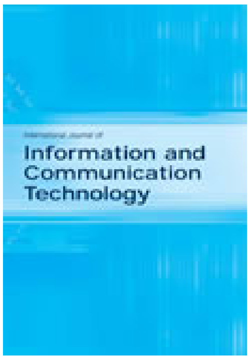
INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research



International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Development of an instructional model for Korean translation in multilingual classroom contexts

Rong Rong, Xiaojian Liu

DOI: 10.1504/IJICT.2025.10074672

Article History:

Received:	05 September 2025
Last revised:	24 September 2025
Accepted:	26 September 2025
Published online:	17 December 2025

Development of an instructional model for Korean translation in multilingual classroom contexts

Rong Rong*

School of Foreign Languages,
Liaodong University,
Dandong, Liaoning, 118001, China
Email: 15941537821@163.com
*Corresponding author

Xiaojian Liu

School of Information Engineering,
Liaodong University,
Dandong, Liaoning, 118001, China
Email: 13942586776@163.com

Abstract: Multilingual classroom contexts pose significant pedagogical challenges, particularly when students have diverse native languages and varying levels of Korean proficiency. This study introduces a data-driven instructional model for Korean translation education that employs machine learning to address learner diversity. The model evaluates translation outputs, identifies learner-specific error patterns, and personalises instruction based on three key variables: native language influence, historical translation accuracy, and individual learning progression. A dataset of Korean translation tasks was collected from university students representing six L1 backgrounds – Chinese, Vietnamese, Arabic, Russian, Japanese, and Spanish. Texts were pre-processed through tokenisation, lemmatisation, and POS tagging, with Word2Vec embeddings used for feature extraction. The proposed Sparrow Search Optimiser Tuned Attention-based Sequence-to-Sequence (SSO-Attn-Seq2Seq) model demonstrated substantial improvements, achieving 88–91% across accuracy, precision, recall, and F1-score. Results highlight its adaptability in handling idiomatic expressions and syntactic variation, providing a scalable solution for multilingual Korean language education.

Keywords: multilingual classroom settings; grammatical variations; languages; SSO-Attn-Seq2Seq.

Reference to this paper should be made as follows: Rong, R. and Liu, X. (2025) 'Development of an instructional model for Korean translation in multilingual classroom contexts', *Int. J. Information and Communication Technology*, Vol. 26, No. 46, pp.114–140.

Biographical notes: Rong Rong is a Lecturer in the School of Foreign Languages at Liaodong University, Dandong, Liaoning, China. Her research interests include multilingual education, translation pedagogy, Korean linguistics, and technology-enhanced language learning. She has published several papers on cross-cultural communication and the application of computational models in translation studies.

Xiaojian Liu is an Associate Professor in the School of Information Engineering at Liaodong University, Dandong, Liaoning, China. His research focuses on artificial intelligence, machine learning applications in language processing, and educational technology. He has contributed to various projects involving intelligent learning systems and natural language processing for multilingual environments.

1 Introduction

Multilingual classrooms were becoming more prevalent in a worldwide educational environment, offering special possibilities and contests for language teaching and translation techniques (Choi, 2024). Teachers must create instructional models that should meet the varied demands of students who speak different native languages as linguistic diversity in society (Han et al., 2021). Korea played an important part in the global economy, culture, and technology, Korean language has become significant internationally. Rising numbers of international students want to learn Korean together with more multicultural families inside Korea have created a strong interest in language courses adapted to learners with different native languages (Yilmaz et al., 2024). To design and apply an instructional model for Korean translation in multilingual classrooms, emphasising both understanding the language and learning the language to work with diverse groups (Gong, 2022). Language educators have debated the role of translation in language teaching for integrating translation for multilingual education. Translation in education was necessary for pedagogy in a long time. The use of Korean in authentic communicative situations helps students to retain the language and become more functional in fluency of speaking (Lee and García, 2021).

Translation tasks were useful in Korean language instruction that help learners to connect their native language's structures with the help of Korean syntax, vocabulary, and culture-specific nuances. Translation helps learners in understanding subtle connotations, idiomatic expressions, and sociolinguistic components term as possible to translate with grammar-based techniques (Antony et al., 2024). A translation-based instructional model was a common platform used for understanding language and collaborative learning in the case of teaching a multilingual classroom. The instructional model provides language development and encourages an inclusive and language-responsive learning environment. The application of this model emphasises collaborative translation tasks and digital translation tools for flexible accommodation of language levels and active participation (Smith, 2024). The Korean translation instructional model transfers beyond simple word-for-word conversion, but combines holistic interpretation of semantic and 16 realistic meaning awareness (Ryu et al., 2024). Developing an instructional model for Korean translation in multilingual classrooms involves creating a multi-layered framework that accounts for cognitive, linguistic, and sociocultural influences.

The cognitive considerations covers how learners take input and output language, both in the native language and in Korean, to decode and encode the meaning of the linguistic systems (Cho et al., 2024). The model linguistically constitutes the differences of Korean phonology, morphologies, syntaxes, and semantics, for identifying the grammatical interference of the words. Korean translation in multilingual classrooms constitutes cultural functions beyond linguistic ones. Language was essentially inseparable from culture and translation that can offer a glimpse into the norms, traditions, and cosmologies of Korean language communities. Sociocultural education helps to search the cultural values and language expressions that were difficult to translate the Korean language directly. The multilingual classrooms were a crucial learning process that requires collaborative translation activities where learners negotiate meaning and strategises reflect on linguistic choices. Language learning was traditionally involved with a great deal of translation training based on the ways of teaching. This led to the decline of translation which was often considered a block immersion and fluency, however communicative language teaching approaches needed translation.

Instructors must be trained to create translanguaging pedagogies and scaffolds for instructing the students through multimodal resources and formative assessment strategies. It helps to emphasise pedagogy in which students participate actively by choosing the texts to translate the cultural connotations and jointly evaluating the translation products. The implementation of an instructional model used for Korean translation in multilingual contexts was further supported by digital platforms (Almusharraf and Bailey, 2023). The structural differences between Korean learners' native languages rely on contrastive analysis, providing students with more difficult translations. Intercultural understanding and respect of core ideas in global citizenship education were promoted. Cultural content available in a student's native language, such as Korean, positively sustains their identity, enhances emotional well-being, and strengthens their connection to their heritage. In addition, Korean-to-other-language translation exercises require learners to think critically about language use, register, and context and hence develop a deeper understanding of both linguistic form and function (Pierson et al., 2021). The instructional model combines linguistic instruction, technological resources, cultural awareness, and learner strategies to address the complexities of multilingual classrooms to ensure appropriate Korean translation skills (Wilang and Duy, 2021).

The translation of the Korean language was heavily impacted by cultural background, particularly in real-time classroom interactions and the model was difficult to appropriately represent the error distinctions. The objective of this research is to create a revolutionary Sparrow Search Optimiser Tuned Attention-based Sequence-to-Sequence (SSO-Attn-Seq2Seq) approach to enhance translation accuracy and adaptability in diverse linguistic contexts. The suggested approach explores multilingual classrooms by analysing learner-specific error patterns and accommodating diverse native language influences. The key contributions of this research are as follows:

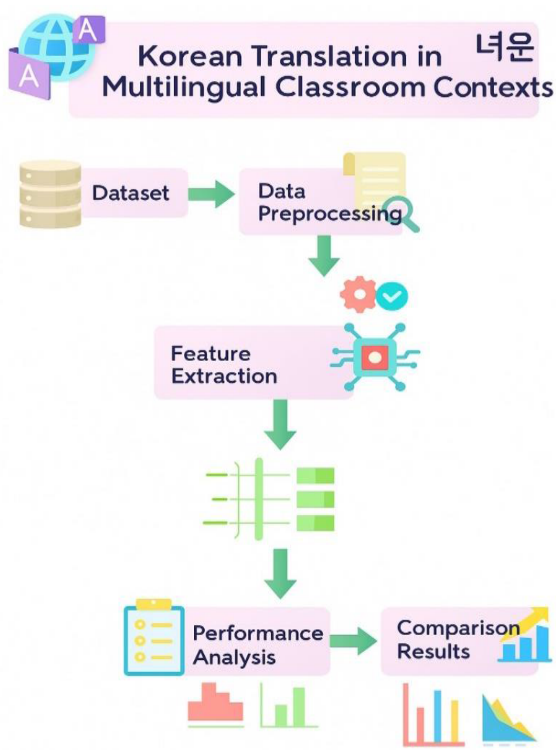
- *Dataset collection:* a Korean language learner errors dataset was collected from Kaggle. This dataset includes Korean sentences paired with six different native languages such as Chinese, Vietnamese, Arabic, Russian, Japanese, and Spanish. This comprehensive dataset used for investigating the linguistic transfer and translation mistakes impacted by native language backgrounds in a multilingual learning setting.

- *Data pre-processing and feature extraction:* tokenisation, lemmatisation, and POS tagging were applied to text data. For every learner output, a Word2Vec model was used to change a vector. The sentence-level features involved averaging word embeddings and error flags. This detailed preprocessing stage and extractions help to create an accurate model to analyse learner translation skills.

- *Optimised classification model:* the SSO-Attn-Seq2Seq model used to enhance the syntactic and semantic complexities inherent in multilingual translation tasks. SSO modifies hyperparameters on its own with the help of efficient training. The attention mechanism helps the model to identify the important parts of being translated. The optimised design works in real-time for education to identify translation behaviours in each user's native language group.
- *Real-time results:* the simulation results evaluate the precision, accuracy, recall, and F1-score for optimising the accurate translations. It performs better in identifying errors and syntactic particles and proving its capability to adaptively support Korean language instruction for diverse learners in real-time learning settings.

The remaining research consists of a literature review of Korean translation instruction, multilingual NLP challenges, and attention-based neural translation models, methodology encompasses (data preprocessing, feature extraction, and SSO-Attn-Seq2Seq model development), and experimental results highlighting superior metrics. The discussion highlights how the model effectively address the native language interference and adapts to learner-specific error patterns, improving personalised instruction. The conclusion emphasises improved efficiency, computational efficiency, and applicability of SSO-Attn-Seq2Seq, establishing a potent solution for real-time translation instruction in diverse classroom settings.

Figure 1 Graphical abstract (see online version for colours)



2 Related works

The relevant literature explores Korean language processing through advanced neural and multilingual models, emphasising dialect-aware translation, emotionally sensitive pedagogy, and cross-lingual adaptation techniques to enhance translation accuracy, instructional authenticity, and inclusivity in Korean language education and technology systems. Kim (2024) examined how the Korean English Language Learners engage with culturally relevant biases utilising the extensive knowledge of reading techniques during the learning process. It demonstrated how the strategies work to create meaningful reading experiences and provide additional information about the teaching literacy and language learning. According to the outcome, the Korean languages able to adapt and integrate the material by using their knowledge bases effectively throughout retelling exercises. Harris et al. (2025) examined how the language instructors teaching circumstances and reactions have been affected by the globalisation of Korean media and traditional assets. The globalisation of Korean learning was driven by pop culture, and its extreme popularity has boosted interest in learning the language. Language instructors need to adapt the evolving demands of their students in the wake of globalisation. The result findings showed the conflicts between culture and non-heritage learners in respect to their influence on the language classroom, for cultural and linguistic authenticity.

Table 1 Summary of related works on Korean translation in multilingual classroom contexts

References	*Technology used*	*Optimisation*	*Result*	*Challenges*
Lim et al. (2022)	Machine translation model, copy mechanism	Applied copy mechanism to improve the efficiency and performance	Improved the scores for both one-to-one and many-to-one models and has	Dialect data scarcity; modelling complexity for multiple dialects
Song (2023)	Korean as second language (KSL)	Analysed emotional vulnerability and interpretation of Korean instruction	Significant performance Identified divergent approaches of Korean teaching linked to English proficiency; anxiety due to perceived English inadequacy	The struggle of KSL instructors used to maintain authority and legitimacy due to English hegemony
Im (2023)	Trans language in English education via student reports	Encouraged active participation and reflection on trans-language practices	Students understood and applied trans language concepts	Encountering ineffective translanguaging moments in practice

Table 1 Summary of related works on Korean translation in multilingual classroom contexts (continued)

References	*Technology used*	*Optimisation*	*Result*	*Challenges*
Fedorova and Nam (2023)	Analysis of Korean Linguistic Landscape Data	Examined multilingual practices, norms, and ideological constructs	Identified pragmatic inequality of languages other than Korean; and the emergence of 'multilingual islands' challenging the monolingual ideology	Persists monolingual ideology; and uneven public language use
Zhang (2023)	Business Korean translation	Focus on word selection	Provided translation characteristics and references for commercials learning	Existing translation theories mostly literary-focused; lack of guidance for translation
Lee et al. (2022)	Content-Language Integrated Learning (CLIL) in Korean for international students	Interviews on perception and satisfaction of Korean-Medium Instruction (KMI) classes	Students improved content knowledge and Korean language skills; overall satisfaction with KMI classes.	Difficulty in rapid improvement of Korean academic proficiency.
Son et al. (2022)	Cross-lingual post-training (XPT) method for Korean language	Selectively and reused English pre-trained model parameters; added adaptation layer	Better performance with small target language datasets compared to target pre-trained models	Low resource status of Korean language; challenges in model adaptation
Park and Padó (2024)	Multidimensional quality metrics (MQM) for evaluation (English-Korean)	Created 1,200-sentence MQM benchmark; multi-task prediction of MQM scores using language models	Provided fine-grained, interpretable quality evaluation;	Typical evaluation reduces quality to a single score; need for multi-dimensional assessment
Kim et al. (2023)	Neural machine translation (NMT) for Korean language	Constructed parallel corpus using comparable corpora; phoneme decomposition	Phoneme decomposition improved Korean translation accuracy	Lack of sufficient Korean training data for NMT models
Bang et al. (2023)	English-Korean speech translation quantity and alignment	Used subtitle timing and bilingual sentence embeddings to enhance sentence alignment	Achieved 0.96 F-measure in sentence alignment; supported end-to-end speech translation	Lack of public English-Korean speech translation corpora; data scarcity

The education of English in Korean classroom environments has contributed to the nation's cultural and linguistic development. This growth has been supported by studies investigating the language and literacy practices related to Korea's multilingual official language (Jang, 2022). According to the outcome, the students used their intervention to choose linguistic and non-linguistic materials that could independently broaden the language repertoires. Lee and Gyogi (2022) examined the key themes of multiple literacies education focused on translation. It emphasises the text's multimodal aspects that help students to comprehend the effects of their choices when translating texts in various classroom settings. It helps students to make educated choices of translations to further utilise and make them aware of the context around the media. The result findings demonstrated that students' reflections on several media-related aspects like readers were used to create multiple literacies. Table 1 presents a summary of related works on Korean translation in multilingual classroom contexts.

2.1 Problem statement

Current advancements in the language processing have enhanced the accuracy of translation education models. Nevertheless, several vital limitations exist in the current literature often struggle to effectively address learner diversity and adapt the unique error patterns influenced by different native languages, limiting their effectiveness in multilingual classroom settings.

- The globalisation of Korean learning was driven by pop culture, and its extreme popularity has boosted interest in learning the language (Harris et al., 2025). It might be restricted to learner groups, thereby omitting different learner backgrounds and regional differences. It lacks in depth analysis of the pedagogical strategies and institutional support systems. The cultural and linguistic reality affects the learners' long-term language proficiency.
- Korean-English language learners representing a specific cultural and linguistic group might restrict the generalisability of the findings. The usage of repeating exercises as the main evaluation strategy might not completely account for the degree of learners' perception of critical texts (Kim, 2024). The cultural relevant biases were recognised but it does not perhaps the learners' identity formation and long-term academic performance.

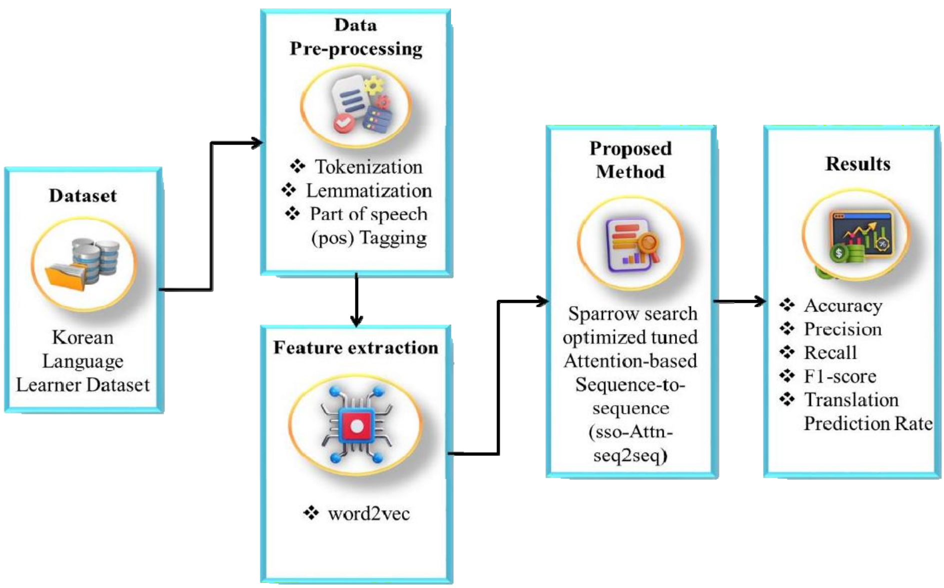
To overcome these issues, the suggested approach, SSO-Attn-Seq2Seq used to enhance Korean translation education by leveraging systems to personalise instruction based on learner-specific data. By analysing translation history, language influence, learning progression, and the model improves translation accuracy, reduces instructional inefficiencies, and enables scalable, adaptive learning in multilingual classrooms.

3 Proposed methodology

The literature review on recent studies in Korean translation education for multilingual classrooms was previously described. It highlighted their shortcomings in adapting to learner diversity and ensuring translation accuracy across varied native language backgrounds, and the proposed model explained the research gap to address these issues.

A Korean language learner dataset was collected from Kaggle. The pre-processing procedure includes tokenisation, lemmatisation, and part-of-speech tagging used to normalise the data. Word embeddings such as Word2Vec were employed to transform words into vector representations. The SSO-Attn-Seq2Seq method used to enhance translation accuracy and adaptability in diverse linguistic contexts. The SSO enhances model parameters, while the Seq2Seq framework strengthened by attention mechanism that guarantees accurate and context-aware translations. The syntactic and semantic difficulties presented in multilingual translation assignments were especially served by the hybrid SSO-Attn-Seq2Seq approach. Figure 2 presents the overall flow of the Korean translation in multilingual classroom contexts.

Figure 2 Overall flow of the Korean translation in multilingual classroom contexts (see online version for colours)



3.1 Dataset

A Korean language learner error by native tongue dataset was collected from Kaggle. This dataset includes Korean sentences paired with learner-generated variants influenced by six different native languages such as Chinese, Vietnamese, Arabic, Russian, Japanese, and Spanish. These native language transfer mistakes were reflected in the learner outputs. It has a variety of sentence structures, including questions, refusals, and descriptive objectives. This dataset was designed for language acquisition, error analysis, and educational technology, particularly for improving Korean language teaching methods. Back-translations and many learner variations were included in the Korean sentence. Learner-written sentences, mistake tags, corrections, and metadata were included in this dataset. This dataset provides a comprehensive resource for analysing language interference and error patterns in Korean language learning. Table 2 depicts the demographic sample data for Korean learner error analysis.

Table 2 Demographic sample data for Korean learner error analysis

ID	*Korean original*	*LI_Language*	*Output learner*	*Back translated Korean*	*Ll_Translation*	*Sentence_type*	*Error*
1	나는	Chinese	나는	나는 학교 갑니다	[Chinese translation	Subject-verb	True
	학교에	Vietnamese	학교		[Vietnamese]	Subject	False
	갑니다	Arabic	갑니다	나는 학교에	갑니다	Verb-	False
2	나는 학교에	Russian	갑니다		학교에 갑니다	Present	False
3	갑니다	Japanese	나는 학교에	학교에 갑니다	[Arabic translation of]	Present	True
4	나는 학교	Russian	나는 집에서	나는 집에서 갑니다	나는학교에	Subject-Verb-	True
5	나는 학교에	Japanese	나는 학교 갑니다	나는 학교에 갑다	[Spanish translation of]	Present	True
6	갑니다	Spanish			나는학교에	Present	True

Note: Here, the dataset was described then next, it describes the data preprocessing procedures.

Source: https://www.kaggle.com/datasets/zoya77/korean-language-learner-errors-by-native-tongue

3.2 *Data preprocessing*

Data processing is used to preprocess the raw data. Unprocessed data was transformed into a readable format by data pre-processing, a necessary step in data exploration. It helps to ensure the missing values and inconsistencies of the lay foundation for accurate insights which makes the data fit for further computational tasks. Data preprocessing includes tokenisation, lemmatisation, and part-of-speech (POS) tagging.

3.2.1 *Tokenisation*

To build a strong instructional model for Korean translation in the multilingual classroom, it was essential to use tokenisation during data preprocessing. Tokenisation separates sentences into individual words or sub-words which helps translation of the Korean language's unique grammar and spacing rules. In multilingual classrooms, students learn different languages, and tokenised data provides simple translation. It helps to solve communication problems and improves how quickly people learn Korean in different classroom settings. Using the natural language processing (NLP) approach of tokenisation, a collection of texts was distinct into meaningful words, clauses, signs, statements, and other parts. Through the conversion of unstructured material into a structured format, tokenisation streamlines text analysis and facilitates the use of NLP techniques. Tokenisation allows for improved comprehension and processing by maintaining the semantic linkages and context of words in sentences. Tokenisation was necessary to convert unprocessed text into format. The tokenised output represents the fine-grained linguistic analysis of Korean learner sentences written by Koreans of different L1 languages. Both of these data reflect morphologically and tokenised learner responses to standard Korean sentences allowing error patterns beyond sentence structure to identify the basis of native language influence. Tokenisation was one of the fundamental blocks in the field of NLP when it comes to breaking sentences into morphemes to compare the syntax and grammar use. Table 3 represents the morpheme-level tokenisation of learner Korean sentences across the L1 background.

Table 3 Morpheme-level tokenisation of learner Korean sentences across L1 background

ID	*L1_Language*	*Learner_Output*	*Tokenised_Output (Okt)*
1	Chinese	나는학교갑니다	[‘나’, ‘는’, ‘학교’, ‘갑니다’]
2	Vietnamese	나는학교에갑니다	[‘나’, ‘는’, ‘학교’, ‘에’, ‘갑니다’]
3	Arabic	나는학교에갑니다	[‘나’, ‘는’, ‘학교’, ‘에’, ‘갑니다’]
4	Russian	나는집에서갑니다	[‘나’, ‘는’, ‘집’, ‘에서’, ‘갑니다’]
5	Japanese	나는학교갑니다	[‘나’, ‘는’, ‘학교’, ‘갑니다’]
6	Spanish	나는학교에 갑다	[‘나’, ‘는’, ‘학교’, ‘에’, ‘갑다’]

3.2.2 Lemmatisation

Lemmatisation a data preprocessing technique, takes a word and turns it into the simplest form (lemma), so that the models perform more accurately. Applying lemmatisation in such instructional models helps to guarantee that all students' different words from various backgrounds were interpreted similarly. Standardising such systems or tools avoids misunderstandings and makes it easier to link Korean with other languages in translation. In multilingual classrooms students express themselves in many ways, lemmatisation helps to correct and translate the terms more accurately. This way of teaching improves how texts were translated and makes it easier for students to understand and talk with various learning environments. By combining variations of the same word, stemming or lemmatisation lowers words to their simplest form, reducing sparseness. It increases the precision of sentiment analysis and pattern detection. The structure t evaluates the text's psychological overtones to help with a mental wellness evaluation.

3.2.3 Part-of-speech (POS) tagging

The POS tagging was a data preprocessing technique that gives every word in a sentence with correct grammatical labels (such as noun, verb, or adjective). POS tagging was especially helpful in multilingual settings for Korean translation because it helped to make translation tools more accurately. In the instructional model, POS tagging helps language learners to translate grammatically familiar. It provides teachers the opportunity to focus on how Korean was different from other languages. Learners benefit from the tagged data because the system creates customised exercises for grammar and vocabulary. It helps to make multilingual teaching more effectively and increases students' ability to translate. In NLP, POS tagging was important for processing data because it assigns noun, verb, adjective, or adverb labels to all words in a sentence. The POS tagging helps the structure of sentences to improve the accuracy of their language work. POS taggers assign a part of speech to each word, especially when the word can be used in more than one way such as either a noun or verb. By indicating that each word in a sentence, models increase their understanding and links between words. POS tagging processes constitute entity recognition, syntactic parsing, machine translation, and sentiment analysis can be done more easily. Encompassing taggers make use of rules, statistics, and machine translation.

Rules that guide rule-based taggers developed by statistical and machine translation taggers to train the annotated text samples. POS tagging ensures that grammatical errors handled by the same source and target languages. Moreover, it reveals common mistakes learners tend to commit. The POS tagging adds more detail to data which was essential for successful NLP solutions. The preprocessing stages of tokenisation, lemmatisation, and POS tagging were completed, and then it describes the feature extraction process.

3.3 Feature extraction

The feature extraction used in machine learning and data analysis to minimise the dimensionality of data. It improves the model performance, reduces the computational cost, and improves the model's accuracy, by features focusing on the most relevant. Preparing data for effective learning and decision-making on complex tasks feature extraction was needed. It cracks unprocessed data into a collection of useful characteristics that make analysis easier while preserving details. Here, the feature extraction includes Word2Vec.

3.3.1 Word2Vec

Word2Vec takes words and changes them into vector forms that show their meaning in relation to other words. It can interpret the fine points of language meaning across different languages. The vector embeddings were passed into the instructional model which constitutes useful information for delivering targeted translations. The instructional model accommodates learners' language backgrounds so that each student can receive suitable instruction. Because of this relationship, students from many native languages easily learn and understand the language. In practice, its features facilitate real-time learning accessibility for all students and provide customised text translation for classroom use. Word2Vec functions as a neural network-based model because it converts words to numerical vectors that reveal semantic meanings between words. The application of this technology in creative writing enables educators to evaluate student writing and enhance both patterns and word selection in generated content. An independent toolkit known as Word2vec was used to create word vectors. Word2vectors have the following properties such as the vector proximity between the Word vectors that might be used to quantify the connection between words.

The gap between two word vectors was smaller and it might be semantically relevant to the words. Word2vec was often used for NLP applications such as text categorisation, sentiment assessment, identical terms, and segmentation. Text summarising constitutes word meanings and word similarities, however the majority of these techniques categorise sentences based on textual characteristics, making them limited to specific texts and lacking generalisability. It discovers vectors representing the set of terms in the vocabulary *S* in equation (1),

$$S = \begin{bmatrix} cx_1 \\ x_2 \\ \vdots \\ x_i \end{bmatrix} \rightarrow \vec{U}_s = \begin{bmatrix} \vec{U}_1 \\ \vec{U}_2 \\ \vdots \\ \vec{U}_j \end{bmatrix} \tag{1}$$

where $U \rightarrow \rightarrow S$ represents the embedded vectors. The text summarising method could maintain the semantic connection between sentences to ascertain the degree based on Word2vec. The Word2Vec outcomes revealed the meaning structure of Korean sentences created by learners and analysing the normal vector representations of their translations. The translation problems influence the native language and changes in grammar influence word links and sentence interpretation. Transforming student production into vectorised form makes it possible to explore how language was learned in different language contexts. Table 4 represents the comparative word embedding results for Korean learner output sentences.

Table 4 Comparative word embedding results for Korean learner output sentences

ID	*Learner_Output*	*Avg Word2Vec (3D)*
1	나는학교갑니다	[0.274, 0.138, 0.612]
2	나는학교에갑니다	[0.259, 0.146, 0.597]
3	나는학교에갑니다	[0.259, 0.146, 0.597]
4	나는집에서갑니다	[0.265, 0.132, 0.563]
5	나는학교갑니다	[0.274, 0.138, 0.612]
6	나는학교에 갑다	[0.280, 0.140, 0.605]

The feature extraction process of Word2Vec used to transform words was completed, and then it describes the proposed model (SSO-Attn-Seq2Seq) to enhance the translation accuracy.

3.4 SSO-Attn-Seq2Seq

The integration of SSO and Attn-Seq2Seq presented a hybrid approach for enhancing Korean translation tasks in multilingual classrooms. SSO was a robust metaheuristic algorithm used for optimal hyperparameters of a neural model based on the social foraging behaviour of sparrows. Applying the Attn-Seq2Seq model encompasses the training phase, and the SSO model further refines the convergence rate along with a more powerful source to target language mappings learned. SSO efficiently avoids local minima through dynamic adjustment of parameters including learning rate, attention weight distribution, and hidden layer size; it outperforms competing methods in terms of generalisability over multilingual data. The Attn-Seq2Seq model was used for sharing the encoder-decoder structure with the integration of the attention mechanism that outperforms the model in incorporating the long dependencies across the sequences. By introducing the attention mechanism, the decoder can be conditioned on the context which more accurately focused on the relevant part of the input sequence during the translation, generating more fluent and contextually relevant text. In educational environments, precision and clarity of the translated text was essential because it directly influence the learner's understanding and take a part in multilingual contexts.

The hybrid model SSO-Attn-Seq2Seq achieves higher translation accuracy for the instructional model to adapt the variability of linguistic inputs frequently encountered across diverse classrooms. The application of this model within language learning platforms was possible and it can provide learners with real-time feedback, vocabulary help, and context-sensitive grammar correction. It supports differentiated instruction by teaching learners with diverse skills and native languages. The SSO-Attn-Seq2Seq hybrid model used for bidirectional translation and assists Korean language learners and Korean-speaking students learn other languages. Students gain a more serious and in-depth grasp of languages, and instructor have access to provide better materials for organising classes and fostering multilingual communication. The hybrid model SSO-Attn-Seq2Seq strengthens the linguistic capabilities of learning to construct an adaptive, accurate, and context-aware translation tool. In Attn-Seq2Seq model, the input sentence was processed on a tokens basis to constitute the encoder. Retain all inserted sentences into a continuous vector representation. The multilingual data was captured by semantic and syntactic information of the input tokens through these vectors. It describes the encoder's parameters. This transformation enables the model to represent the whole input sequence in a format and the decoder generates accurate translations at later times. The input sentence was summarised into a sequence of rich feature vectors for carrying necessary linguistic content to translation. The encoder productivity was expressed in equation (2).

$$g_j = Encode\left(w_j; \theta_f\right) \tag{2}$$

where w_j represents the location of the input token, g_j indicates as encoder and θ_f depicts the weight of the encoder network. The encoder attention score computation makes a decision about the input sequence to concentrate on generating a particular word output. It prescribes the amount of each input tokened to generate an output token and the decoder manipulates the multilingual contexts. The effect of this mechanism increases the translation accuracy, especially in the translation of multilingual sentences. The attention score computation was expressed in equation (3).

$$a_{ji} = \frac{\exp\left(score\left(t_{i-1}, g_j\right)\right)}{\sum_{k=1}^{S_w} \exp\left(score\left(t_{i-1}, g_j\right)\right)} \quad (3)$$

where a_{ji} indicates the weight of the encoder hidden state, t_{i-1} represents the decoder output token, S_w illustrates the length of the input sequence, and $score(\cdot)$ indicates the function of alignment between the decoder state and encoder output (4).

$$t_i = Decode\left(Z_{i-1}, t_{i-1}, \theta_c\right) \quad (4)$$

Here, t_i represents the hidden state of the decoder, and z_i depicts the output token generated by the decoder. The model attempts to escape local minima by using the update vector and simulating sparrow foraging vigilance behaviour. The step size was controlled and ensured an adequate compromise between convergence speed and instability. Finally, the resulting optimisation of the model improves the model's generalisation power over multiple frames of multilingual contexts in a classroom translation task. The update rule was expressed in equation (5).

$$\theta^{(n+1)} = \theta^{(n)} + \eta \times SSO\ update(\nabla \mathcal{L}) \quad (5)$$

Here, $\mathcal{L}$ indicates the function of loss, η indicates the learning rate and *SSO–update* represents the updated vector.

3.4.1 Attention-based sequence-to-sequence (Attn-Seq2Seq)

In multilingual classes, the Attn-Seq2Seq model makes it possible for students to speak educational content in Korean, even if students speak other native languages. The translation included in the lesson plan, so that educators can encourage students to gain new vocabulary, learn grammar, and build sentences using different interactive activities. The Attn-Seq2Seq model used in certain language patterns for the classroom, making its translations more accurate with use. Attn-Seq2Seq was an innovative neural network capable to translate text by studying how the input and output sequences were related to one another. Because of its ability to focus on specific words, the model creates sentences that were more accurate and natural. The attention-based sequence-to-sequence approach originates from machine translation, it converts a source of language's input sequence into a target language's output sequence. The Attn-Seq2Seq model converts an length outcome into a sequence of D from a S-length input observation of sequence $\mathcal{P}$. The Attn-Seq2Seq models use the chain rule of conditional probabilities to construct the posterior distribution $O(\mathcal{D}/\mathcal{P})$ for long short-term memory (LSTM) model and a linguistic model was expressed in equation (6).

$$\left(D/p\right) = O\left(d_1/p\right)O\left(d_2/d_1, p\right)...\left(d_2/d_1/d_1, ..., d_{j-1}, p\right) \quad (6)$$

The Attn-Seq2Seq models work well on machine translation since the input and output sequences typically have comparable lengths and irregular alignments. The alignment between the input and output sub-word unit sequences was comparatively local and rigorously monotonic when the input observation sequence was frequently longer. Additionally, it was more difficult to integrate rule-based information, such as lexicon systems frequently needed more training data. The encoder produces a series of hidden states *genc*, of length. This sequence serves as a high-level, summarised representation of the input data for tasks like classification or decoding expressed in equation (7).

$$g_1^{enc}, g_2^{enc}, ..., g_s^{enc} = Encoder(w_1, w_2, ..., w_s) \tag{7}$$

The attention module receives the input decoder from the hidden state as *gdec*, and all encoder 5 hidden states as *gdec*. The attention module compares each encoder's hidden state *gdec* with the current decoder's hidden state to determine alignment weights. These weights were obtained from a compatibility score, to establish the relative importance of each input element in the model's output generation was expressed in equation (8).

$$a_{j,s} = \frac{\exp\left(Score\left(g_j^{dec}, g_s^{enc}\right)\right)}{\sum_{k=1}^{S} \exp\left(score\left(g_j^{dec}, g_k^{enc}\right)\right)} \tag{8}$$

The Attn-Seq2Seq modules were distinct into various categories based on the various compatibility score functions expressed in equation (9).

$$Score\left(g_j^{dec}, g_s^{enc}\right) = v_a^T \tanh\left(W_a\left[g_j^{dec}, g_s^{enc}\right]\right)$$

The Attn-Seq2Seq model computes to produce a context vector *dj*. By efficiently condensing the pertinent encoder data required for the present prediction, this vector helps the decoder provide outputs that were more precise and contextually aware were expressed in equation (10).

$$d_j = \sum_{s-1}^{s} \alpha_j, sg_s^{enc} \tag{10}$$

The inputs for the decoder were the prior context vector d_{j-1} was previously labelled from prediction z_{j-1}. It was used to calculate the current hidden state *genc*, and which helps to produce the subsequent output in the sequence by inspiring from contextual data, and previous predictions were expressed in equation (11).

$$g_j^{dec} = f_{dec}\left(z_{j-1}, d_{j-1}\right) \tag{11}$$

In baseline design, z_{j-1} and d_{j-1} were the inputs of the Attn-Seq2Seq model. Although the rooted calculation includes the label data from the last step to complete the decoder hidden state information and the context vector that combines all encoder hidden state data expressed in equation (12).

$$\hat{g}_j^{dec} = f_{dec}\left(z_{j-1}, g_{j-1}^{dec}\right) \tag{12}$$

The Attn-Seq2Seq model initially computes the context vector ci using *gdec*, after which it derives the cognitive hidden state $\hat{g}^{dec}$ in equation (13) as follows:

$$\tilde{g}_j^{dec} = \tanh\left(W_c\left[d_j; g_j^{dec}\right]\right) \tag{13}$$

A probability distribution across potential labels for the current output produced by the projection layer, followed by the soft max function expressed in equation (14). is theoretically accurate if you want to show the probability distribution across output labels produced by soft maxing the projected feature vector $X_p g^{cfd}$. The phrase generates a normalised probability distribution across potential output labels, where g^{cfd} stands for the context-dependent feature descriptor or hidden state, and X_p indicates the projection layer output or embedding matrix.

$$O\left(Z_j / Z_{1:j-1,w}\right) = softmax\left(X_p g^{cfd}\right) \tag{14}$$

The prediction of the sequence at every state generates distribution, which shows the probability of each label.

3.4.2 Sparrow search optimiser

Sparrow search optimiser (SSO) algorithm was inspired by teamwork, food searching and protection strategies constituted by swarms of sparrows. It was used to adjust essential features in translation models or ways of instruction so that learning can become more accurate and adaptable. SSO was incorporated with teaching strategies to match translation activities, vocabulary learning tasks, and sentence structures for students according to their learning level and background. The model delivers content at difficult levels and adapts for multilingual learners. Analysis and refinement of translation tasks supported by feedback make the model better at helping learners improve their thinking skills and achieve better outcomes. It drives students to participate by incorporating techniques such as reviewing translations with associates and using the language with context. The SSO allows teachers to bridge different languages and help all students to learn Korean well in multilingual diverse classrooms. The feeding and anti-predation behaviours of sparrows serve as the primary inspiration of SSO, a revolutionary meta-heuristic technique. Sparrows were a frequent bird in the natural world that coexists with individuals. The sparrow often lives in groups with a distinct division of labour, and its acute mouth was small and powerful. The rest of the population depends on the food information provided by the former sparrows to seek food, while certain sparrows constitute to prey victims and supply foraging places and instructions for the entire population.

The population immediately begins back feeding once a sparrow in the population detects danger and promotes a sound alert. Based on the intelligence and memory of sparrows, the SSO proposal accurately mimics the cooperative mechanism of sparrow populations during daily feeding. The individuals were assigned to find the food known as discoverers. The joiners used to obtain the food by following the discoverers. To increase the effectiveness in obtaining food, a subgroup of joiners observes the discoverers constantly and plans their movements to compete for the food. It was possible to switch the roles of discoverer and joiner to ensure the total population. The sparrow's position was saved as an algorithmic solution. The sparrow's starting locations depicted by matrix were expressed in equation (15) as follows:

$$\begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,c} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,c} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m,1} & w_{m,2} & \cdots & w_{m,c} \end{bmatrix} \tag{15}$$

where c represents the problem's dimension and m indicates the sparrow's population. The sparrow's fitness value is expressed in equation (16).

$$E(w) = \begin{matrix} e([w_{1,1} \quad w_{1,2} \quad .... \quad w_1]) \\ e([w_{2,1} \quad w_{2,2} \quad ... \quad w_2]) \\ e([w_{m,1} \quad w_{m,2} \quad ... \quad w_{m,c}]) \end{matrix} \tag{16}$$

Here, the fitness function indicated as *e*. A wider diversity of foods can be found since the discoverer must navigate the foraging path for the entire community. The discoverer's position was updated in the following manner during the iterative process expressed in equation (17).

$$W_{j,i}^{S+1} = \begin{cases} W_{j,i}^{s} EXP\left(\dfrac{j}{a, Marxitem}\right), & Q_2 < TS, \\ W_{j,i}^{s} + R.K, & Q_2 < TS. \end{cases} \tag{17}$$

where W indicates the current location of the j^{th} sparrow, and Max item represents the maximum number of algorithmic iterations; S indicates the current iteration; a indicates the uniform number; and R represents the random number. The population's sparrows sound an alert if they come across a hunter or other threat. The discoverer will guide the population to safer foraging locations when the alert value exceeds for safety. The updated location of the follower was expressed in equation (18).

$$W_{j,i}^{S+1} = \begin{cases} R.EXP\left(\dfrac{W_{worst} - W_{j,i}^{S}}{j^2}\right), j > \dfrac{m}{2} \\ W_{j,i}^{s+1} + \left|W_{j,i}^{S} - W_{j,i}^{S+1}\right|.B^{+}.K, else \end{cases} \tag{18}$$

Here, W_{wrost} represents global position, and W_{s+1} indicates the global optimal value of discover. The sparrow flies to different locations to gather food and boost energy. When $j > m$, it 2 indicates the joiner maintained at the best location and selects a foraging spot close to W_a. Sparrows often constantly keep an eye on the colony as a whole and advise everyone to backfeed in times of danger. The modified position was expressed in equation (19).

$$W_{j,i}^{S+1} = \begin{cases} W_{j,i}^{S} + \beta\left|W_{j,i}^{S} - W_{j,i}^{S}\right|, e_j \neq e_h, \\ W_{j,i}^{S} + L\dfrac{W_{j,i}^{S} - W^{S}worst,i}{\left(e_j - \epsilon_{wrost}\right) + \varepsilon} e_j = e_g \end{cases} \tag{19}$$

where L represents the random number from [–1, 1]. A sparrow's current fitness value expressed as 0, e_g, and fe_h for obtaining the global fitness values. The SSO-Attn-Seq2Seq hybrid model intersects the periphery of linguistic, computational intelligence and pedagogical innovation learning. This model leverages the optimisation strength of the SSO and the linguistic precision of the Attn-Seq2Seq to ensure multi-lingual Korean language education. Moreover, it enhances the quality of translation in technical and matches the learning of equity, engagement, and autonomy of the learners. Algorithm 1 shows the proposed SSO-Attn-Seq2Seq model working procedure.

Algorithm 1: SSO-Attn-Seq2Seq

```
import numpy as np
```

```
import torch
import torch. nn as nn
import torch. optim as optim
from some_translation_dataset import load_multilingual_korean_data
from some_bleu_evaluation import compute_bleu_score
class Encoder(nn. Module):
def init (self, input_dim, emb_dim, hid_dim, n_layers, dropout):
super().init ()
def forward (self, src): return outputs, hidden
class Attention(nn. Module): defbiggr init (self, hid_dim): super ().biggr initbiggr() def forward (self, hidden,
encoder_outputs): return attention_weights, context
class Decoder(nn. Module): def init (self, output_dim, emb_dim, hid_dim, n_layers, dropout, attention):
super(). init ()
def forward (self, input, hidden, encoder_outputs):
return output, hidden, attention_weights
class Seq2Seq (nn. Module): def init (self, encoder, decoder): super(). init ()
self. encoder = encoder
self. decoder = decoder def forward(self, src, trg): return outputs
class Sparrow: def init (self, param_vector):
self. position = param_vector
self. fitness = None
def initialise_population(pop_size, param_bounds):
population = [for _ in range(pop_size)]
params = np. array([np. random. uniform (low, high) for (low, high) in param_bounds])
return population def fitness_function (params, train_data, val_data): pass def update_positions(population): pass def main (): train_data, val_data, test_data = load_multilingual_korean_data()param_bounds = [(1e – 4, 1e – 2), (128, 512), (0.1, 0.5)]
pop_size = 10
max_iter = 20
population = initialise_population(pop_size, param_bounds)
best_sparrow = None
best_fitness = -np. inf for iteration in range(max_iter): print (f"Iteration {iteration + 1}/{max_iter}")
for sparrow in population: if sparrow. fitness is None: sparrow. fitness = fitness_function(sparrow. position, train_data, val_data) print(f"Fitness: {sparrow. fitness}")
if sparrow. fitness > best_fitness: best_fitness = sparrow. fitness best_sparrow = sparrow update_positions(population)
print ("Best hyperparameters found: ", best_sparrow. position)
print("Best BLEU score: ", best_fitness)
if name == " main ":
main()
```

4 Results

The proposed model's strategy and its applications in enhancing Korean translation education through the integration of SSO-Attn-Seq2Seq were described previously. It effectively addresses the linguistic complexities in multilingual classrooms through a scalable and data-driven method with various language backgrounds. Experimental data indicate that the suggested model SSO-Attn-Seq2Seq improves Korean translation accuracy in multilingual classroom settings. The model's reliable performance was attributed to the integration of SSO and Attn-Seq2Seq, which effectively precise context handling and learner-specific error correction. These results demonstrate the effectiveness in enhancing adaptive instruction, translation quality, and pedagogical efficiency across linguistically diverse learner groups.

4.1 Experimental setup

The experiment setup was implemented using Python 3.10.1 to realise the proposed method. This Python version was selected for its compatibility and improved performance over previous versions. It provides an in-depth explanation of the experimental outcomes. It helps to enhance the accuracy and adaptability of Korean translation in multilingual educational contexts.

4.2 Performance evaluation of the proposed method

The error rate heatmap of L1 language vs. sentence type illustrated in Figure 3, represents the error rates for different sentence types across various languages. The colour gradient ranges from yellow, indicating lower error rates, to dark red, signifying higher error rates. The x-plane indicates the sentence type and y-plane indicates the L1 languages. The mistake rate of the matching L1 and phrase type was represented by numerical values in each cell of the heatmap. In heatmap, Subject-Verb-Future sentences show a pattern of high errors in every language, whereas Adjectives have lower errors. This visualising patterns and obstacles can strengthen the language teaching methods and practices. The heatmap conveniently illustrates how a language can influence sentence mistakes and their interests.

The word usage trends in the Korean language corpus represented in Figure 4, where x-plane indicates the frequency of words, and y-plane lists the Korean words. Frequency ranges were shown by using a graded colour scale that changes from dark purple to light green. By visualising the language data in a normal format supports the analysis of language and makes it easier to find patterns in texts. This graph helps to judge the characteristics of a text such as how frequently particular words were used. In addition, the analysis can be used in machine translation models to polish translation systems or advance tasks based on their content.

The error rate comparison across different L1 Languages was illustrated in Figure 5. The x-plane indicates the L1 language and the y-plane indicates the error rates. The Chinese, Japanese, and Russian has 0.9 error rate, Spanish has 0.75 error rate, Arabic has 0.56 error rate and Vietnamese has 0.5 error rate. This seems to show that errors speakers have significantly affect the native language, and ways of teaching languages during language learning. These findings provide information on the sound elements that affect learning another language regardless of the first language speakers. It helps students to speak another language, paying attention to the language-specific correctness. This visualisation improves the way of teaching, so that students in multilingual classes can better understand and communicate.

Figure 3 Error rate heatmap: comparison of sentence-type errors across L1 languages (see online version for colours)

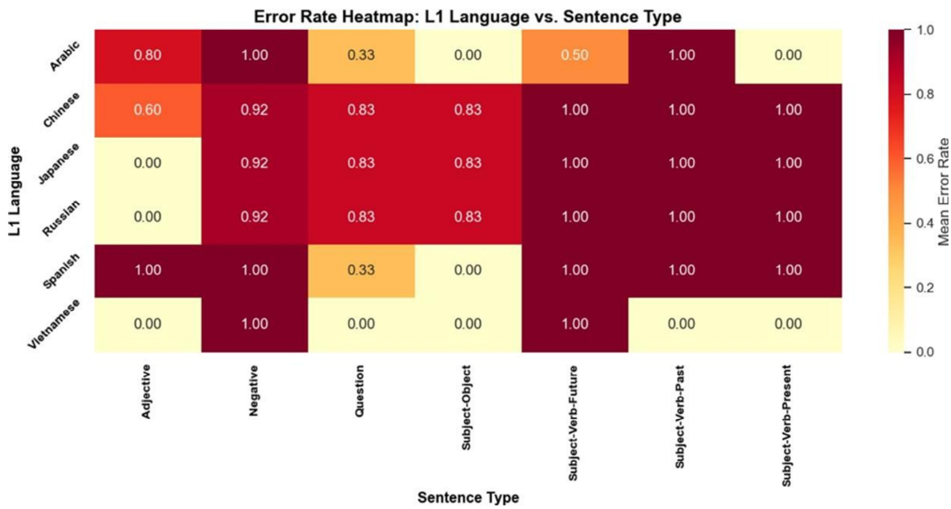


Figure 4 Analysing word usage trends in Korean language corpus (see online version for colours)

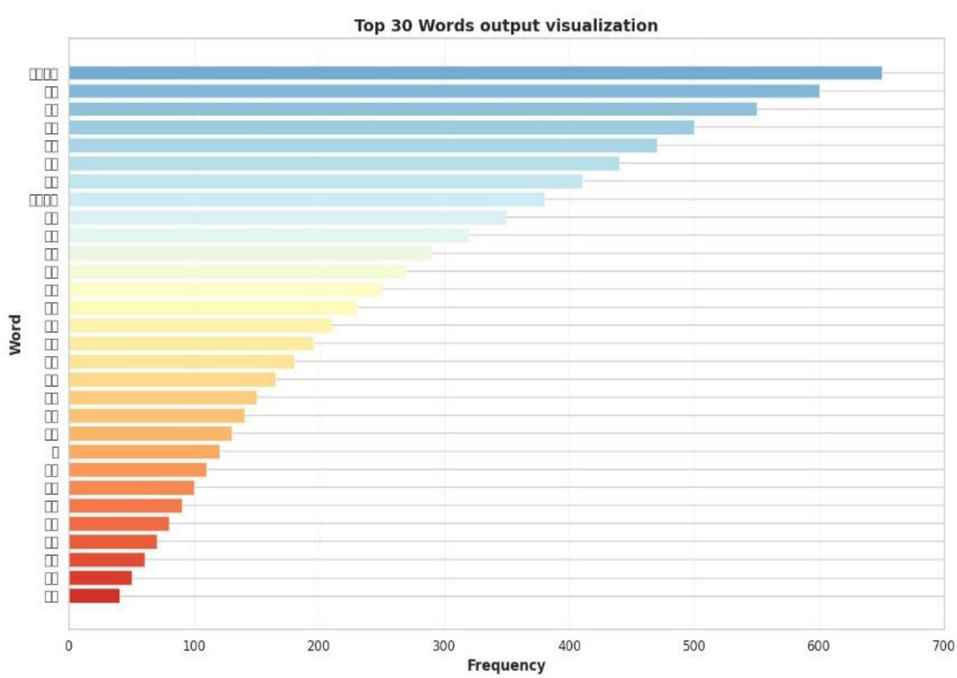


Figure 5 Error rate comparison across different L1 languages (see online version for colours)

4.3 *Metrics for evaluating the effectiveness of the proposed model*

- *Interaction ratio:* it measures the engagement or active participation between learners and the instructional system during the translation learning process known as the interaction ratio. It evaluates how often and how effectively students engage with the translation model or the learning platforms. It shows the system's ability to support learner attention maintenance needed for comprehending the complex skills of language translation. Finally, an interaction ratio becomes a meaningful metric of how effectively adapts educational technologies.
- *Translation prediction rate: i*t indicates how efficient and how accurate the machine learning model's predictions or generations of translation give a time span. It measures the time to translate from input text to completed translated output. This metric is used to evaluate the adaptability of the model to diverse linguistic inputs and can be used in multilingual domains where native languages of learners influence the performance of translation. It helps in determining both the computationally efficient and practical usability of automated language education systems.

The interaction ratio and translation prediction rate of the SSO-Attn-Seq2Seq model as illustrated in Figure 6. The interaction ratio increases from 78.8% to 94.1% indicating the degree of involvement or participation throughout the translation process. Similarly, the translation prediction rate indicates the effectiveness and precision of the translation model improved from 78.5% to 95.2%. This model performs better and students participate more successfully, and suggesting that the translation system was scalable and more effective in larger multilingual classroom situations. Table 5 illustrates the comparative analysis of interaction and prediction rates by student group size in multilingual translation settings.

Figure 6 Interaction ratio and translation prediction rate of SSO-Attn-Seq2Seq model (see online version for colours)

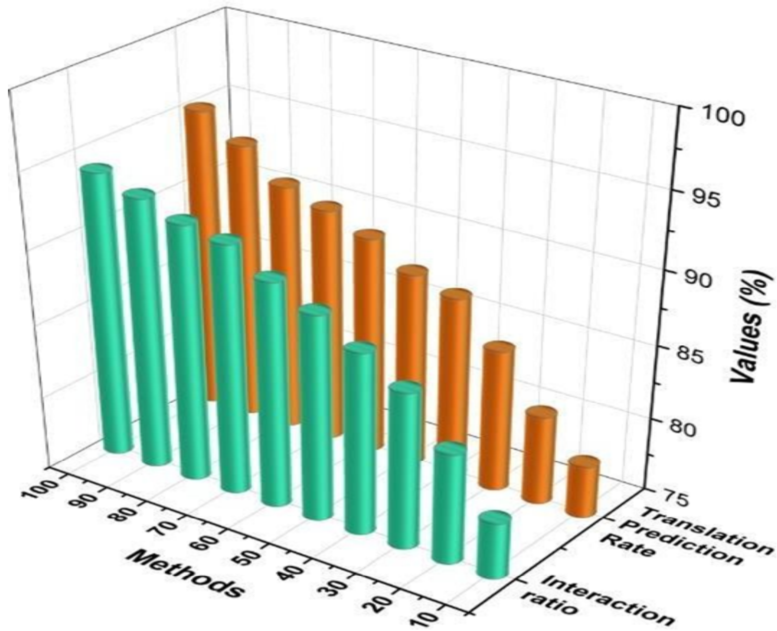


Table 5 Comparative analysis of interaction and prediction rates by student group size in multilingual translation settings

Number of students	*Interaction ratio (%)*	*Translation prediction rate (%)*
10	78.8	78.5
20	82.5	81
30	85.6	8.4
40	87.3	87.3
50	88.9	88.1
60	90.2	89.7
70	91.8	90.8
80	92.3	91.6
90	93.2	93.6
100	94.1	95.2

- *Accuracy:* measures the overall rate of correct predictions. Accuracy measures the percentage of the translated sentences that were actually translated among all the changes made by the model prediction.
- *Precision:* it indicates how many actual positive cases exist within all the predicted positive outcomes. This metric was essential to ensure the translations of the model predicts are correct or false as possible, staying away from low-quality translations, especially for error patterns unique to learners.

- *Recall:* measures the model's ability to detect all actual instances of a class. Recall measures how the correct translations identified. This metric is used to detect the model for accurate translation, detecting language nuances to provide comprehensive support for students with a wide variety of native languages to learn.
- *Fl-score:* balances precision and recallusing their harmonic mean. It offers a balanced assessment of the model's accuracy and completeness with respect to problems of translation and achieves a good translation to minimise errors. The performance metrics of the proposed SSO-Attn-Seq2Seq model across six different 11 languages was illustrated in Table 6. The model achieves 90.5%, F1-score of 89.8%, precision of 90.2%, and recall of 87% of Chinese indicating a strong overall translation correctness. These findings demonstrated how well the model handles a variety of language patterns and learner backgrounds in a multilingual translation context. These metrics confirm the model's effectiveness in acquiring Korean translations for diverse native language backgrounds. The evaluation of SSO-Attn-Seq2Seq models, precision, F1-score, accuracy and recall for translation effectiveness across different Languages was illustrated in Figure 7.

Figure 7 Evaluation of SSO-Attn-Seq2Seq models, precision, F1-score, accuracy, and recall for translation effectiveness across different languages (see online version for colours)

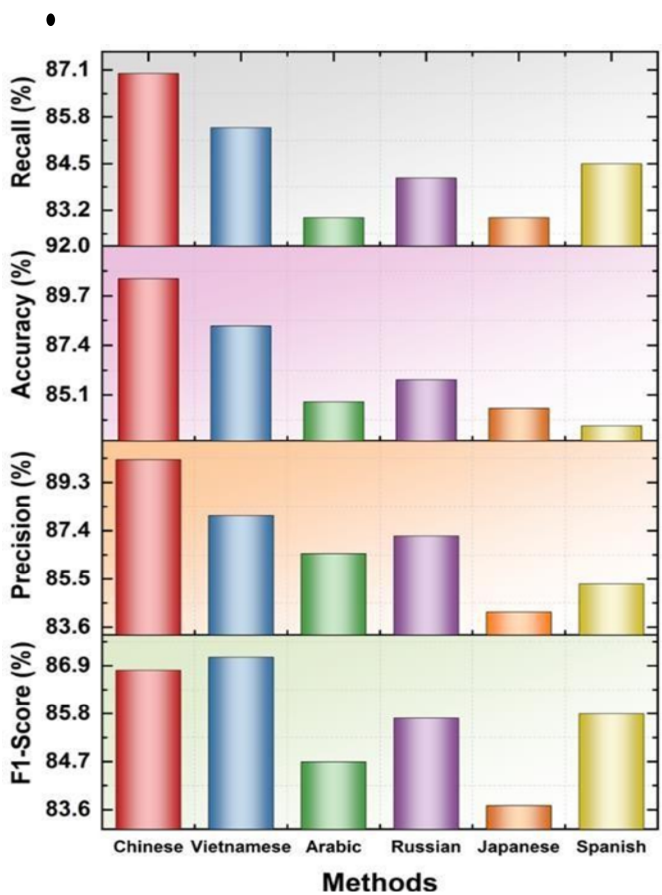


Table 6 Comparative performance of SSO-Attn-Seq2Seq models, and recall

Languages	*F1-Score (%)*	*Precision (%)*	*Accuracy (%)*	*Recall (%)*
Chinese	89.8	90.2	90.5	87.0
Vietnamese	87.1	88.0	88.3	85.5
Arabic	84.7	86.5	84.8	83.0
Russian	85.6	87.2	85.5	84.1
Japanese	83.7	84.4	84.9	83.0
Spanish	85.8	85.3	83.7	84.5

The performance metrics of a multilingual translation model evaluated with six different languages. By combining the average calculation of six languages to obtain the superior results in terms of accuracy (86.28%), F1-score (86.11%), recall (84.51%), and precision (86.93%) to indicate a strong translation correctness.

5 Discussion

A translation-based instructional model was a common platform used for understanding language and collaborative learning in the case of teaching a multilingual classroom. The Engagement with Language (EWL) framework showed the diversity of teaching methods and classroom behaviours in Korea (Lim and Kester, 2023). Since the participant's sample was quite limited, it might be hard to use the findings in other regions or contexts. Arranging teacher responses into disposition, acceptance, non-responsiveness and refusal could overstate the simplicity of classroom environments. It lacks the students' opinions about language play and the impact of how much individuals motivated for involvement. Additionally, the result findings of the EWL framework might differ the language development or teaching approaches. The EFL environment, instructors who speak English as their first language were more competent and legitimate to teach English exclusively (Yang and Jang, 2022). The ethnographic approach provides a lot of detail, its interpretation can be affected by the personal stance. It explores the interpretations of Korean bilingual teachers, with few perspectives considered from students or administrators which might show the impact of the language policy.

It was recognised that native-speakers, gendered nationalism, and professionalism have strong links, but the analysis might not analysed separately. Translanguaging techniques in vlogs produced by EFL for instructors (Im and Park, 2025). Korean language instructor teachers might prevent the cultural or educational influences on translanguaging with other nations. The content analysis was planned, it requires to interpretation of multimodal and translingual parts. Moreover, participants' experience with technology used for vlog production might determine the quality and involvement in multimodal communication. Nishioka (2023) investigated the self-directed language learning practices of learners. It limits the learning platform with the probable omitting of digital resources or social media that potentially takes part in self-instruction. The participant's self-reported practices effectively serve as a cornerstone for their analysis which could result in more subjectivity or bias. The broader impact of formal instruction or institutional support and the language instruction model was not explored in language translation. To overcome these limitations, the suggested SSO-Attn-Seq2Seq model enhances translation accuracy by adapting diverse learner outlines.

The SSO fine-tunes the model parameters, and the attention-based Seq2Seq framework ensures context-aware for accurate translations. By analysing learner-specific patterns such as native language influence and learning history, and the SSO-Attn-Seq2Seq hybrid model delivers personalised instruction with high accuracy and efficiency. A multilingual environment encourages insertion, and provides support for diverse teaching that helps Korean language learners to better understand the content. This model helps to make different languages and cultures more visible by making it simple for students to access the same information. Overall, the SSO-Attn-Seq2Seq approach offers a scalable and intelligent solution for real-time educational applications in linguistically diverse classrooms.

6 Conclusions

Language educators have debated the role of translation in language teaching for integrating translation into multilingual education. The multilingual classrooms was a crucial learning process that requires collaborative translation activities where learners negotiate meaning, strategise, and reflect on linguistic choices. Significant pedagogical obstacles arise in multilingual classroom settings, especially when students speak different native tongues and have differing degrees of Korean competence. The model analyses translation outputs, identifies learner-specific error patterns, and personalises instruction based on three core variables such as native language influence, historical translation accuracy, and individual learning progression. The dataset was collected from Kaggle. It includes six distinct L1 backgrounds such as Chinese, Vietnamese, Arabic, Russian, Japanese, and Spanish. Pre-processing included tokenisation, lemmatisation, and part-of-speech tagging to normalise the texts. Feature extraction includes word embeddings such as Word2Vec to transform words into vector representations. To improve the translation accuracy, the suggested model used SSO-Attn-Seq2Seq. The SSO model fine-tunes the model parameters, while the attention-enhanced Seq2Seq framework ensures context-aware and accurate translations. The syntactic and semantic complexity presented in multilingual translation assignments was especially well-served by the SSO-Attn-Seq2Seq technique. Extensive experiments demonstrated that the proposed SSO-Attn-Seq2Seq approach performs better than multimodal baseline architectures, achieving superior results in terms of accuracy (86.28%), F1-score (86.11%), recall 1(84.51%), and precision (86.93%) to indicate a strong translation correctness. highlight the potential of deep learning to enhance Korean translation instruction and offer a scalable, adaptive solution for multilingual learning environments.

Declarations

All authors declare that they have no conflicts of interest.

References

Almusharraf, A. and Bailey, D. (2023) 'Machine translation in language acquisition: a study on EFL students' perceptions and practices in Saudi Arabia and South Korea', *Journal of Computer Assisted Learning*, Vol. 39, No. 6, pp.1988–2003, https://doi.org/10.1111/jcal.12857.

Antony, S., Ramnath, R. and Ellikkal, A. (2024) 'Examining students' perspectives on pedagogical translanguaging in the multilingual classroom context', *IAFOR Journal of Education*, Vol. 12, No. 1, pp.199–223, https://doi.org/10.22492/ije.12.1.08.

Bang, J.U., Maeng, J.G., Park, J., Yun, S. and Kim, S.H. (2023) 'English-Korean speech translation corpus (EnKoST-C), Construction procedure and evaluation results', *ETRI J.*, Vol. 45, No. 1, pp.18–27, https://doi.org/10.4218/etrij.2021-0336.

Cho, E., Park, H.J.L. and Looney, S.D. (2024) 'Multimodality in third turn repetitions: evaluation, mitigation, and the pursuit of responses in a Korean-as-foreign-language classroom', *Language Teaching Research*, https://doi.org/10.1177/13621688241241037.

Choi, M.S. (2024) 'A Korean sojourning multilingual child's diary as translanguaging spaces', *Journal of Language, Identity & Education*, pp.1–17, https://doi.org/10.1080/15348458.2023.2298393.

Fedorova, K. and Nam, H.H. (2023) '"Multilingual islands in the monolingual sea": foreign languages in the South Korean linguistic landscape', *Open Linguist*, Vol. 9, No. 1, 20220238, http://dx.doi.org/10.1515/opli-2022-0238.

Gong, Y. (2022) 'Study on machine translation teaching model based on translation parallel corpus and exploitation for multimedia Asian information processing', *ACM Trans Asian Low-Resour Lang Inf Process*, https://doi.org/10.1145/3523282.

Han, S., Liang, D. and Haneda, M. (2021) 'A case study of two South Korean middle school EFL teachers' practices: instructional stances and use of classroom materials', *Classroom Discourse*, Vol. 12, Nos. 1–2, pp.56–74, https://doi.org/10.1080/19463014.2020.1869572.

Harris, S., Chen-Wu, H., Sue, E. and Lee, J.S. (2025) 'Affordances, challenges, and possibilities of Hallyu in the classroom: Korean language instructors' perspectives', *Language, Culture and Curriculum*, Vol. 38, No. 1, pp.38–56, https://doi.org/10.1080/07908318.2024.2381447.

Im, J.H. (2023) 'Translanguaging in action: a case study on two Korean students' engagement in translanguaging in radio-based English education', *J. English Teach. Movies Media.*, Vol. 24, No. 4, p.4156, https://doi.org/10.16875/stem.2023.24.4.41.

Im, J.H. and Park, G. (2025) 'Vlogs as Korean college students' translanguaging space', *STEM J.*, Vol. 26, No. 1, pp.29–45, https://doi.org/10.16875/stem.2025.26.1.29.

Jang, J. (2022) 'Translanguaging as an agentive action: a longitudinal case study of Uzbek EFL learners in South Korea', *Tesl-Ej*, Vol. 26, No. 3, N3, https://doi.org/10.55593/ej.26103a6.

Kim, H., Tosho, H., Moon, S., Okazaki, N. and Komachi, M. (2023) 'North Korean neural machine translation through South Korean resources', *ACM Trans Asian Low-Resour Lang Inf Process*, Vol. 22, No. 9, pp.1–22, https://doi.org/10.1145/3608947.

Kim, S.A. (2024) 'Harnessing Korean-English learners' funds of knowledge through culturally responsive readings', *International Journal of Language and Literary Studies*, Vol. 6, No. 3, pp.245–259, https://doi.org/10.36892/ijlls.v6i3.1704.

Lee, C. and García, G.E. (2021) 'Understanding Korean-American first-graders written translanguaging practices', *Linguistics and Education*, Vol. 66, p.100998, https://doi.org/10.1016/j.linged.2021.100998.

Lee, V. and Gyogi, E. (2022) 'Translation for multiliteracies: case studies of EFL and JFL classrooms', *Pedagogies: An International Journal*, Vol. 17, No. 2, pp.103–118, https://doi.org/10.1080/1554480X.2020.1826948.

Lee, Y.J., Kim, H.J. and Li, Y. (2022) 'International students' experiences of content language integrated learning in a Korean university: focusing on Korean as a medium of instruction', *Int. J. Educ. Methodol.*, Vol. 8, No. 3, pp.421–429, http://dx.doi.org/10.26803/ijlter.20.9.3.

Lim, E. and Kester, K. (2023) 'Korean middle school teachers' perceptions and teaching practices of multicultural education: a qualitative case study', *Int. J. Multicult. Educ.*, Vol. 25, No. 3, pp.67–87, https://doi.org/10.18251/ijme.v25i3.3603.

Lim, S., Park, C. and Yang, Y. (2022) 'Deep learning-based Korean dialect machine translation research considering linguistics features and service', *Journal of the Korea Convergence Society*, Vol. 13, No. 2, pp.21–29, https://doi.org/10.15207/JKCS.2022.13.02.021.

Nishioka, H. (2023) 'My Korean language teachers are YouTubers: learning Korean via self-instruction', *Comput. Assist. Lang. Learn.*, Vol. 36, No. 3, pp.346–374, https://doi.org/10.1080/09588221.2021.1928227.

Park, D. and Padó, S. (2024) *Multi-dimensional Machine Translation Evaluation: Model Evaluation and Resource for Korean*, arXiv preprint arXiv:2403.12666, https://doi.org/10.48550/arXiv.2403.12666.

Pierson, A.E., Clark, D.B. and Brady, C.E. (2021) 'Scientific modeling and translanguaging: a multilingual and multimodal approach to support science learning and engagement', *Sci. Educ.*, Vol. 105, No. 4, pp.776–813, https://doi.org/10.1002/sce.21622.

Ryu, J., Lee, K. and Yang, S. (2024) 'Korean-as-a-foreign-language learners' engagement with machine translation output', *RELC Journal*, https://doi.org/10.1177/00336882241298006.

Smith, C.A. (2024) 'Mapping reflexivity in situ: a multimodal exploration of negotiated textbook discourses in Korean university EFL classrooms', *Language Teaching Research*, Vol. 28, No. 4, pp.1368–1390, https://doi.org/10.1177/13621688211024932.

Son, S., Park, C., Lee, J., Shim, M., Lee, C., Park, K. and Lim, H. (2022) 'Korean and multilingual language models study for cross-lingual post-training (XPT)', *J. Korea Converg. Soc.*, Vol. 13, No. 3, pp.77–89, https://doi.org/10.15207/JKCS.2022.13.03.077.

Song, J. (2023) 'Korean language teachers' vulnerability over English competency in Korean-only classrooms', *Asian-Pacific J. Second Foreign Lang. Educ.*, Vol. 8, No. 1, p.16, https://doi.org/10.1186/s40862-023-00185-6.

Wilang, J.D. and Duy, T.V. (2021) 'Relationships of language learning variables in the acquisition of third languages in a multilingual context', *International Journal of Evaluation and Research in Education*, Vol. 10, No. 4, pp.1117–1124, http://dx.doi.org/10.11591/ijere.v10i4.21594.

Yang, J. and Jang, I.C. (2022) 'The everyday politics of English-only policy in an EFL language school: practices, ideologies, and identities of Korean bilingual teachers', *Int. J. Biling Educ. Biling*, Vol. 25, No. 3, pp.1088–1100, https://doi.org/10.1080/13670050.2020.1740165.

Yilmaz, T., Li, S. and Lee, Y.J. (2024) 'A multilingualism-as-resource approach in a graduate-level English-medium instruction (EMI) program at a Korean university', *Journal of Multilingual and Multicultural Development*, pp.1–14, https://doi.org/10.1080/01434632.2024.2314659.

Zhang, Z. (2023) 'Computational technologies for the construction of business Korean translation corpus based on association rules mining', *ACM Trans. Asian Low-Resour. Lang. Inf. Process*, Vol. 29, No. 4, https://doi.org/10.1145/36035002375-4702.