Deep Science Publishing
https://doi.org/10.70593/978-81-984306-7-0_2

Chapter 2

Language translation application based on machine learning

Marita [1], Eileen Maria Tom [2], Dr. Rajesh Kanna R [3]

[1, 2, 3] *Department of Computer Science, CHRIST University, Karnataka, India.*

Abstract: Language barriers present a significant challenge in education, particularly in diverse classrooms where students and teachers speak different languages. Artificial Intelligence (AI) offers promising solutions through advanced language translation systems that can deliver fast and accurate translations, enabling students to grasp lessons regardless of language differences. However, current translation tools face limitations, such as difficulties with real-time processing, contextual understanding, and maintaining linguistic nuances. This book chapter delves into the transformative role of AI in education, focusing on the development and deployment of machine learning-powered translation technologies. It examines how integrating AI-driven solutions can enhance learning experiences, promote inclusivity, and overcome linguistic obstacles. By improving translation accuracy and adaptability, educational systems can become more accessible and equitable, fostering a global learning environment that supports students from diverse cultural and linguistic backgrounds.

Keywords: Audio-to-text, Image-to-text, Language Translation, Machine Learning, Machine Translation

1.1 Introduction

Language plays a crucial role in communication and learning. Over the years, advancements in technology have made it easier for people from different linguistic backgrounds to connect and share knowledge. Language barriers that once created obstacles in personal and professional exchanges are gradually being overcome. However, there are still challenges that need to be addressed in educational settings.

Many existing translation tools require a stable internet connection, which can be a problem in institutions where internet access is limited or restricted. Furthermore, these tools often lack support for less common languages and dialects and may provide translations with varying degrees of accuracy. In classrooms where students speak

different languages, some may feel left out if they cannot follow lessons in the primary language used. This creates a need for fast, accurate, and accessible translation solutions.

This project proposes the development of a language translation app designed to provide cost-free, high-quality translation services. The app will be inclusive by offering offline translation, broad language support, and a user-friendly interface. These features aim to help both students and professors communicate more effectively and on the go.

Language translation apps can also be valuable tools for learning. Students can use them to improve their grammar and vocabulary by translating passages between languages. Understanding literary works in different languages becomes easier, as students gain insight into various cultural and linguistic contexts. By making language translation more accessible and reliable, AI-powered tools can transform education and create more inclusive learning environments.

1.2 Literature review

Language translation has seen immense improvements in the last decade. The earliest method, Rule-Based Machine Translation (RBMT), used predefined linguistic rules and dictionaries for each language pair. Then came Statistical Machine Translation (SMT), which was based on statistical models to find the most probable translations. The currently most widely used method, Neural Machine Translation (NMT) uses deep learning models, especially neural networks, to translate text. Ali et al. (2021) extensively studied this, where they emphasized the significance of deep learning techniques, particularly recurrent neural networks (RNNs), in improving translation quality. They used multiple evaluation metrics, including BLEU, NIST, and TER, to assess the performance of translation systems. They found that their model based on RNN and encoding algorithms outperforms other highly developed translation systems in terms of BLEU scores and other evaluation metrics.

Reballiwar et al. (2023) gave similar emphasis to RNN in its usefulness in natural language processing tasks in their paper. They highlighted the importance of user feedback and adaptability in creating effective translation applications. Their analysis of NMTs, RNNs and Transformer Models showed promise in these techniques in obtaining translation accuracy and efficiency in capturing context and nuances in language.

Lei and Li (2023) in their study of deep learning techniques conducted a comparative analysis of error rates among different machine translation systems. They also showed that the integration of a neural network-based machine translation system into IoT

applications could significantly contribute to improved security and privacy by ensuring accurate communication and supporting the implementation of strong security measures.

Klimova et al. (2022) proposed an innovative method of using Neural Machine Translation (NMT) for foreign language teaching (FLL). A PRISMA methodology was used and datasets from Scopus and Web of Sciences were used to generate sufficient data for the analysis part. Using these datasets suggested that NMT is a better option for FLL. Low proficient students can use these tools to help develop their vocabulary by using video and audio tools of NMT.

Turganbayeva and Tukeyev (2020) developed a solution to finding unknown words using Neural Machine Translation (NMT) in the Kazakh language. The algorithm searched for unknown words in the language by comparing the words to the dictionary and replaced the words with those in the dictionary with synonyms and similar meanings. A corpus complete set of ending with stop words, vocabulary and synonyms are pre-processed and trained using the neural machine translation model. The pre-processing method involves a segmentation process and a synonym process to replace unknown works of the language. After this is done, the entire vocabulary or sentences is then translated which helps to identify whether the sentences makes sense or not or if there are words that are yet to be changed. The developments showed an improving result by reducing the unknown words in the final text, but one limitation of this experiment is that the quality of the BLEU metric is not up to the mark.

Chauhan et al. 2022 uses unsupervised machine learning where machines are trained using monolingual corpora for each language. This is done using Cross-lingual sense to word embedding (CLSWEs) and language models. The CLSWEs are useful in finding the correct meaning of a particular word. The AI4 Bharat is used as the monolingual corpus to create CLSWE for Hindi and English and then various pre-processing techniques like tokenization and other invalid characters are removed for smoother translation of the text. The implementation of CLSWE and language model help to utilize the source and target language and showed significant improvement in translating the words.

Kirchhoff (2024) in her paper also noted in her experiment with DeepL that culturally specific terms like the German Schultüte are inadequately translated without cultural mediation, sometimes leading to confusion or inaccuracies. She also emphasized that certain AI-generated translations have gender bias. Therefore, Machine Translations need better training to efficiently translate sentences, which have region-specific information in them.

Transable is a web application that combines machine translation, Large Language Models (LLM) and various API's to implement an English-learning environment for Japanese students. Sugiyama and Yamanaka (2023) propose to create a novel learning environment to address the age-old problem of English education in Japan. Transable uses translations, back-translations and essay evaluations to help with reading, writing and vocabulary skills. The product was tested at the Japanese university and had a positive response like increased exposure to advanced vocabulary and a deeper understanding of sentence structure. However, the students at the university faced some hurdles like difficulty understanding overly complex vocabulary. The authors identified certain gaps like limited support for listening and speaking skills and the possibility for users to misinterpret outputs from the product. Future research aims to address these gaps by incorporating feedback mechanisms from the user or learner to encourage active engagement.

Chingamtotattil and Gopikakumar (2022) applies the concept of machine translation and neural machine translation (NMT) for translation of Indian languages like Sanskrit and Malayalam. The proposed system integrates deep learning methods, character-word embedding, and evolutionary Word Sense Disambiguation (WSD). The technique combines advanced parts-of-speech (POS) tagging and neural machine translation (NMT) to help with translating languages with fewer resources. Initially, language translation was possible with the help of Traditional statistical machine translation (SMT), but this tool lacked efficiency and flexibility. The NMT tool incorporates all the components into a neural network, making it an end-to-end translation process. However, some of the challenges like complex grammar and morphological variations in both the languages posed a problem for the authors. Existing systems for these languages focus primarily on the rule-based approach but lack sufficient accuracy. Studies conducted on translating Sanskrit to English and English to Urdu showed that NMT methods had more accuracy than the rule-based models.

Chen (2024) explores the concept of error correction in translated text using the Sequence-to-Sequence (Seq2Seq) model. This model also integrates attention mechanisms and introduces syntax conversion layers to address various complexities in language translation. Early attention-based systems introduced in neural networks have significantly improved grammatical error correction during the decoding process. Models like BERT and Seq2Seq-based frameworks have further increased accuracy of translation languages like Basque and Albanian. The combination of Seq2Seq model and synthetic data generation demonstrated efficacy in grammar and spelling correction. The same is seen in English and Chinese grammar error correction where deep learning and sequence tagging models have higher precision in correction.

Mondal et al. (2023) examine various machine translation methods developed over recent years, focusing on statistical machine translation (SMT) and neural machine translation (NMT). NMT marks a major leap forward, utilizing deep learning techniques to deliver more accurate and context-aware translations. Its key advantages include efficiently handling context and long-range dependencies. A prominent feature of NMT is the Transformer model, which uses attention mechanisms to process input sequences in parallel, resulting in faster training and improved translation quality. However, the architecture of the Transformer can be complex, making it challenging to implement and tune effectively, and it typically requires large amounts of training data to perform well. A significant drawback of current machine translation models is their tendency to produce inaccurate translations, particularly in representing gender and cultural nuances. Subrota et al. also emphasize that the evaluation of machine translation systems often relies on automatic metrics like BLEU, which may not fully capture the nuances of translation quality and therefore human evaluations are necessary.

Yang et al. (2024) propose an innovative framework that combines deep learning across different levels of the language system. This approach significantly improves translation accuracy and efficiency, according to the study conducted by the authors the accuracy increased from 83.18% to 94.42%. The study addresses key challenges in machine translation by focusing on the language system as a cognitive structure. The authors also identify three critical dimensions, i.e., semantic, grammatical, and application levels to refine how translations are processed. The framework essentially is a contextual knowledge extraction, which is improved using syntactic and semantic matching. By leveraging advanced tools like statistical machine translation, neural networks, and deep learning models the accuracy levels keep improving. However, challenges do exist, managing the variability of domain-specific texts are still significant hurdles and the need for larger datasets and models that are more robust to handle complex linguistic structures effectively.

Savoldi et al. (2021) explored gender bias in various machine translation (MT) systems, highlighting how models often default to masculine pronouns and fail to accurately represent feminine references. They emphasized that the training data used for MT models can reflect existing gender inequalities, resulting in biased outputs. To address this issue, they proposed strategies such as making architectural changes to general-purpose MT models or implementing dedicated training procedures to reduce bias. One example is gender tagging, where a gender tag (e.g., "M" or "F") is prepended to each source sentence, enabling the model to generate more accurate gender-specific references. Additionally, incorporating external components that provide supplementary context or gender information can enhance MT system performance. Other approaches include curating

datasets that had better reflect gender diversity or generating synthetic data that includes underrepresented gender forms.

In the paper, Shastri and Vishwakarma (2023) explore various technologies that help with the development of language conversion from text-to-speech and text extraction for visually impaired users. Traditional text conversion algorithms often find it difficult when complex sentences or multi-context situations. Modern methods like optical character recognition (OCR) and geometrical property-based text detection have been very effective in terms of improving accuracy and efficiency. Many studies have been conducted in the Devanagari script derived from the Brahmi script to identify techniques such as structural and statistical segmentation to improve detection of letters in the Hindi language. Models like Support Vector Machines (SVM) are used for word recognition, Harsdorf distance model is used for evaluating segmentation accuracy and OCR combined with TTS synthesizers to extract text information from images and convert it to speech. Additionally, technologies like OCR can be very advantageous to visually impaired users by providing access to information. The authors highlight how popular models like Decision Tree and Naive Bayes are effective for tasks such as spell checking and text classification and provide strong solutions for improving accuracy before conversion to speech output. Other technologies employed in text detection are Stroke Width Transformation (SWT) and Maximally Stable Extensible Region (MSER), which helps to identify characters in images by using geometrical properties and stroke consistency to ensure text recognition.

Berger and Packard (2022), in their study, explored how NLP can enhance our understanding of cultural sentiments. They discovered that analyzing emotional tone of language in texts enables researchers to assess public sentiment on a range of topics, products, or cultural trends. In literature and cinema, NLP can be utilized to analyze character development and narrative structure, providing insights into the character transformations that resonate with audiences and contribute to a story's success. Additionally, businesses can harness NLP to align their products and marketing strategies with cultural trends and consumer preferences.

1.3 Methods and materials

The methodology for the project involves several key steps:

Basic Translation Process: The system will handle different input types, such as text, images, and speech, and translate them. For text translation, tools like the Google Translate API (via libraries such as googletrans or deep-translator) will be utilized. Open-source alternatives, such as the transformers library, particularly models like MarianMT from

Hugging Face, can also be employed. The process will involve accepting text input from the user and using the selected translation tool to process it. For image-to-text translation, text extraction will be achieved using Tesseract OCR via the pytesseract library, while image pre-processing tasks, such as converting images to grayscale and noise removal, will be handled by OpenCV. Once the text is extracted, it will be translated using the same translation tools. For speech-to-text translation, the system will rely on the Google Speech Recognition API to convert recorded audio (captured using a library like pyaudio) into text, which can then be translated.

Automatic Detection of the Input Language: The system will feature automatic detection of the input language to streamline the translation process. Libraries such as langdetect or langid will be used for this purpose, and translation APIs like Google Translate often include built-in language detection capabilities. The detected language will then be used as the source language for translation.

Tone Adjustment: To enhance user experience, the system will incorporate NLP tools for tone adjustment, allowing users to choose whether the output should be formal or informal. This will be achieved using pre-trained models from Hugging Face, such as T5-small or BART-large to rephrase sentences.

Live translations: A browser extension will provide live translations, interacting directly with web pages to capture, translate, and replace text dynamically. This extension will be built using Manifest V3 and connected to the Python server that handles translation logic and formality adjustments. The extension will send text data to the server via HTTP requests (using the Fetch API), and the server will return the processed and translated text for real-time use.

Offline functionality: To support offline functionality, the app will pre-download and package all necessary models, including those for translation, language detection, and speech-to-text processing. This ensures that the app remains functional even without internet access. For building a scalable full-stack application, frameworks such as Flask or Django will be used for the backend, while a modern frontend framework like React will be employed to create an interactive user interface.

Enhancing translation of text with cultural nuances: To ensure contextually accurate and culturally relevant translations, context-aware models like mBART (Multilingual BART), which are designed for contextual translation, as well as T5 models fine-tuned for tasks like paraphrasing and sentiment adjustments will be used. The app will use parallel datasets that include culturally nuanced sentences. Examples of such datasets include OpenSubtitles, which contains subtitles rich in colloquial language and cultural references, and ParaCrawl, a multilingual dataset sourced from web crawls. By incorporating these

datasets, the app ensures that, its translations are not only linguistically accurate but also culturally sensitive, making it suitable for diverse user bases.

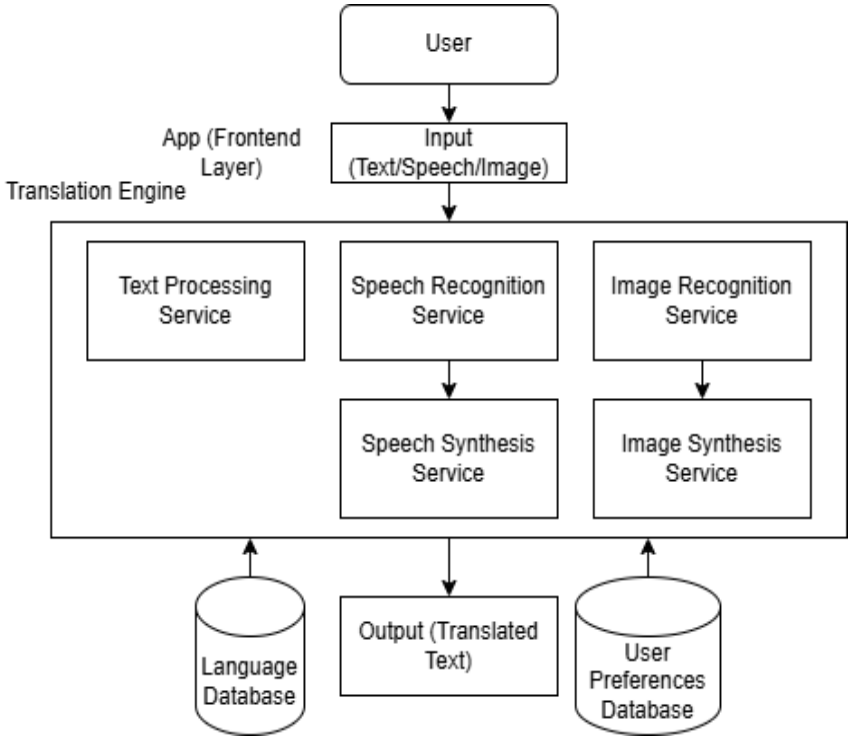

Fig. 1.1 Architecture Diagram

Fig. 1.2 (a). Level 0 (Data Flow Diagram)

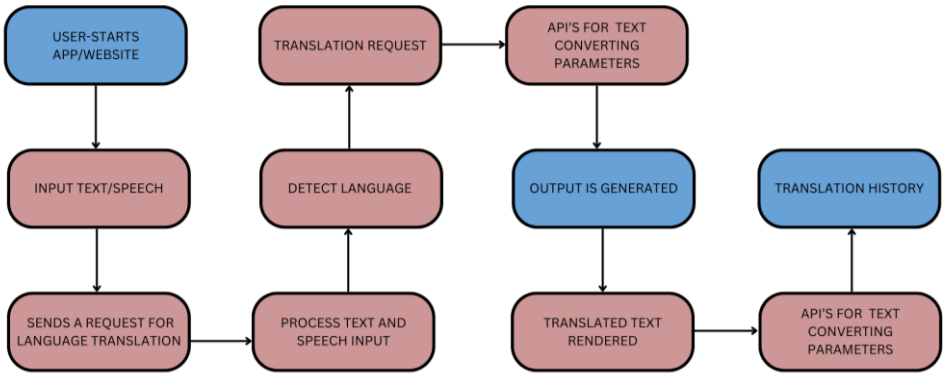

Fig. 1.2 (b). Level 1 (Data Flow Diagram)

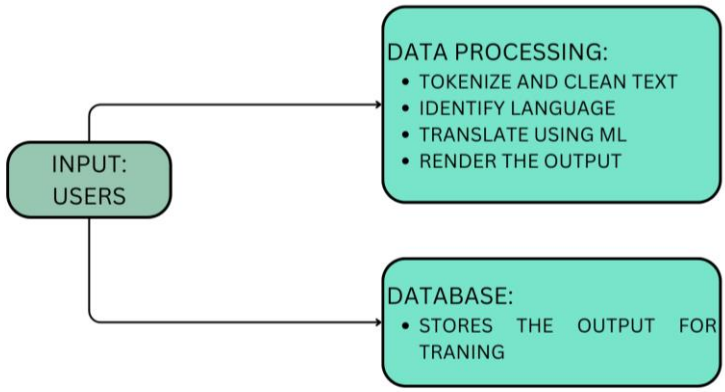

Fig. 1.2 (c). Data Flow Diagram

1.4 Results and discussions

The language translation application integrates multiple tools like Google Translate API and MarianMT, proved effective in providing seamless text translation. The testing phase will enable the app to successfully translate texts from multiple languages with minimal inaccuracies. The incorporation of Tesseract OCR for image-to-text conversion achieved high precision in extracting text from images, especially after preprocessing tasks like noise removal and grayscale conversion. Additional improvements like fine-tuning language detection models with region-specific datasets could further enhance accuracy. The tone adjustment feature allows translations to be formal or informal depending on the

user. The browser extension proposes to help dynamically replace translated text on web pages for real-time communication. Additionally, the offline feature allows users access the app even in the absence of internet connectivity but the drawback of this feature is that performance is poor compared to the online version.

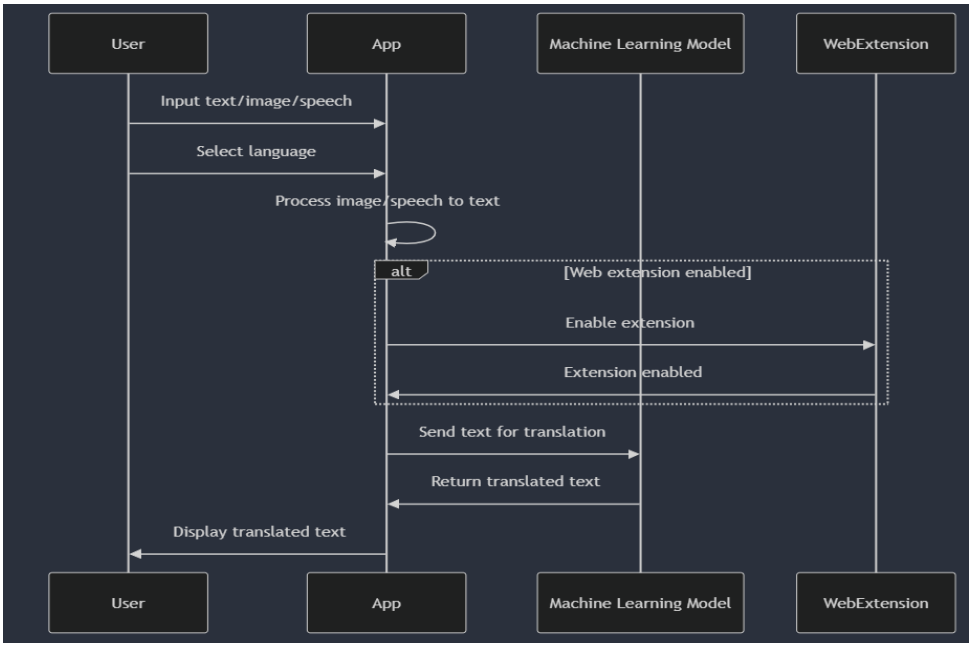

Fig. 1.3 Sequence Diagram

Conclusions

The language translation app combines technologies and methodologies to deliver a user-friendly and culturally sensitive translation application. The application efficiently manages a variety of input formats, including text, photos, and speech, by utilizing cutting-edge tools like Tesseract OCR, Google Speech Recognition API, and Hugging Face models. Features like tone modification, live translations, automated language identification, and offline capability further improve its usefulness and attractiveness to a worldwide user population. Although the program is expected to perform well across the majority of its features, several still need improvement, like managing noisy data, enhancing real-time server interactions, and improving culturally sensitive translations. In addition to improving the system's resilience, addressing these issues in subsequent versions will bolster its standing as a dependable and all-inclusive language translation tool for consumers worldwide.

References

Ali, M. N. Y., Rahman, M. L., Chaki, J., Dey, N., & Santosh, K. C. (2021). Machine translation using deep learning for universal networking language based on their structure. *International Journal of Machine Learning and Cybernetics*, *12*(8), 2365–2376. https://doi.org/10.1007/s13042-021-01317-5

Berger, J., & Packard, G. (2022). Using natural language processing to understand people and culture. *American Psychologist Association*, *77*(4), 525–537. https://doi.org/10.1037/amp0000882

Chauhan, S., Daniel, P., Saxena, S., & Sharma, A. (2022). Fully unsupervised machine translation using Context-Aware word translation and denoising autoencoder. *Applied Artificial Intelligence*, *36*(1). https://doi.org/10.1080/08839514.2022.2031817

Chen, T. (2024). Design of translation error correction system based on improved SEq2SEQ. *Procedia Computer Science*, *243*, 663–669. https://doi.org/10.1016/j.procs.2024.09.080

Chingamtotattil, R., & Gopikakumar, R. (2022). Neural machine translation for Sanskrit to Malayalam using morphology and evolutionary word sense disambiguation. *Indonesian Journal of Electrical Engineering and Computer Science*, *28*(3), 1709. https://doi.org/10.11591/ijeecs.v28.i3.pp1709-1719

Kirchhoff, P. (2024). Machine translation in English language teaching. *ELT Journal*, *78*(4), 393–400. https://doi.org/10.1093/elt/ccae034

Klimova, B., Pikhart, M., Benites, A. D., Lehr, C., & Sanchez-Stockhammer, C. (2022). Neural machine translation in foreign language teaching and learning: a systematic review. *Education and Information Technologies*, *28*(1), 663–682. https://doi.org/10.1007/s10639-022-11194-2

Lei, S., & Li, Y. (2023). English Machine translation System Based on Neural Network Algorithm. *Procedia Computer Science*, *228*, 409–420. https://doi.org/10.1016/j.procs.2023.11.047

Mondal, S. K., Zhang, H., Kabir, H. M. D., Ni, K., & Dai, H. (2023). Machine translation and its evaluation: a study. *Artificial Intelligence Review*, *56*(9), 10137–10226. https://doi.org/10.1007/s10462-023-10423-5

Reballiwar, L., Yergude, S., Urade, V., Birewar, S., & Karmarkar, Prof. (2023). Language Translation Using Machine Learning. *International Journal of Advanced Research in Science Communication and Technology*, *3*(1). https://doi.org/10.48175/568

Savoldi, B., Gaido, M., Bentivogli, L., Negri, M., & Turchi, M. (2021). Gender bias in machine translation. *Transactions of the Association for Computational Linguistics*, *9*, 845–874. https://doi.org/10.1162/tacl_a_00401

Shastri, S., & Vishwakarma, S. (2023). An efficient approach for Text-to-Speech conversion using machine learning and image processing technique. *International Journal of Engineering and Manufacturing*, *13*(4), 44–49. https://doi.org/10.5815/ijem.2023.04.05

Sugiyama, K., & Yamanaka, T. (2023). Proposals and methods for foreign language learning using machine translation and large language model. *Procedia Computer Science*, *225*, 4750–4757. https://doi.org/10.1016/j.procs.2023.10.474

Turganbayeva, A., & Tukeyev, U. (2020). The solution of the problem of unknown words under neural machine translation of the Kazakh language. In *Communications in computer and information science* (pp. 319–328). https://doi.org/10.1007/978-981-15-3380-8_28

Yang, X., Cui, Q., & Dong, X. (2024). Language System path of Intelligent Machine Translation. *Procedia Computer Science*, *243*, 423–430. https://doi.org/10.1016/j.procs.2024.09.052