



Memrise and the Pursuit of Proficiency: A Phenomenological Study on Students' Experiences and Independent Learning Strategies for TOEFL Vocabulary

Dimitria Rahayu¹, Awwalia Fitrotin Izza²

^{1,2}Pendidikan Bahasa Inggris, Universitas Islam Negeri Walisongo Semarang, Jawa Tengah

Corresponding E-Mail: dimitria.shadika@gmail.com

Received: 2025-12-06 Accepted: 2025-12-18

DOI: 10.24256/ideas.v13i2.8711

Abstract

This phenomenological study investigated the lived vocabulary learning experiences of seven high-stakes TOEFL candidates (N=7) who used the Memrise Spaced Repetition System (SRS) app. Through Interpretive Phenomenological Analysis (IPA), the study confirmed that while gamified features offer strong structural support for consistency and efficient recognition-based memorization, the learning process is structurally incomplete within the app. The central finding is a persistent recognition–production gap, meaning vocabulary learned passively in the app fails to transfer directly to the active, contextual production required by the TOEFL exam. Success, therefore, hinged on learners' active, metacognitive effort and the implementation of external compensatory strategies, such as dedicated sentence journaling and custom course creation. These findings suggest the motivational benefits of gamification are conditional and can risk a focus on volume over learning quality. Results have significant implications for Mobile Assisted Language Learning (MALL) design, calling for integrated features that foster contextual application. Ultimately, this work benefits app designers, language teachers, and autonomous learners by specifying the structural and metacognitive requirements for effective vocabulary learning via SRS applications.

Keywords: *Mobile Assisted Language Learning (MALL); Spaced Repetition System (SRS); Vocabulary Acquisition; TOEFL; Gamification; Metacognition*

Introduction

In today's interconnected world, English stands out as a key international language, serving as the primary medium for conveying information across science, technology, and culture. The demand for English proficiency, particularly demonstrated through high-stakes tests like the Test of English as a Foreign Language (TOEFL), remains a critical benchmark for global academic and professional advancement. A significant barrier to achieving a high TOEFL score is the vast, often specialized, vocabulary required for success across the integrated skills of reading, listening, and writing. Traditional classroom instruction, especially in the Indonesian higher education context, often struggles to provide the sheer volume of exposure and sustained practice necessary for deep vocabulary acquisition and retention. Consequently, students rely heavily on structured, intensive preparation methods to achieve the necessary score improvements (Rifiyanti et al., 2023).

This preparation challenge has spurred a critical reliance on Mobile-Assisted Language Learning (MALL) applications. MALL offers flexibility and personalized content, yet undergraduate students often struggle with self-regulation, maintaining consistent study habits, and effectively integrating digital tools into their autonomous learning strategies (Ma & Chiu, 2024; Netta & Trisnawati, 2019). The efficacy of these tools, therefore, hinges not only on their technical features but also on how students manage cognitive challenges, such as vocabulary memorization, in an unsupervised environment. This gap between the readily available technology and effective student implementation forms the fundamental problem this research addresses.

Recent scholarly research has extensively explored the impact of digital tools, particularly those utilizing gamified spaced repetition systems (SRS), on vocabulary acquisition. Investigations into applications like Memrise, Quizlet, and Anki consistently affirm their value in bolstering lexical proficiency by employing spaced repetition, which leads to improved word recall and learner motivation (Agung Cahyono et al., 2023; Baniara et al., 2024). Specific quantitative studies focusing on Memrise highlight its gamification features, such as points and streaks, as significant boosters of students' intrinsic motivation and engagement in self-directed learning (Deputri et al., 2023). Comparative research further suggests that Memrise usage leads to notable improvements in vocabulary breadth and retention (Nguyen et al., 2023), often outperforming conventional paper-based approaches.

Nonetheless, the limitation of this extensive literature lies precisely in its positivist methodology. The bulk of existing research emphasizes quantifiable results, such as pre- and post-test scores or feature evaluations, which confirms if and how much students learn. While the literature affirms the effectiveness of SRS and MALL, it often treats the student's process as a black box, providing performance metrics without fully unpacking the complex, nuanced psychological and practical realities of independent learning (Zohoorian et al., 2022).

The discernible gap, therefore, lies at the intersection of effective digital tools and the subjective, lived experience of the learner in a high-stakes context. We know that students encounter both internal and external obstacles in digital learning and strategically combine resources (Muryani & Yunus, 2024), but the existing body of knowledge lacks phenomenological depth. The current state-of-the-art fails to deeply unpack the specific independent learning strategies students develop, the meaning they ascribe to the pursuit of proficiency, or how they negotiate their autonomy when preparing for an externally mandated, high-level examination like the TOEFL using a highly structured application.

This study argues that the missing link is a detailed understanding of the strategies, challenges, and meanings students construct while using Memrise. A purely quantitative analysis fails to capture the internal and external pressures of TOEFL preparation, which are critical to understanding success or failure in independent learning. This gap necessitates a shift towards a qualitative, interpretive paradigm, specifically, a phenomenological approach, to uncover the essence of this unique learning experience in the higher education setting.

Based on the identified gap, this study poses the following research question: What is the essence of university students' lived experiences and the independent learning strategies they employ when using the Memrise application to improve their TOEFL vocabulary?. The primary objective of this research is to conduct an in-depth, phenomenological exploration to understand the nature of students' challenges, successes, and autonomous learning behaviors.

The novelty of this research is twofold: (1) It uniquely combines the technological context of Memrise and the high-stakes outcome of TOEFL preparation within a single study; and (2) it employs a phenomenological methodology to move beyond effectiveness metrics and provide rich, thick descriptions of the meaning students attribute to this independent digital learning journey, offering foundational insights for both pedagogical practices and future MALL app design.

Method

This research employed a qualitative, interpretive phenomenological design. The choice of phenomenology was directly aligned with the research question, which sought to understand the "essence" of university students lived experiences when engaging with the Memrise application for high-stakes TOEFL vocabulary preparation. Unlike quantitative approaches that measure causality or frequency, phenomenology aims to describe, understand, and interpret the meanings individuals attribute to a particular phenomenon (in this case, digital independent vocabulary learning under high-stakes pressure). The design focused on gathering rich, detailed descriptions to uncover the common, core structure of the participants' experiences, thereby revealing the *how* and *what* of their learning journey.

To strengthen qualitative rigor, researchers acknowledged their position throughout the study. Researchers had backgrounds in applied linguistics and were familiar with the MALL application, which allowed for informed questioning but required the suspension of existing assumptions to avoid influencing participants' narratives. Ethical approval was obtained prior to data collection. All participants provided oral consent after receiving a detailed explanation of the research objectives, procedures, and their right to withdraw at any time. To ensure anonymity and confidentiality, all identifying information was removed during transcription, and pseudonyms were used in all research data and publication materials.

The subjects of this research were selected using purposive sampling. This technique was chosen to ensure that all participants had directly and deeply experienced the phenomenon being studied. The target population consisted of university undergraduate students who met specific inclusion criteria: (1) actively using the Memrise application specifically for improving their English vocabulary related to the TOEFL test; (2) having used the application consistently for a minimum period of at least three months to ensure they had established a routine and encountered challenges; and (3) actively preparing for or intending to take the TOEFL examination. A small sample size of seven participants (N=7) was ultimately recruited, which is standard practice in phenomenological studies, allowing for in-depth engagement and the generation of rich data from each individual.

The primary data collection technique was in-depth, semi-structured interviewing. According to Sugiyono (2019), a semi-structured interview is designed to discover problems more openly, allowing the interviewee to express their opinions and ideas freely. Interviews were the cornerstone of this phenomenological research, as they allowed participants to articulate their experiences in their own words, providing the nuanced details necessary for interpretation.

An interview guide was developed, framed around the research objectives to explore three main areas: (a) the participants' experiences (e.g., initial impressions, satisfaction, emotional responses to gamification); (b) the challenges they faced (e.g., retention issues, technical difficulties, lack of motivation); and (c) the specific independent learning strategies they employed (e.g., customization of courses, use of Memrise time management).

All interviews were conducted face-to-face in a quiet, private setting on campus and averaged 60–90 minutes in duration. All sessions were audio-recorded and fully transcribed verbatim. To enhance the trustworthiness of the data, a secondary collection technique was employed: participants were asked to provide reflective journals or screen recordings of their typical Memrise sessions. This supplementary data served to contextualize and triangulate the verbal accounts provided during the interviews, ensuring a deeper understanding of the phenomenon as it unfolded in the students' natural learning environment.

The transcribed interview data were analyzed using Interpretive Phenomenological Analysis (IPA). The analysis followed the systematic steps commonly outlined in phenomenological research, such as those detailed by Moustakas (1994) in Figure 1. The process began with immersion, where the researcher read and reread the transcripts multiple times to achieve a deep understanding of the participants' narratives. Next, the data was broken down into significant statements or meaning units that related directly to the phenomenon of interest.

These statements were then transformed into formulated meanings or themes (a process Moustakas termed *horizontalization*). The formulated meanings were subsequently clustered into emergent themes that reflected common patterns, strategies, or challenges shared across the participants. The final step involved synthesizing these themes to construct an essential structure—a detailed, comprehensive description of the core experience (the essence) of using Memrise for TOEFL vocabulary in the context of independent higher education learning.

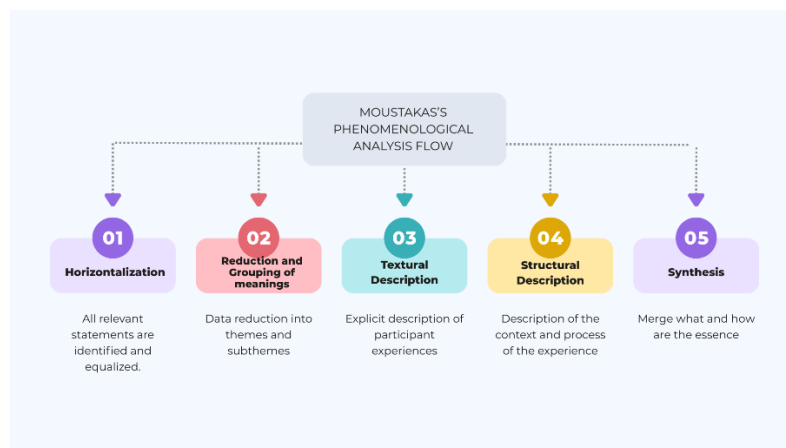


Figure 1. Moustakas's Analysis Steps

Results

The Interpretive Phenomenological Analysis (IPA) of the seven participants' narratives revealed three superordinate thematic clusters that capture the essence of their lived experience using the Memrise application for TOEFL vocabulary preparation. These clusters illuminate the tension between the app's structural efficiency and the learners' subjective, strategic efforts required for true proficiency.

The Duality of Motivation (System vs. Self)

This cluster describes how Memrise's internal mechanics, particularly its gamification and Spaced Repetition System (SRS), provided structural consistency but introduced external competitive pressures that sometimes conflicted with genuine academic goals.

Reliance on System Metrics for Consistency

Most participants (Participant 1, Participant 2, Participant 5, Participant 7) framed Memrise as a structured external force that successfully managed the logistics of learning. This reliance removed the cognitive burden of planning and sustained routine.

Participant 1 described this structural benefit, *"It's my digital drill sergeant, it forces consistency..."*

Similarly, Participant 7 trusted the internal organization, *"I trust its SRS algorithm to tell me what to study next, so I don't have to plan the minutiae of studying."*

Motivation was derived from simple, quantitative metrics such as the streak counter (Participant 1) and the progress bar (Participant 2).

Conflict between Competition and Depth

While system metrics promoted routine, external competitive features introduced pressure.

Participant 4 was uniquely driven by the leaderboard and high points. This pursuit of quantity led to self-confessed compromises on learning quality, *"Sometimes I prioritize the answer needed to get the point, which is memorization."*

Conversely, for serious learners like Participant 5, these competitive features were dismissed, *"Completely neutral. I barely notice the points or levels. They are just numerical noise. My focus is purely on the word bank and the number of words pending review."*

Micro-Strategies for Fatigue

Participants developed specific strategies to maintain consistency when facing fatigue. Participant 1 and Participant 5 employed tactical changes in modality,

"If the typing reviews are making me tired, I switch to multiple-choice," (Participant 1)

or using the 'speed review' mode *"just to save the streak"* (Participant 5).

Participant 6 employed external means, *"I promise myself a small, external reward if I complete the session."*

The Recognition-Production Gap (The Learning Trap)

A core shared experience was the struggle to convert the passive recognition knowledge acquired in the app into the active production required by the TOEFL's writing and speaking sections.

The Shallow Knowledge Phenomenon

All participants acknowledged the memorization trap inherent in flashcard-based systems, where the app's testing format encouraged superficial knowledge without ensuring contextual understanding.

Participant 1 articulated this duality, *"For the complex, abstract vocabulary, I realize I'm often just memorizing the visual form... it's a shallow knowledge for about 30% of the words."*

Participant 6 admitted to compromising quality, *"I learn just enough to pass the app's test."*

Manual Compensation for Production

Highly motivated learners actively implemented manual, external strategies to force production, compensating for the app's structural limitations.

Participant 7, for example, developed a personal rule, *"If I get a word right five times in a row, I immediately write it down in a physical sentence-journal to force production."*

Similarly, Participant 5 utilized the learned vocabulary in external contexts, such as work emails or discussions, to test true knowledge.

Deficiency in Context and Collocation

Participants repeatedly criticized the app's inability to test for collocation (words that typically accompany the target word) and for isolating the vocabulary from natural language use.

Participant 6 noted the lack of transfer, *"I learn the words in isolation, and when I see them in a full paragraph in a TOEFL passage, they look unfamiliar."*

Participant 1 concluded, *"great for recognition, but poor for production. The words stay locked inside the app's ecosystem."*

The typing feature was seen as problematic (Participant 4), as it was

perceived as testing typing speed rather than vocabulary recall: *"The learning mode, where you have to type the word perfectly. That slows me down too much."*

Autonomous Management of Content and Modality

This cluster highlights the proactive strategies learners employed to manage the community-driven content and integrate the app into their diverse study schedules, demonstrating significant personal control over the digital material.

Customization as Ownership

Participants engaged in significant content selection and customization to gain control and ensure relevance.

Participant 3 found the strongest motivation in creating their own courses, viewing the app as a digital workshop, *"It's my digital workshop. It's where I deconstruct vocabulary and rebuild it in a format that my brain can absorb efficiently."*

Participant 7 and Participant 2 meticulously filtered community content, noting they *"don't trust just any community-made course; it has to be authoritative,"* (Participant 2).

Resistance to Distracting Features

Learners frequently bypassed features that interrupted their focus or lacked academic rigor.

Participant 1 and Participant 6 expressed frustration with the distraction of *"constant pushing of non-TOEFL related courses."*

Furthermore, Participant 5 and Participant 7 found the visual *memrise* distracting, preferring those that provided *"a specific linguistic rule about the word"* (Participant 5) or concise example sentences.

Varied Time Management

The integration of Memrise into study was highly individualized. Some relied on fixed, concentrated blocks, such as the 30 minutes scheduled,

"every morning after breakfast" (Participant 2)

In contrast, others embraced opportunistic, micro-learning, using the app strictly as a filler activity,

"on the bus, waiting in line, or the 10 minutes before bed," (Participant 1).

The frequent appearance of quotations from Participant 1, Participant 5, and Participant 7 reflects their high level of verbal articulation and self-awareness regarding their learning processes and metacognitive strategies, which is common among highly motivated, high-stakes language learners in qualitative research.

Theme	Challenge within the App	Learner's External Compensatory Strategy
The Duality of Motivation	Gamified metrics risk prioritizing quantity (points/streaks) over quality.	Switching review modality (e.g., from typing to multiple-choice) to manage fatigue while maintaining consistency.
The Recognition-Production Gap	App's format yields shallow, recognition-based knowledge that does not transfer to active use.	Manual sentence journaling and using the vocabulary in real-world contexts (e.g., work emails) to force production.
Autonomous Management of Content	Community-driven content lacks necessary authority, focus, or academic rigor.	Creating customized courses from scratch; Meticulously filtering and selecting only authoritative content.

Discussion

The findings of this phenomenological study reveal the complex, nuanced experience of autonomous language learning mediated by a gamified Spaced Repetition System (SRS) application. The core essence of this experience is defined by the student's continuous effort to negotiate and compensate for the application's inherent structural limitations within the high-stakes environment of TOEFL preparation. This discussion focuses on two core arguments: the conditional nature of gamification efficacy and the metacognitive effort required to bridge the recognition-production gap.

The Conditional Efficacy of Gamification

The results add phenomenological depth to prior literature (Hanus & Fox, 2015; Philpott & Son, 2022) by detailing *how* gamification is received. We found that the motivational value of gamification is not uniform; it shifts based on the learner's focus. System-internal metrics (like streak and SRS counters) were confirmed as a positive source of consistency (Participant 2, Participant 7). However, the external, competitive features (like leaderboards) were a potential source of pressure and shallow learning.

As Participant 4 indicated, the pursuit of points and ranks led to self-confessed compromises on learning quality, sometimes resulting in burnout. This outcome aligns with findings that competitive game elements function primarily as extrinsic incentives, effective for quantity but not necessarily for the quality of deep performance (Mekler et al., 2017). This suggests a critical implication for MALL design: the competitive overlay is primarily beneficial for learners with a highly external locus of control. For the majority focused on internal academic goals (Participant 3, Participant 5, Participant 7), the progress bar and internal mastering counts are more meaningful success metrics than external ranks, as they reinforce mastery over simple compliance.

The Recognition-Production Gap and Metacognitive Compensation

The most significant and widespread challenge identified is the application's inability to automatically facilitate the transfer of vocabulary from passive recognition (tested by the app's flashcard format) to active production (required by TOEFL Writing/Speaking). This challenge confirms the fundamental pedagogical gap between the technology's capability (efficient recognition-based memorization) and the student's high-stakes academic requirement (fluent application).

Our core finding is that successful independent learning relies heavily on the learner's metacognitive awareness to identify the app's weakness and implement structured, external countermeasures. This directly refutes the notion of autonomous learning as a purely self-contained, in-app process. When faced with the "shallow knowledge phenomenon," students actively developed compensatory strategies: the deliberate, manual acts of writing sentences (Participant 7, Participant 5) and creating custom content (Participant 3). As Participant 1 succinctly highlighted, *"The words stay locked inside the app's ecosystem."* This failure to facilitate transfer is consistent with critiques that gamified language learning tools often present "chocolate-covered broccoli," adding game elements to repetitive tasks without addressing the need for collaboration, context exploration, and complex language skills (Azzouz Boudadi & Gutiérrez-Colón, 2020; Luo, 2023). The manual bridging effort is the hidden curriculum of digital vocabulary preparation.

Theoretical and Practical Implications

This study offers two novel contributions to the understanding of Mobile-Assisted Language Learning (MALL) and Self-Regulated Learning (SRL):

1. Advancing MALL Theory: Beyond Opportunistic Learning

While micro-learning (Participant 1) provides flexibility, our findings suggest that mastery, especially in high-stakes contexts, necessitates dedicated, structured scheduling. The most successful learners (Participant 2, Participant 7, Participant 3) preferred fixed, concentrated blocks of time over fragmented micro-sessions, stating it helped "*establish a better focus*." This directly supports SRL theory, which links successful learning to the development of Goal Setting and Planning abilities (Alhalafawy & Tawfiq Zaki, 2022). Educators must, therefore, guide students toward structured integration rather than relying solely on notification-driven, opportunistic use.

2. Implications for App Design and Content Authority

The observed effort in selecting and customizing authoritative courses (Participant 2, Participant 3, Participant 7) highlights a critical deficiency in the community-driven content model. Future research should explore the development of integrated features that allow learners to immediately verify, cross-reference, and contextualize community *mems* and definitions with official academic usage (collocation). This shift in design focus—from merely encouraging *input* to systematically facilitating contextual application and output—is necessary to minimize the content-quality hurdles and truly support the advanced learner's goals.

Conclusion

This phenomenological study explored the lived experiences of high-stakes language learners utilizing the Memrise Spaced Repetition System (SRS) application for TOEFL vocabulary preparation. The findings confirm that while gamified technology provides strong structural support for consistency and efficient recognition-based memorization, the learning process remains fundamentally incomplete within the app's isolated ecosystem.

The central contribution of this research is the conceptualization of the learner's active, metacognitive effort required to negotiate the inherent recognition-production gap. Successful autonomous learning was shown to be heavily reliant on compensatory, often manual and external, strategies, such as dedicated sentence journaling and custom content creation, to bridge the divide between passive knowledge acquired through flashcards and the active, contextual production demanded by the TOEFL exam. Furthermore, the efficacy of gamification proved highly conditional: system metrics successfully supported routine, but competitive features risked motivating surface-level learning, emphasizing volume over true

academic quality.

These results hold significant implications for Mobile Assisted Language Learning (MALL) design. Future vocabulary applications must move beyond simple SRS functionality to integrate features that systematically foster contextual application and collocation competence. Educators, in turn, should shift their guidance from merely recommending digital tools to actively coaching students on the necessary external, production-focused strategies required to convert in-app memorization into true, usable academic proficiency.

References

- Agung Cahyono, M. F., Puspita Sari, H., & Primasari, Y. (2023). A SOLUTION TO THE DIFFICULTY OF VOCABULARY MASTERY THROUGH MEMRISE. *JARES (Journal of Academic Research and Sciences)*, 8(2), 10–16. <https://doi.org/10.35457/jares.v8i2.2658>
- Alhalafawy, W. S., & Tawfiq Zaki, M. Z. (2022). How Has Gamification Within Digital Platforms Affected Self-Regulated Learning Skills During the COVID-19 Pandemic? Mixed-Methods Research. *International Journal of Emerging Technologies in Learning (IJET)*, 17(06), 123–151. <https://doi.org/10.3991/ijet.v17i06.28885>
- Azzouz Boudadi, N., & Gutiérrez-Colón, M. (2020). Effect of Gamification on Students' Motivation and Learning Achievement in Second Language Acquisition within Higher Education: A Literature Review 2011-2019. *The EuroCALL Review*, 28(1), 40. <https://doi.org/10.4995/eurocall.2020.12974>
- Baniara, M., Halim, A., & Abidin, A. (2024). Students' Perceptions of Memrise App Use in Learning Vocabulary. *AL LUGHAWIYAAT*, 4(2). <https://doi.org/10.31332/alg.v4i2.4250>
- Deputri, A. B., Hamidiyah, A., & Baihaqi, A. (2023). THE USE OF THE MEMRISE APPLICATION ON STUDENTS' MOTIVATION IN LEARNING ENGLISH VOCABULARY. *GENTA BAHTERA: Jurnal Ilmiah Kebahasaan Dan Kesastraan*, 9(1). <https://doi.org/10.47269/gb.v9i1.243>
- Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education*, 80, 152–161. <https://doi.org/10.1016/j.compedu.2014.08.019>
- Luo, Z. (2023). The Effectiveness of Gamified Tools for Foreign Language Learning (FLL): A Systematic Review. *Behavioral Sciences*, 13(4), 331. <https://doi.org/10.3390/bs13040331>
- Ma, Q., & Chiu, M. M. (2024). Self-regulated and collaborative personalized vocabulary learning approach in MALL. *Language Learning & Technology*, 28(1), 1–27. <https://doi.org/10.64152/10125/73579>
- Mekler, E. D., Brühlmann, F., Tuch, A. N., & Opwis, K. (2017). Towards understanding

- the effects of individual gamification elements on intrinsic motivation and performance. *Computers in Human Behavior*, 71, 525–534. <https://doi.org/10.1016/j.chb.2015.08.048>
- Moustakas, C. (1994). *Phenomenological research methods*. SAGE Publications, Inc. <https://doi.org/10.4135/9781412995658>
- Muryani, R. N., & Yunus, M. (2024). STUDENTS' DIGITAL VOCABULARY LEARNING: EXPERIENCES, BARRIERS, AND ADAPTIVE STRATEGIES EXPLORATION. *English Review: Journal of English Education*, 12(1), 31–42. <https://doi.org/10.25134/erjee.v12i1.9328>
- Netta, A., & Trisnawati, I. K. (2019). ACEHNESE UNDERGRADUATE STUDENTS' STRATEGIES IN PREPARING FOR TOEFL PREDICTION: A PRELIMINARY STUDY. *Englisia Journal*, 7(1). <https://doi.org/10.22373/ej.v7i1.5779>
- Nguyen, A. T., Nguyen, T. T., Le, T. T., Phuong, H. Y., Pham, T. T., Huynh, T. A. T., & Nguyen, H. T. (2023). Effects of Memrise on Vietnamese EFL Students' Vocabulary: A Case Study at a College in a Rural Area. *Electronic Journal of E-Learning*, 21(5), 450–460. <https://doi.org/10.34190/ejel.21.5.3066>
- Philpott, A., & Son, J.-B. (2022). Leaderboards in an EFL course: Student performance and motivation. *Computers & Education*, 190, 104605. <https://doi.org/10.1016/j.compedu.2022.104605>
- Rifiyanti, H., Dewi, D. U., & Putra, F. M. (2023). The Role of TOEFL Preparation Courses in Improving Test Scores. *Foremost Journal*, 4(2), 99–104. <https://doi.org/10.33592/foremost.v4i2.3615>
- Sugiyono. (2019). *Metode Penelitian dan Pengembangan*. Bandung: CV. Alfabeta.
- Zohoorian, Z., Noorbakhsh, M., & Zeraatpisheh, M. (2022). EFL Learners' Vocabulary Achievement and Autonomy: Using the Memrise Mobile Application. *Indonesian Journal of EFL and Linguistics*, 7(2), 233. <https://doi.org/10.21462/ijefl.v7i2.487>