed the target students' learning self-efficacy, which reached only an upper-average level. Secondly, testing the correlations between learning self-efficacy, learning regulation, and resilience shows that the three factors are positively correlated. This dictates that they jointly impact the learning process in a positive direction. Thirdly, academic years' experience at college appeared not to increase students' learning self-efficacy substantially. These findings provide more insights into the complexity of those factors impacting the learning process and EFL learning in particular. More importantly, it calls for teachers' ongoing strategies in specific contexts to elevate student learning self-efficacy in conjunction with learning regulation and resilience. Despite its strengths in the findings, the present study is open to further research within EFL settings and beyond because the three factors under discussion are shared across learning disciplines and countries. Different strategies to positively impact these factors are also encouraged for trials to benefit learners as much as possible.

Funding: This research is not funded.

Acknowledgment: 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 text readability and language. The author reviewed and edited the content as needed. The author takes full responsibility for the content of the publication.

About the Author:
Dr. Do Minh Hung has been teaching English at Dong Thap University, Cao Lanh City, Dong Thap Province, Vietnam, for more than 30 years. He holds a doctoral degree in Linguistics and has published several journal articles, both locally and internationally, in the areas of English instruction.
https://orcid.org/0000-0001-9140-8404

References

Akanni, A. A., & Oduaran, C. A. (2018). Perceived social support and life satisfaction among freshmen: Mediating roles of academic self-efficacy and adjustment. *Journal of Psychology in Africa, 28*(2), 89–93. 10.1080/14330237.2018.1454582

Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management, 38*(1), 9-44. http://doi.org/10.1177/0149206311410606

Bandhu, D., Mohan, M.M., Nittala, N.A., Jadhav, P., Bhadauria, A., & Saxena, K.K. (2024). Theories of motivation: A comprehensive analysis of human behavior drivers. *Acta Psychologica, 244*, 104177, 1-16. https://doi.org/10.1016/j.actpsy.2024.104177

Bong, M., & Skaalvik, E.M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, *15*(1), 1–40.

Chen, J. (2022). The effectiveness of self-regulated learning (SRL) interventions on L2 learning achievement, strategy employment, and self-efficacy: A meta-analytic study. *Frontiers in Psychology*. 13,1021101, 1-17. http://doi.org/10.3389/fpsyg.2022.1021101

Chu, W., Yan, Y., Wang, H., & Liu, H. (2024). Visiting the studies of resilience in language learning: From concepts to themes. *Acta Psychologica 244*, 104208, 1-11. https://doi.org/10.1016/j.actpsy.2024.104208

Do, M.H. (2022). EFL students' self-regulated learning strategies during the COVID-19 pandemic. *Arab World English Journal (AWEJ) 2nd Special Issue on Covid-19 Challenges* (2) 22-34. https://dx.doi.org/10.24093/awej/covid2.2

Do, M.H. (2024). Cultivating EFL students' emotional domain via translation class. *Pakistan Journal of Life and Social Sciences*, *22*(2), 2011-2022. https://pjlss.edu.pk/pdf_files/2024_2/2011-2022.pdf

Do, M.H., Le, T.N.A., Vo, P.T.N., Pham, V.T., & Bui, T.T. (2025). Measuring Vietnamese-speaking English as a foreign language students' socio-emotional skills. *International Journal of Evaluation and Research in Education (IJERE), 14*(1), 739-748. http://doi.org/10.11591/ijere.v14i1.30099

Duan, S., Han, X., Li, X., & Liu, H. (2024). Unveiling student academic resilience in language learning: a structural equation modeling approach. *BMC Psychology*, *12*(177), 1-12. http://doi.org/10.1186/s40359-024-01665-1

Genç, G., Kuluşaklı, E., & Aydın, S. (2016). Exploring EFL learners' perceived self-efficacy and beliefs on English language learning. *Australian Journal of Teacher Education, 41*(2), 53-68. https://doi.org/10.14221/ajte.2016v41n2.4

Gong, W., & Xu, C. (2024). Teacher support as predictors of Chinese EFL learners' classroom flow: the mediating role of academic self-efficacy. *Frontiers in Psychology*, *15*, 1452146. http://doi.org/10.3389/fpsyg.2024.1452146

Ha, T.T., Hien, N.T.T., An, N.T., & Huong, N.T.M. (2025). Enhancing the quality of English language education for students at art universities in Vietnam through a competency-based approach. *Forum for Linguistic Studies. 7*(1), 1–14. https://doi.org/10.30564/fls.v7i1.7739

Hair, J.F., Black, W.C., Anderson, R.E., & Babin, B.J. (2019). *Multivariate data analysis* (8th ed.). London: Cengage Learning.

Hair, J.F., Howard, M.C., & Nitzl, C. (2020). Assessing measurement model quality in PLS-SEM using confirmatory composite analysis. *Journal of Business Research, 109*, 101–110. https://doi.org/10.1016/j.jbusres.2019.11.069

Huang, W. C. (2024). Influence of enhanced self-efficacy on English performance through English medium instruction: a mixed-methods research. *Asian-Pacific Journal of Second and Foreign Language Education, 9*, Article 32.https://sfleducation.springeropen.com/articles/10.1186/s40862-024-00255-3

Lestari, Z. W., Mujiyanto, J., Saleh, M., & Yusuf, S. (2020). Students' self-efficacy in learning English: A case study at a vocational high school. *Advances in Social Science, Education, and Humanities Research, 443*, 418-425.

Li, J., Wang, C., Zhao, & Y., Li, Y. (2023). Boosting learners' confidence in learning English: Can self-efficacy-based intervention make a difference? *TESOL Quarterly, 58*(4), 1518-1547. https://onlinelibrary.wiley.com/doi/10.1002/tesq.3292

Lin, Y., Mohamad, M., & Mahmud, M.I. (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://dx.doi.org/10.24093/awej/vol16no1.4

Luo, Q., Chen, L., Yu, D., & Zhang, K. (2023). The mediating role of learning engagement between self-efficacy and academic achievement among Chinese college students. *Psychology Research and Behavior Management,16,* 1533–1543.

Ma, X., Zhang, H., Zhou, X., & Bo, L. (2024). Investigating the learning self-efficacy of a fully online teaching environment among undergraduate Chinese medical students: a cross-sectional study. *BMC Medical Education, 24*(912), 1-11.

Maier. A., Schaitz, C., Kröner, J., Berger, A., Keller, F., Beschoner, P., Connemann, B., & Sosic-Vasic, Z. (2021) The association between test anxiety, self-efficacy, and mental images among university students: Results from an online survey. *Frontiers in Psychiatry 12*, 618108. https://bmcmededuc.biomedcentral.com/articles/10.1186/s12909-024-05890-5

Mahmud, A., Susilowati, N., & Sari, P.N. (2024). Assessment methods and learning styles for learning satisfaction: the role of self-efficacy. *International Journal of Evaluation and Research in Education (IJERE), 13*(6), 3752-3762. http://doi.org/10.11591/ijere.v13i6.29757

Mohebbi, A. (2025). Enabling learner independence and self-regulation in language education using AI tools: a systematic review. *Cogent Education, 12*(1), 2433814, 1-18. https://doi.org/10.1080/2331186X.2024.2433814

Nie, Y., Lau, S., & Liau, A.K. (2012). Role of academic self-efficacy in moderating the relation between task importance and test anxiety. *Learning and Individual Differences, 21*(6), 736–741.

Polrak, M., Kedcham, A., & Chansri, C. (2023). Measuring properties of an English language self-efficacy scale for EFL undergraduate students in Thailand. *NIDA Journal of Language and Communication, 28(*43), 68–91.

Schunk, D.H., & DiBenedetto, M.K. (2016). Self-efficacy theory in education. In K.R. Wentzel & D.B. Miele, *Handbook of motivation at school* (pp. 34-54). Routledge.

Torres, M.E, García, W.C, Sanchez, L.Z, García, S.B, Lozada, O., Sucapuca, F.E., & Aranzábal, D.F. (2023) Psychometric properties and invariance of an English self-efficacy scale for university students in Peru. *Frontiers in Psychology, 14*, 1187342. http://doi.org/10.3389/fpsyg.2023.1187342

Truong, T.N.N., & Wang, C. (2019). Understanding Vietnamese college students' self-efficacy beliefs in learning English as a foreign language. *System, 84*, 123-132, https://doi.org/10.1016/j.system.2019.06.007

Truong, X.M., & Luu, T.M.V. (2024). Identifying sources of English-speaking self-efficacy among Vietnamese EFL university learners. *AsiaCALL Online Journal, 15*(1), 105-116. https://doi.org/10.54855/acoj.241517

Van-Zyl, L., Klibert, J., Shankland, R., See-To, E., & Rothmann, S. (2022). The General academic self-efficacy scale: Psychometric properties, longitudinal invariance, and criterion validity. *Journal of Psychoeducational Assessment, 40*(6), 777–789. http://doi.org/10.1177/07342829221097174

Waddington, J. (2023). Key concepts in ELT: Self-efficacy. *ELT Journal, 77*(2), 237-240. https://doi.org/10.1093/elt/ccac046

Wang, C., Schwab, G., Fenn, P., & Chang, M. (2013). Self-efficacy and self-regulated learning strategies for English language learners: Comparison between Chinese and German college students. *Journal of Educational and Developmental Psychology, 3*(1), 173-191.

Xu, M., Wang, C., Chen, X., Sun, T., & Ma, X. (2022). Improving self-efficacy beliefs and English language proficiency through a summer intensive program. *System, 107*(7) https://doi.org/10.1016/j.system.2022.102797

Xu, J., & Wang, Y. (2025). Variations in perceived self-efficacy in information literacy and their impacts on self-regulated English writing strategies. *Journal of Librarianship and Information Science.* https://doi.org/10.1177/09610006251320090

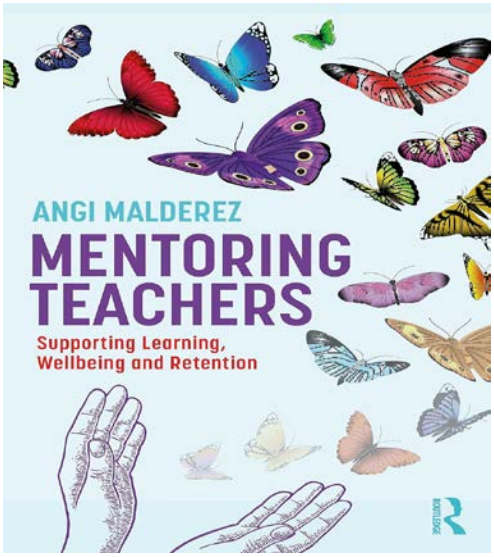
Zhang, L., Jiang, Y., & Chen, S. (2023). Longitudinal interrelations among self-efficacy, interest value, and effort cost in adolescent students' English achievement and future choice intentions. *Contemporary Educational Psychology,* 102176. https://doi.org/10.1016/j.cedpsych.2023.102176

Zimmerman, B.J., Bandura, A., & Martinez-Pons, M. (1992) Self-motivation for academic attainment: the role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 29*, 663– 676.

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

Arab World English Journal (AWEJ) Volume 16. Number 3, September 2025 Pp. 472-475
DOI: https://dx.doi.org/10.24093/awej/vol16no3.29

Book Review
Mentoring Teachers: Supporting Learning, Wellbeing and Retention



Title: **Mentoring Teachers: Supporting Learning, Wellbeing and Retention**
Author: **Angi Malderez**
Year of Publication: **2024**
Place of publication: **United Kingdom**
Publisher: **Routledge**
Pages:**149**
Reviewer: **Hamdan F. Alzahrani**

How to cite: Alzahrani, H. F. (2025). Mentoring Teachers: Supporting Learning, Wellbeing and Retention. By Malderez, A. *Arab World English Journal, 16*(3): 472-475. https://dx.doi.org/10.24093/awej/vol16no3.29

Mentoring Teachers: Supporting Learning, Wellbeing and Retention presents a form of nonjudgmental mentoring that is supported by both practical and theoretical evidence. The author puts forward that mentoring must be non-judgmental and based on trust and understanding where mentees are scaffolded throughout, with mentors taking the five roles of a supporter, an acculturator, a model, a sponsor and an educator to help mentees develop as informed reflective practitioners, who are confident and capable of making their own decisions and judgements.

The book has ten chapters preceded by an introduction. The author starts her book with some clarifications of her views of teaching and mentoring, and how she sees her professional identity as a teacher, a mentor as well as a researcher in the area. The first few pages set the scene for a reader-friendly book by avoiding jargon and through an accessible writing style. One of the key ideas presented in the introduction, and that is referred to frequently in the book is Systematic Informed Reflective Practice (SIRP). Another core idea that the author asserts is that mentors need to analyse the features of mentoring in their contexts. Context is defined in this book as "People, Place and (point in) Time" (Wedell and Malderez, 2013, p.15). According to the author, the contextual analysis can help mentors identify features of their contexts that can facilitate or impede mentoring.

After preparing readers for the five roles of a mentor by elucidating the conceptual meaning of mentoring in this book, and by considering context, the author discusses each role in a sperate chapter. The book shows how a mentor can support a mentee throughout a mentoring programme. Active listening is a key skill that a mentor needs to possess to effectively support mentees. This skill is highlighted by the author where she draws on the levels of listening in coaching by citing van Nieuwerburgh (2020), and adding her own observations and commentaries to each level. One level that stands out is that a mentor needs to reflect back to what a mentee has said by paraphrasing and rechecking their understanding. It could be noticed that the nonjudgmental listening is a joint feature in mentoring and coaching as well as in Cooperative Development (CD) where a listener in CD is labeled as an understander to reflect the non-judgmental nature of such a role (see Edge,1992).

The author also discusses how introducing mentees to their new institutions and communities of practice is a vital role for mentors. The writer points out that the role of an acculturator is much more needed by new career teachers who need to know their micro context and its norms. Also, a beginning teacher needs to be introduced to local, national and international communities of practice that can help them further their professional learning. As for long serving teachers, if they are new to an institution, then they might need what a beginning teacher does with regard to introducing them to the immediate context and its norms as well as to local communities of practice.

Besides the roles of a supporter and acculturator, the book sheds light on the model role of mentors, where they display the professional side of a teacher. The author suggests some ways for mentors to achieve that. One salient feature of this role emphasised in this book is that mentors need to show how they take care of their own ongoing learning and make their own decisions and plans visible to their mentees. For example, mentors can share with mentees their experiential knowledge about an aspect of their teaching, and mentors can talk their mentees through their decisions and informed practices. The author also considers how mentors should use their professional capital that stems mainly from their knowledge and professional contacts to facilitate the growth of a mentee. Sharing knowledge with mentees is a balancing act according to the author in that mentors need to provide their mentees with some knowledge, especially about methods and strategies to teach something in the right time. However, mentors should be aware of when they should refrain from doing so to allow a mentee to develop at their own pace and develop the ability and the skill to make their own professional judgments and decisions.

Special attention in the book is given to the educator role of mentors, where they can scaffold mentees to develop into informed reflective practitioners. The author proposes a model of mentoring that can be used in the post-lesson mentoring meetings to enhance a mentee's SIRP,

where ideally a mentor has not observed a mentee. The five-step model has some distinctive features such as it is led by the mentee and the mentor takes the role of a facilitator of reflection and helps the mentee to refine his or her own understanding of successful, challenging or puzzling experiences. The mentor also helps mentees to develop their noticing skills by asking some questions when invited to do so. The mentor is not judgmental and should listen actively and if asked, he or she could add some ideas. Mentees need to reach their own informed conclusions and plans for future interventions based on the process and outcome of the post-lesson mentoring meeting. In the educator role, the author also deals with the pre-lesson mentoring which focuses on scaffolding mentees by means of a flexible guide to help them plan their lessons.

After discussing the five roles of a mentor, the author problematises the assessment of mentees by highlighting challenges when mentors are tasked with such a role or when mentees are to be assessed by others (managers, supervisors, trainers). The author shows how mentors can support their mentees to navigate challenges associated with observation.

Given that mentoring requires knowledge, skills and learning from experience, mentors need to take care of their ongoing learning. The author recommends that mentors use the framework of SIRP for their own learning. They can do that on their own by asking themselves questions relating to each step in the model. Also, they can do that with another mentor and/or with a group of mentors or even with their mentees. However, the author makes some caveats when doing so with mentees. The author also encourages mentors to engage with research-based ideas which have useful practical implications, such as scaffolding, noticing, reflective practice. As springboard for teacher learning, the book highlights both real life stories related to teaching and mentoring, and allegorical stories which can be a source of thrill and learning. As for the latter, the author presents three allegorical stories and invites readers to get the learnt lessons pertaining to mentoring. The author leaves readers with a positive note of three successful stories for three mentors where such stories reflect that mentoring can be rewarding for the profession and mentors themselves despite the challenges that can encounter mentors.

All things considered, this book has many strengths. For instance, the author conceptualises mentoring based on her wide experience as a teacher, a mentor and as a well-known researcher in the area of mentoring. This is evident throughout the book where she vividly reflects on her own teaching and mentoring experiences, and also refers to some research where needed to theorise such experiences. Besides, the book provides insights and practical ideas to mentors of both new career teachers and experienced teachers, where many comparisons of how to mentor them are explicit in most chapters. On the other hand, while the author compares mentoring and coaching briefly in a few paragraphs at the beginning of the book, I think, as a reader, the relationship between mentoring and coaching is worth more discussion, and hence more space (perhaps a chapter). This is so for two reasons. Firstly, the author mentions that she hopes that her book will play a role in rebuilding the reputation of mentoring, as it is badly impacted by judgmental mentoring, and which , according to the author, explains why some have opted for coaching, and it also justifies the addition of coaching to mentoring. Secondly, mentoring and coaching as terms are almost always used adjacently, and some key researchers in the area provide an encompassing definition of both (see Maxwell et al, 2022). All in all, this book is a valuable source for mentors and other professionals who want to have an idea of supporting, acculturating, sponsoring, modelling, and educating teachers.

Reviewer:
Dr. Hamdan F. Alzahrani
King Abdulaziz University, Jeddah, Saudi Arabia

References

Edge, J. (1992). "Cooperative Development." *ELT Journal*, 46(1), 62–70. https://doi.org/10.1093/elt/46.1.62

Maxwell, B., Hobson, A.J., and Manning, C., (2022). Mentoring and coaching trainee and early career teachers: conceptual review. The National Institute of Teaching.

van Nieuwerburgh, C. (2020). *An Introduction to Coaching Skills: A Practical Guide*. London: Sage

Wedell, M., & Malderez, A. (2013). *Understanding Language Classroom Contexts: The Starting Point for Change*. Bloomsbury Academic.