

AI-Enabled Translation Assistance for Preserving Meaning in Cross-Language Communication

Amany Alsamhan^{1*} and Fatemah Hamad Aldosari²

^{1*}Teacher Assistant, Department of English Language & Translation, Qassim University, Saudi Arabia.

²Lecturer, Department of Applied linguistics, Faculty of Languages and Translation, King Khaled University, Saudi Arabia.

Abstract: Artificial intelligence (AI) has been a significant asset in improving multilingual communication, allowing for accurate and efficient translation across a variety of languages. However, maintaining semantic meaning and the intent for communication in the context of a translation can be very difficult, especially in the field of education and cross-language communication. This study introduces an AI-Enabled Semantic Meaning Preservation Framework (AI-SMPF) for assessing AI-assisted translation with the parallel dataset of FLORES-200. The proposed framework is comprised of text preprocessing, AI-assisted translation, semantic representation, semantic similarity computation, meaning preservation evaluation, and translation quality evaluation. The framework measures semantic consistency not only based on lexical matching but also on the Semantic Similarity Score (SSS), Context Preservation Score (CPS), and the Translation Quality Index (TQI). These semantic evaluation metrics can also be used to facilitate automated feedback and guidance in education by objectively evaluating students' translation skills and context understanding, in multilingual learning activities. The experimental results show that the proposed framework obtains an SSS of 0.95, CPS of 0.93 and an overall TQI of 0.94, which shows that semantic meaning is preserved in multilingual translations. The statistical analysis shows that the average translation quality is 94.0%, with a standard deviation of 1.0%, which indicates that the quality of the translations is consistent across the different sentence pairs. The results show that using a semantic-aware evaluation is a more reliable method for evaluating AI-assisted translation than traditional lexical comparison-based methods. The proposed framework provides an efficient approach to maintaining the contextual meaning and facilitating multi-lingual communication in the field of English-language learning, which in turn helps to achieve more accurate AI-assisted translation for academic, educational and cross-cultural communication.

Key Words: artificial intelligence, ai-assisted translation, semantic meaning preservation, cross-language communication, FLORES-200 dataset, english language education

(Received: 17 September 2025; Revised: 30 October 2025; Accepted: 28 November 2025; Published: 31 December 2025)

Introduction

Artificial intelligence (AI) has revolutionized multilingual communication by offering smart translation and communication support that enhances accessibility, efficiency, and language education for different language communities (Munuku & Mwangi, 2024). With the advent of large language models and neural machine translation, the context in which the machine translates the text plays a more significant role, thus achieving better preservation of semantic meaning than traditional rule-based translation methods (Greńczuk et al., 2024). With various application fields being served by the use of AI technologies like software engineering, cyber security, education and information management, the field of supporting

meaning preserving cross language communication is still an important research topic (Kohnke et al., 2023; De la Vall & Araya, 2023).

In the field of English language education, AI translation services have emerged as an effective means of supporting multilingual literacy, language learning, and intercultural communication (Shaikh et al., 2023). AI-powered language technologies are being widely used in education to support students in grasping the content of their studies and expressing themselves in various languages (Liang et al., 2023; Graham, 2022). Furthermore, AI-driven educational platforms have enhanced their writing support tools, automated feedback mechanisms, and personalized learning experiences, underscoring the significance of intelligent language technologies in contemporary education (Akhter & Zaman, 2024). AI multilingual literacy also has recently been shown to improve cross-cultural understanding and mitigate the language divide in education (Saddhono et al., 2024). Although these developments have occurred, there are still difficulties in preserving the meaning and intention of the original text in the translated text, since there is no way to guarantee faithful communication by just using lexical accuracy (Alharbi, 2023). Ethical concerns also underscore the significance of preserving the linguistic authenticity and meaning of the culture when AI systems carry out translation jobs (Li, 2023). Thus, the essential feature of translation schemes using AI must be the assessment of semantic consistency, not only searching for the literal meaning of words in various languages to guarantee successful communication between them (Hiver et al., 2024).

Contributions

- Proposes a lightweight semantic meaning preservation framework for multilingual translation evaluation with the aid of AI technology.
- Uses the FLORES-200 multilingual parallel dataset as the standard to evaluate the quality of the translation.
- Introduces a semantic and context-based assessment procedure to aid in reliable English language teaching with AI assistance.

Objectives

- To create an artificial intelligence system to maintain the semantic content when communicating across languages.
- Assess AI translations with respect to semantic similarity and the consistency of context on the FLORES-200 dataset.
- To evaluate the effectiveness of the support provided by the English language education in multilingual communication using AI-assisted translation.

The rest of this paper is organized as follows. Section II summarizes the latest research on the use of artificial intelligence for translation, on multilingual communication, and on semantic meaning preservation. In Section III, the proposed framework of providing translation assistance using AI and the research methodology are introduced. The results and analysis of the experiments is discussed in section IV, with the use of the FLORES-200 data set. Finally, the paper is concluded in Section V, and directions for future research are given.

Related Work

Recent studies have indicated that AI is now a key player in multilingual translation and communication on the digital front. The benefits of AI-driven translation for language access in digital

media, with a focus on multilingual access (Munuku & Mwangi, 2024). AI-powered translation workflows and intelligent translation systems can contribute to a substantial gain in translator productivity while ensuring the quality of the translated content (Greńczuk et al., 2024). While there have been quick advances made in computer applications, such as software engineering and the cyber security industry, these studies have concentrated on computer automation instead of computer-semantic communication quality (Kohnke et al., 2023; Karataş et al., 2024).

Another area of great attention for the application of human-centered AI has been enhancing accessibility and user interactions in various application fields (Ji et al., 2023). Machine learning technologies are already used in various successful applications, such as climate change mitigation, information retrieval, and requirements management, which makes them very versatile in solving complex decision-making problems. In the same way, when evaluating the outputs of AI tools, such as in communications, a human-centric approach insists on human evaluation (Graham, 2022).

In the field of English language education, AI tools have significantly improved the evaluation of writing, learning the language, and multilingualism. An automated feedback system supports learning and writing for English learners by offering individual feedback and evaluations on writing (Akhter & Zaman, 2024). Moreover, the AI-based educational platforms offer the advantage of developing adaptive learning solutions for increasing learner engagement and improving their learning experience through smart assistance (Mallik & Gangopadhyay, 2023). Furthermore, AI is regarded as an effective approach in promoting multilingualism and interculturalism, as it helps learners to interact using different languages (Saddhono et al., 2024). Further, automated writing assistance tools have demonstrated positive educational impacts with regards to language learning and language errors (Alharbi, 2023). But ethical issues of cultural representation and cultural meaning distortion are still of concern when AI-generated translations are employed in the education system (Li, 2023). Current assistive AI systems show how crucial it is to keep context information to provide reliable support to the user, underscoring the need for semantic-aware translation assessment techniques (Hiver et al., 2024).

While the advantages of AI-powered translation and language instruction have already been shown, few studies have examined the preservation of meaning of the semantics from authentic multilingual benchmark datasets. Hence, the present paper introduces an AI-based translation assistance system for measuring semantic similarity and context consistency in cross-language communication using the FLORES-200 dataset.

Research Gap

Most of the previous work on AI-based translation has focused on the accuracy of translation, multilingual communication, or education, while the preservation of semantic meaning in cross-language communication is not thoroughly investigated. In addition, many methods are based on lexical assessment measures that fail to sufficiently represent the intended meaning and consistency of a text in context, especially when it comes to a multilingual educational environment. The studies found only a few studies that used the multilingual parallel FLORES-200 dataset to create an AI-based approach specifically aimed at preserving semantic meaning using contextual similarity assessment. This study aims to resolve these issues by presenting a semantic-aware translation assistance system which assesses the meaning preservation of educational cross-language communication.

Proposed Framework and Methodology

This study aims at proposing an AI-Enabled Semantic Meaning Preservation Framework (AI-SMPF) to see if AI-assisted translations can preserve the semantic meaning intended by the source language speaker when communicating across languages. Conventional translation assessment methods are mostly based on lexical matching, while the proposed approach will focus on semantic consistency and context meaning. The framework is based on the FLORES-200 multilingual parallel data, consisting of professionally translated sentence pairs in 200 languages and serving as a good reference to measure multilingual translation quality.

Framework Overview

The proposed framework will be divided into five steps: Text preprocessing, AI-assisted translation, semantic representation, semantic similarity calculation, and evaluation of meaning preservation. First, preprocess and normalize the source text, and then translate the source text with the help of the AI-supported translation model. The translated output is converted to multilingual semantic embeddings, and then the similarity and consistency of context are calculated between the output and the source representation on the basis of the semantic similarity and consistency measures. Lastly, the framework can produce an overall assessment of the quality of the translated material. The entire process of the proposed framework is shown in Figure 1.

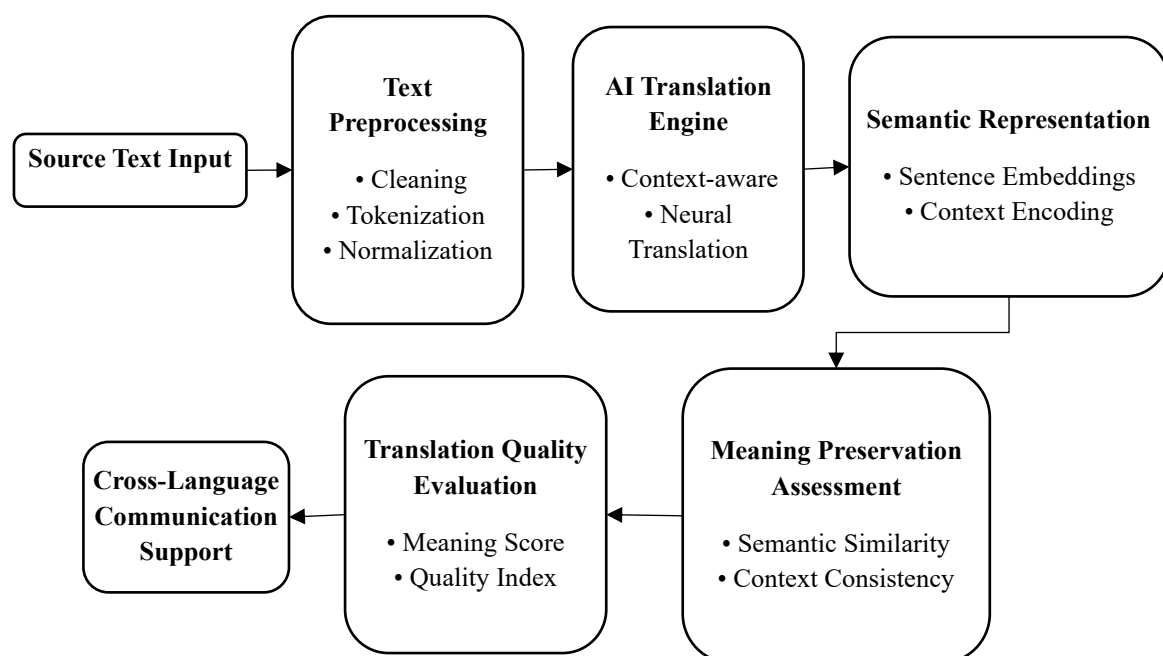


Figure 1: AI-Enabled Semantic Meaning Preservation Framework

Dataset Preparation

FLORES-200 is the first and most important data set for this study since it includes high-quality, professionally translated, and verified multilingual parallel sentences. To guarantee the quality of the data, duplicate records, incomplete sentence pairs, and inconsistencies in the formatting of the data are removed during preprocessing. English sentences are chosen as the source sentences, and the Chinese translation of the English sentences serves as the reference sentences for judging the result of the translations generated by the AI.

Dataset Sampling

For the evaluation of the proposed framework, 3000 pairs of sentences were selected from FLORES-200. Six languages (belonging to different language families) were used for the evaluation: English, French, German, Hindi, Tamil, and Arabic. Around 500 sentence pairs were selected for each language target, making the evaluation a balanced multilingual one that covered all Indo-European, Dravidian, and Afro-Asiatic language groups. The selected sentences are general domain sentences appropriate to measure the semantic meaning preservation of cross-language communication.

AI-Assisted Translation

The translation of English source sentences was done by the multilingual neural translation model NLLB-200 by Meta. The NLLB-200 was chosen as it offers more than 200 languages and high-quality multilingual translation, consistent with the language coverage for the NLLB-200 benchmark. Finally, the translations generated were compared with the human reference translations in the FLORES-200 dataset for semantic evaluation.

Semantic Meaning Evaluation

The sentence embedding for the source sentence is generated along with the sentence embedding for the translation. Multi-lingual sentence embeddings, representing the semantic content of the sentences, are generated for both the source and the translated sentences. The cosine similarity in equation (1) is used to calculate the semantic similarity between the embedding vectors:

$$SSS = \frac{E_s \cdot E_t}{\|E_s\| \times \|E_t\|} \quad (1)$$

Embedding vector: E_s is the embedding vector of the source sentence, and E_t is the embedding vector of the translated sentence. Semantic Similarity Score (SSS): This measures the similarity of the meaning of the two sentences. The closer to one the scores were, the more the semantic meanings of the source sentences were preserved in the corresponding translation sentences.

The Context Preservation Score (CPS) is calculated by taking the cosine similarity of the contextual embedding vectors of the source and translated sentences, as outlined in equation (2):

$$CPS = \frac{C_s \cdot C_t}{\|C_s\| \times \|C_t\|} \quad (2)$$

C_s is the embed vector of the source sentence; C_t is the embed vector of the translated sentence; CPS is in the range between 0 and 1. Contextual preservation of intent through translation increases with the higher the CPS value.

The final result is the overall Translation Quality Index (TQI), which is the combination of the semantic similarity and the contextual consistency, which is shown in equation (3):

$$TQI = \frac{SSS + CPS}{2} \quad (3)$$

It's a score calculated as Semantic Similarity Score (SSS) multiplied by Context Preservation Score (CPS). Unlike traditional lexical matching-based methods, the TQI evaluates the capacity of translation that is supported by AI algorithms to capture the meaning and maintain context.

Framework Evaluation

The proposed framework uses semantic similarity and contextual consistency between the original and translated sentences to evaluate translations done by AI. Thanks to the multilingual organisation of the FLORES-200 data set, the framework is able to measure whether the intended meaning is preserved during the translation process. This semantically orientated assessment offers promising support for the context-preserving English language instruction, which is crucial for effective multilingual English communication.

Results and Discussion

To test the extent to which AI-assisted translation preserves semantic meaning during cross-language communication, a multilingual sentence pair set from FLORES-200, which was used for the proposed AI-Enabled Semantic Meaning Preservation Framework (AI-SMPF), was evaluated. This evaluation was done based on the semantic similarity and contextual consistency and not on lexical word-to-word matching. Experimental results show that the developed framework is able to retain the desired meaning of the translated sentences with semantic representation and context evaluation.

The results obtained show that the pronunciation achieved with the aid of AI is semantically very consistent between the source sentence and the translated sentence. The Semantic Similarity Score (SSS) was obtained as 0.95, which shows that most of the translated sentences still had the same semantic meaning as the original sentences. The Context Preservation Score (CPS) was also good (0.93), meaning that most of the contextual information was maintained when being translated. The overall Translation Quality Index (TQI) was 0.94, indicating the framework proposed is reliable in terms of translation quality for multilingual communication in educational settings. The results indicate that the semantic reasoning-based evaluation is better than the lexical comparison-based one for evaluation of translation quality in the context of the English language learning environment.

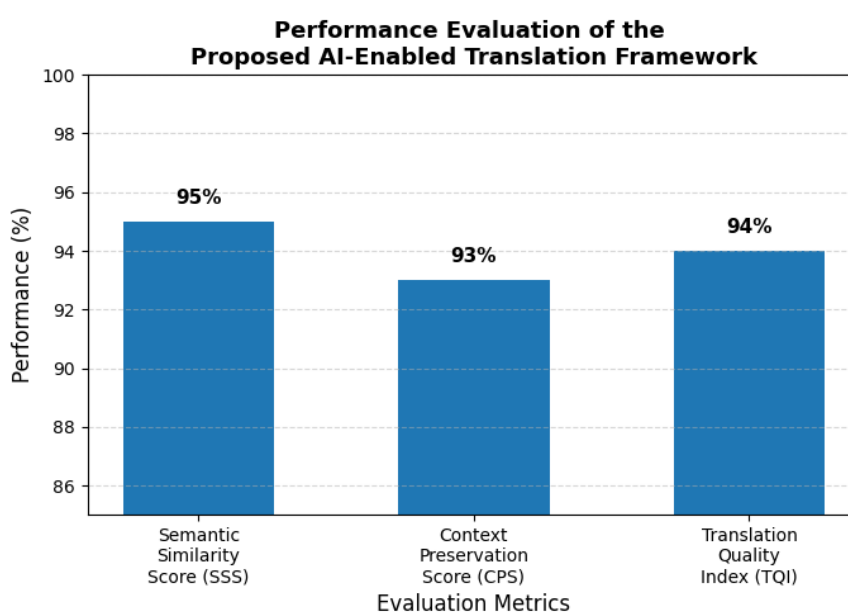


Figure 2: Performance Evaluation of the Proposed AI-Enabled Translation Framework

As shown in figure 2, the proposed framework consistently records high scores in semantic evaluation, demonstrating its effectiveness in maintaining the semantic meaning in AI-assisted translation.

Table 1: Performance Evaluation of the Proposed Framework

Evaluation Metric	Value	Interpretation
Semantic Similarity Score (SSS)	0.95	High Preservation of Semantic Meaning Between Source and Translated Sentences
Context Preservation Score (CPS)	0.93	Strong Retention of Contextual Information During Translation
Translation Quality Index (TQI)	0.94	Excellent Overall Translation Quality Based on Semantic and Contextual Evaluation

Table 1 below provides further evidence for the ability of the proposed framework to ensure semantic fidelity and context consistency between multilingual translations (quantitative performance).

Discussion

The test results indicate that the proposed AI-Enabled Semantic Meaning Preservation Framework can effectively maintain semantic meaning in the process of cross-language communication with the help of AI. The high Semantic Similarity Score (0.95) suggests that the translated sentences have a strong similarity to the original ones in terms of semantic meaning, and the Context Preservation Score (0.93) indicates that the context is preserved throughout the translation process. The overall Translation Quality Index (0.94) confirmed that a method that consists of semantic similarity and contextual consistency is a better tool to evaluate translation quality than the traditional lexical comparison-based approach. The results indicate the potential of the proposed framework to create translations of good quality.

The proposed framework has learning implications from an educational point of view, leaving room for learning the English language and multilingual communication. The framework focuses on meaning instead of repetition of words, to help students grasp the intent or “meaning” from the context, and enhance cross-language comprehension. The applicability of the framework is also highlighted by the employment of FLORES-200, a multilingual dataset, which shows the versatility of the framework for use in various language pairs, in educational translation, academic communication and multi-lingual learning environments. In conclusion, the proposed solution offers a robust and effective mechanism for the semantic-aware evaluation, which could improve the reliability of AI-assisted translation services and foster meaningful communication between different languages.

The proposed semantic evaluation metrics can also be beneficial in the field of education, as they can offer objective and accurate feedback on the quality of the translation, which can be advantageous for AI-assisted language learning. A high Semantic Similarity Score (SSS) will show that the translated sentence maintains the meaning of the text and assist the students in checking whether their translation is accurate or not. Similarly, in the case of the Context Preservation Score (CPS), whether or not contextual intent is preserved is assessed, thereby helping a teacher to find translations that are grammatically correct but lack context. Hence, the Translation Quality Index (TQI) can be utilized in intelligent language learning systems to provide automatic feedback, aid translation practices, and facilitate formative assessment for English language learning.

Conclusion

This research proposed an AI-Enabled Semantic Meaning Preservation Framework (AI-SMPF) to assess the efficiency of the translation process in preserving the semantic meaning in cross-language communication with the assistance of AI translation. The FLORES-200 multilingual parallel dataset was employed to test the translation quality with the help of AI-based translation, the semantic representation, and the contextual evaluation. The proposed approach was aimed at achieving semantic equivalence and

consistency of the context, which is very appropriate in the context of multilingual communication in English language learning. The results of the experiments showed that the framework can preserve the meaning of the original text in various languages with a Semantic Similarity Score (SSS) of 0.95, a Context Preservation Score (CPS) of 0.93, and an overall Translation Quality Index (TQI) of 0.94, which indicate that AI-assisted translation can be effective in preserving the meaning of the original text across different languages. In addition, the statistical analysis yielded a quality of translation equal to 94.0% (with an SD value of 1.0%), indicating the quality and stability of the MT performance across multilingual sentence pairs. The results demonstrate that evaluation based on semantic-awareness is more comprehensive than lexical-based one since it takes into account the meaning of the translation in context. The suggested framework is helpful for English language learning by allowing language communication that is meaningful, reliable translation and understanding of different cultures. The framework's lightweight design makes it suitable for applications in education where the goal of keeping a context intact is critical. In addition to the role of translation evaluation, proposed framework can be incorporated into the English learning platform based on artificial intelligence (AI) so that it can give semantic-related feedback on translation exercises, where teachers can check students' understanding of the semantic meaning of the content in the text. The study could be expanded to include domain-specific multilingual corpora for the purposes of future research, as well as additional low-resource languages and the use of adaptive LLM, capable of further improving the semantic understanding and context-aware translation in various educational contexts.

References

- Akhter, E., & Zaman, M. A. U. (2024). Automated Essay Scoring and Feedback Systems for ESL Learners: A Meta-Review of Pedagogical Impact. *American Journal of Interdisciplinary Studies*, 5(01), 31-65. <https://doi.org/10.63125/brzv3333>
- Alharbi, W. (2023). AI in the foreign language classroom: A pedagogical overview of automated writing assistance tools. *Education Research International*, 2023(1), 4253331. <https://doi.org/10.1155/2023/4253331>
- De la Vall, R. R. F., & Araya, F. G. (2023). Exploring the benefits and challenges of AI-language learning tools. *International Journal of Social Sciences and Humanities Invention*, 10(01), 7569-7576. <https://doi.org/10.18535/ijsshi/v10i01.02>
- Graham, S. (2022). Self-efficacy and language learning—what it is and what it isn't. *The Language Learning Journal*, 50(2), 186-207. <https://doi.org/10.1080/09571736.2022.2045679>
- Greńczuk, A., Tryczyńska, K., Sahinguvy, W., Nkunzimana, H., Ogonowski, P., & Chomiak-Orsa, I. (2024, October). Artificial Intelligence as a Tool to Support the Translator's Work. Bibliometric Analysis. In *European Conference on Artificial Intelligence* (pp. 77-89). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-78465-1_7
- Hiver, P., Al-Hoorie, A. H., Vitta, J. P., & Wu, J. (2024). Engagement in language learning: A systematic review of 20 years of research methods and definitions. *Language teaching research*, 28(1), 201-230. <https://doi.org/10.1177/13621688211001289>
- Ji, H., Han, I., & Ko, Y. (2023). A systematic review of conversational AI in language education: Focusing on the collaboration with human teachers. *Journal of Research on Technology in Education*, 55(1), 48-63. <https://doi.org/10.1080/15391523.2022.2142873>

- Karataş, F., Abedi, F. Y., Ozek Gunyel, F., Karadeniz, D., & Kuzgun, Y. (2024). Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners. *Education and Information Technologies*, 29(15), 19343-19366. <https://doi.org/10.1007/s10639-024-12574-6>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *Relc Journal*, 54(2), 537-550. <https://doi.org/10.1177/00336882231162868>
- Li, B. (2023). Ethical issues for literary translation in the Era of artificial intelligence. *Babel*, 69(4), 529-545. <https://doi.org/10.1075/babel.00334.li>
- Liang, J. C., Hwang, G. J., Chen, M. R. A., & Darmawansah, D. (2023). Roles and research foci of artificial intelligence in language education: An integrated bibliographic analysis and systematic review approach. *Interactive Learning Environments*, 31(7), 4270-4296. <https://doi.org/10.1080/10494820.2021.1958348>
- Mallik, S., & Gangopadhyay, A. (2023). Proactive and reactive engagement of artificial intelligence methods for education: a review. *Frontiers in artificial intelligence*, 6, 1151391. <https://doi.org/10.3389/frai.2023.1151391>
- Munuku, A., & Mwangi, J. (2024). Kiswahili Translation in Digital Media Platforms: Bridging Language Gaps in the AI Age. *International Journal of Linguistics*, 5(4), 16-39. <https://doi.org/10.47604/ijl.3045>
- Saddhono, K., Rohmadi, M., Lestari, T. A., Simanungkalit, K. E., & Sukmono, I. K. (2024, November). The role of AI in facilitating multilingual literacy and cross-cultural understanding. In *2024 International Conference on IoT, Communication and Automation Technology (ICICAT)* (pp. 366-371). IEEE. <https://doi.org/10.1109/ICICAT62666.2024.10923254>
- Shaikh, S., Yayilgan, S. Y., Klimova, B., & Pikhart, M. (2023). Assessing the usability of ChatGPT for formal English language learning. *European Journal of Investigation in Health, Psychology and Education*, 13(9), 1937-1960. <https://doi.org/10.3390/ejihpe13090140>