YORUBA-ENGLISH NUMBER CONVERSION USING TRANSFORMER MODEL

LAWRENCE BUNMI ADEWOLE
Department of Computer Science,
Federal University Oye-Ekiti,
Ekiti State, Nigeria
lawrence.adewole@fuoye.edu.ng|

FAYOKE MARY MESE
Department of Computer Science,
Federal University Oye-Ekiti,
Ekiti State, Nigeria
fayokemese@gmail.com

STEPHEN ALABA MOGAJI
Department of Cybersecurity,
Federal University Oye-Ekiti,
Ekiti State, Nigeria
stephen.mogaji@fuoye.edu.ng

TIMILEHIN VINCENT ADEWOLE
Department of Software Engineering
Obafemi Awolowo University, Ile-Ife,
Osun State, Nigeria
tvadewole@oauife.edu.ng

*Corresponding Author: lawrence.adewole@fuoye.edu.ng
(+2348066407161)

Abstract
Numbers are essential part of communication, yet machine translation models often struggle with its accurate translation, especially in low-resource languages. Yorùbá language, spoken by over 40 million people, has no published work that addressed the computational translation of Yorùbá numerals into English or Arabic digits. This paper presents a Transformer-based model for the conversion of Yorùbá cardinal numerals to their English equivalents. The research utilised the fine-tuned Flan-T5 'small' model for its computational efficiency and adaptability across text-to-text tasks. A dataset of 50,000 Yorùbá numerals was systematically generated using a rule-based algorithm and partitioned into 80% training and 20% testing sets. The model demonstrated remarkable computational efficiency, performing inference on

30 samples per second. Performance was evaluated using multiple metrics: accuracy, Character Error Rate (CER), Word Error Rate (WER), and Bilingual Evaluation Understudy (BLEU) score. The fine-tuned model achieved exceptional results: 99.96% accuracy, 0.000068 CER, 0.00010 WER, and 0.9999 BLEU. The near-perfect accuracy confirms the model's ability to correctly translate the vast majority of Yorùbá numerals. The extremely low CER reflects precise character-level generation, while the minimal WER indicates outstanding performance in predicting complete numeral words, essential for accurate translation. The BLEU score approaching 100% demonstrates that model outputs are nearly identical to reference translations, further validating translation fidelity. This work constitutes the first computational model for Yorùbá-to-English numeral translation, achieving state-of-the-art performance. The model is readily applicable to downstream NLP tasks, particularly text normalisation in text-to-speech systems, thereby contributing to language technology development for Yorùbá and serving as a template for similar low-resource languages.

Keywords: Transformer, Flan-T5, Numeral Translation, Machine Translation, Yoruba Language

Introduction

Numerals are important aspect of human communication, serving as the basis for expressing quantities, financial transactions, temporal references, and mathematical concepts across all languages (Eludiora & Abubakar, 2021). In an increasingly interconnected world, the accurate translation of numerals between languages is essential for cross-border information exchange, international commerce, scientific collaboration, and the development of inclusive digital technologies. Numerals, however, present unique challenges for computational linguistics due to their language-specific structural properties, morphological complexity, and the precision required for accurate translation, where a single error can fundamentally alter meaning (Akinadé & Ọdẹjọbi, 2014).

The Yoruba language has been classified as a potentially endangered language, meaning that, it has fewer linguistic resources available, such as digital content, annotated texts, and language technology tools (Mirela, 2024). This decline can be attributed to several factors, including the impact of colonialism, the dominance of English as a lingua franca in Nigeria, and

the perception of Yoruba as a "vernacular" language –(Akinkurolere & Akinfenwa, 2018; Osuntokun, 2024). The scarcity of resources dedicated to the Yoruba language significantly hampers the development of robust language models and tools. This limitation affects the creation of accurate and comprehensive linguistic technologies and also exacerbates the challenges in preserving and promoting the language. As a result, the lack of adequate resources and technological support contributes to the gradual decline of Yoruba, making it increasingly difficult for speakers to access and utilise digital tools that could aid in its revitalisation and daily use. Efforts to preserve and promote Yoruba culture and language are crucial.

Numerical conversion between languages significantly impacts commerce, especially in the context of cross-border e-commerce. In the consumer perception, precise translation and numerical conversion significantly boost consumer trust and perception. When prices and product details are accurately converted and displayed in the local language, it enhances the shopping experience and strengthens brand confidence "(Tang & Li, 2023). Also, language and numerical accuracy is essential in localization of news content within the media space, marketing and advertising as it boosts engagement and conversion rates (Nina, 2022). Consumers often compare numerical values, such as prices and product specifications, across different platforms. Accurate conversion guarantees that comparisons are fair and relevant, thereby impacting purchasing decisions "(Kara et al., 2020). Prioritising precise numerical conversion and localisation enables businesses to boost their global market competitiveness, elevate customer satisfaction, and foster growth.

Many Yoruba oral traditions, such as proverbs and folklore, incorporate numerals. Translating these numerals can help document and preserve these traditions for future generations (Adetomiwa, 2023). Facilitating the translation of Yoruba numerals can promote cultural exchange by making Yoruba cultural practices more accessible to a global

audience. This can lead to a greater appreciation and understanding of Yoruba heritage, and also foster a sense of pride and identity among community members '(Adetomiwa, 2023; Akinlade et al., 2021).

Review of Related Literature

Yoruba to English numeral conversion involves translating numerical expressions from the former (Yoruba) language to the later (English) language; this is particularly difficult because of the structural and syntactic differences between the two languages. Recent developments in natural language processing (NLP) have introduced transformer models '"(Ferraris et al., 2025), which have shown significant promise in handling such complex translation tasks. In order to navigate between these two languages, it is therefore expedient to consider the numerical systems of both languages.

English and Yoruba Numeral System

The English language employs a base-10 system, characterised by its straightforward and linear structure. Numbers are typically expressed in a simple and direct manner, such as "twenty-one" for 21 or "thirty-five" for 35. Yoruba numerals, on the other hand, are unique and frequently use a base-20 system. Numbers as 20, 200, 2,000, and 20,000 are used for counting in the system. Many numbers are factored as multiples of these numbers. Numbers are often formed by adding and subtracting from multiples of these base numbers (Adetomiwa, 2023), with preference for subtraction over addition (Akinadé & Ọdẹjọbí, 2014). For instance, 176 is obtained by multiplying 20 by 8, and then subtracting 4 (i.e. (20 x 8) – 4); 500 is obtained by multiplying 200 by 3, and then subtracting 100. (i.e. (200 x 3) – 100), the number 34 in Yoruba is thus, expressed as *"ẹrindínlà*

Yoruba: "ookan" - English: "one"
Yoruba: "ogun" - English: "twenty"
Yoruba: "aadota" - English: "fifty"

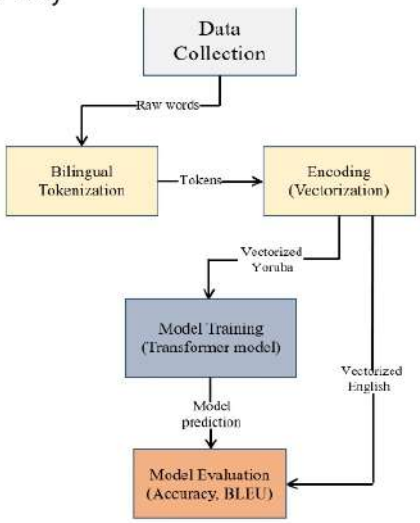

Figure 1: System Overview

Tokenisation

The individual characters or sub-words of both Yoruba and English numerals were split into smaller units called tokens. The process was applied to facilitate their conversion into numerical vectors suitable for input into the deep learning model. The tokenisation was done using the tokeniser associated with the Flan-T5 model, which is Byte-Pair Encoding (BPE) tokeniser, designed to handle multiple languages, including Yoruba. The FLAN-T5 is a sequence-to-sequence (seq2seq) model, which means it consists of an encoder and a decoder. The encoder processes the input text and converts it into a series of hidden states, while the decoder generates the output text based on these hidden states.

To ensure uniformity in input data, all tokenised sequences were truncated or padded to a fixed length of 64 tokens. This length was chosen based on the maximum length of the numeral sequences in the dataset. This

process ensured that all sequences were of equal length, allowing for efficient batch processing during model training.

Model Selection and Training

The Flan-T5 model, a Transformer-based system pre-trained on diverse text, was chosen for numeral translation due to its versatility in various NLP tasks. The 'small' variant was selected because it balances accuracy with efficiency, requiring less computational power and memory while retaining the core capabilities of larger versions. This makes it ideal for resource-limited applications that still demand high-performance. Its encoder-decoder architecture, powered by self-attention mechanisms, enables effective handling of translation tasks by capturing context and meaning. The encoder processes Yoruba numerals into semantic representations, while the decoder generates corresponding English outputs. Self-attention ensures the model focuses on relevant input segments, crucial for accurately translating varied numeral formats and contexts.

Fine-tuning adapted the pre-trained Flan-T5 to the specific task of numeral translation. A specialised dataset of Yoruba-English numeral pairs was used, allowing the model to learn translation rules and patterns. The dataset was split into training (80%, 40,000 pairs) and validation (20%, 10,000 pairs) sets using random splitting to ensure representativeness and reduce bias. Training involved five epochs, a learning rate of 0.0001, the Adam optimiser, and cross-entropy loss. These settings optimised the model's parameters to minimise prediction errors and improve accuracy. The system was implemented using PyTorch, which offers flexibility through dynamic computational graphs that allow real-time adjustments during fine-tuning. This streamlined the process and enabled efficient adaptation of the model. Training loss and validation loss were used to evaluate the performance of the proposed system. The Yoruba numerals to English numerals translation model was evaluated using standard metrics, which include accuracy, word error rate, character error rate, and bilingual evaluation understudy.

Results and Discussion

The fine-tuned Flan-T5 model achieved an impressive accuracy of 99.96% on the validation set. This means that out of all the predictions made by the model during the validation phase, 99.96% were correct. The result shows a high accuracy which indicates that the model was able to correctly translate the vast majority of Yoruba numerals into their corresponding English numerals. This remarkable performance is a testament to the model's robustness and the effectiveness of the fine-tuning process, which involves adjusting the model's parameters based on a specific dataset to enhance its accuracy and generalisation capabilities. Achieving such a high accuracy on the validation set suggests that the model is well-prepared to perform similarly well on unseen data, making it a reliable tool for practical applications.

Character Error Rate (CER)

Character Error Rate (CER) is a metric used to assess the number of character-level errors made by the model in its predictions. It is calculated as the ratio of the number of character substitutions, insertions, and deletions that is required to transform the predicted sequence into the correct sequence, to the total number of characters in the correct sequence. The model achieved a CER of **0.000068**. This extremely low CER reflects the model's precision at the character level, meaning it made very few character-level errors in its predictions.

Word Error Rate (WER)

Word Error Rate (WER) is a metric similar to CER but applied at the word level. It measures the percentage of words in the predicted sequences that are incorrect, including substitutions, insertions, and deletions. The WER for the model was calculated to be **0.00010**. The low WER indicates that the model performed exceptionally well at predicting entire words, which is crucial for accurate numeral translation.

BLEU Score

The BLEU (Bilingual Evaluation Understudy) score is a metric used to evaluate the quality of text that has been machine-translated from one language to another. It compares the predicted output with one or more reference translations and calculates a score based on the number of n-grams (sequences of words) that match between the predicted and reference texts. The BLEU score achieved by the model was **99.99.** A BLEU score close to 100% suggests that the model's translations were nearly identical to the reference translations, further confirming the model's high accuracy. The summary of the results is analysed in Table 1.

Table 1: Results Summary

Metric	Value
Accuracy	99.96%
CER	0.000068
WER	0.00010
BLEU score	0.9999

Error Analysis

Despite the high overall performance, the model did make a small number of errors. Analysing these errors helps to understand the limitations of the model and provides insights into potential areas for improvement. Table 2 showed that out of 10,000 samples in the validation set, the model misclassified 4 samples. These misclassified samples were examined to identify patterns or commonalities that may have led to errors.

The review of the misclassified samples revealed that these errors predominantly occurred in numerals with less complex structures. Three out of the four errors by the model occurred on Yoruba numerals with simple structures rather than more complex ones. This bias may be attributed to the composition of the dataset.

Table 1: Misclassified Samples

Input	Actual	Predicted
Igba	two hundred	two thousand
oke kan o le ni egberun	twenty-one thousand	twenty thousand
Eeje	Seven	one hundred and seven
Eejo	Eight	Six

The dataset had fewer examples of simple numerals (0 to 200) compared to more complex numerals, leading the model to encounter and learn more from complex examples. As a result, the model became better at handling complex numerals but occasionally struggled with simpler ones. While this bias did not affect the overall accuracy in any significant way, it is a consideration for the generalisation of the model, especially in scenarios where the translation of simpler numerals is crucial.

Inference Speed

Inference speed, which refers to the time it takes for a model to process input data and generate output, is a critical factor in evaluating the practical applicability of the model. This is particularly in real-time or large-scale translation systems, where rapid and efficient processing is essential. For instance, in real-time translation applications, users expect immediate responses, and any delay can significantly impact the user experience. Similarly, in large-scale systems that handle vast amounts of data, slow inference speeds can lead to bottlenecks and reduced overall efficiency. Therefore, a model with high inference speed ensures that translations are delivered promptly and efficiently, making it more suitable for practical, real-world applications where time and computational resources are critical.

The model achieved an inference speed of 35 samples per second. This speed is quite efficient, indicating that the model can translate Yoruba

numerals to English numerals at a rate suitable for real-time applications. However, for very large-scale deployment, further optimisation may be necessary to meet higher throughput demands.

Visualisation of Results

Visualisation provides a clear and intuitive understanding of the model's performance. By representing data graphically, it allows users to quickly grasp how well the model is performing. This visual approach can highlight patterns, trends, and anomalies that might not be immediately apparent when looking at numerical metrics alone. For instance, a confusion matrix can show which classes are frequently misclassified, while a loss curve can indicate whether the model is overfitting or under-fitting. These visual tools help in diagnosing issues, understanding model behavior, and making informed decisions about further improvements. Overall, visualisation enhances the interpretability of the model's results, making it easier to communicate findings and insights to a broader audience.

Training and Validation Loss Curves

The training and validation loss curves were plotted to monitor the model's performance during fine-tuning. Loss curves are graphical representations that show how the loss (or error) of a model changes over time during training. Typically, there are two types of loss curve; the training loss curve which shows the loss on the training dataset, and the validation loss curve which shows the loss on the validation dataset, and is used to evaluate the model's performance on unseen data.

The loss curves, as shown in Figures 2 and 3 showed a steady decline over the 5 epochs, with the validation loss closely tracking the training loss. This suggests that the model was learning effectively and improving its performance as it trains. That the validation loss closely tracks the training loss, suggests that the model is generalising well to new, unseen data. In other words, the model's performance on the validation set is similar to its performance on the training set. This is a good sign because it indicates that

the model is without overfitting (i.e., it is not just memorising the training data but is also learning patterns that generalise to new data).

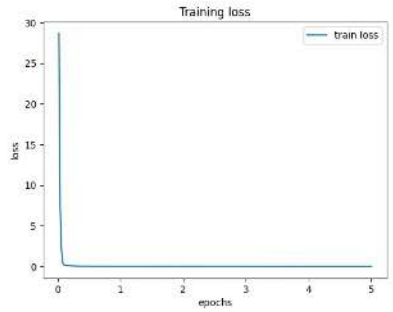

Figure 2: Training Loss

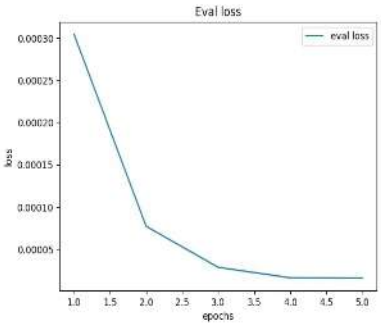

Figure 3: Validation Loss

Summary of Results

The results demonstrate that the fine-tuned Flan-T5 model is highly effective at translating Yoruba numerals to English numerals, with an accuracy of 99.96%, a Character Error Rate of 0.000068, a Word Error Rate of 0.00010, and a BiLingual Evaluation Understudy score of 99.99. The model's performance is further supported by its fast inference speed of 35 samples per second. While the model exhibited an infinitesimal bias towards more complex numerals, this did not significantly impact its overall accuracy. The visualisation of results provided additional insights into the model's learning process and error patterns.

Discussion of Key Findings

The study's results demonstrate the effectiveness of the fine-tuned Flan-T5 model in translating Yoruba numerals to English numerals. The high accuracy, low error rates, and near-perfect BLEU score all underscore the model's capability in this task. Several key findings are discussed below Effectiveness of Fine-Tuning. One of the most significant findings is the success of the fine-tuning process. By adapting the pre-trained Flan-T5 model to the specific task of numeral translation, the model achieved 99.96% accuracy, which is exceptional for a task involving language translation.

The success of the fine-tuning process indicates that pre-trained models like Flan-T5 can be effectively adapted to specialised tasks with limited training data. This opens up possibilities for applying similar approaches to other translation tasks or different domains of language processing.

The error analysis revealed that the model made very few mistakes, and those that did occur were primarily associated with simpler numerals. This suggests that the model was better at handling complex structures but occasionally struggled with simpler cases. The slight bias observed towards complex numerals indicates that the dataset's composition plays a crucial role in shaping the model's learning. It emphasises the importance of balanced datasets in training models that generalise well across different levels of complexity. In future work, balancing the dataset more evenly or employing data augmentation techniques for simpler numerals could help in reducing this bias.

The model's inference speed of 35 samples per second is adequate for many real-time applications, demonstrating that the model is not only accurate but also efficient. The inference speed suggests that the model can be integrated into real-world applications where rapid translation is required, such as in educational tools or digital assistants for Yoruba speakers.

This study makes several contributions to the field of natural language processing, particularly in the context of low-resource languages like Yoruba. This include: successful application and fine-tuning of the Flan-T5 model for the task of numeral translation in a low-resource language, demonstrating its versatility and adaptability, and the creation of parallel dataset for further studies on translation.

Conclusion and Future Direction

This study successfully developed an efficient system for translating Yoruba numerals to English numerals by fine-tuning a pre-trained Flan-T5 model. The results demonstrate the model's effectiveness, with high accuracy

and low error rates, making it a valuable tool for educational and technological applications involving Yoruba numerals. The study highlights the potential of using advanced NLP models for low-resource languages, and also underscores the importance of dataset quality and model optimisation.

To enhance the system and expand its impact, focus should be on improving dataset quality with more data and high-quality annotations. This include the enhancement of the dataset with more Yoruba numerals, especially those with rare or complex structures. Additionally, the dataset was somewhat imbalanced, with fewer examples of simpler numerals. The dataset also only included one variation for each Yoruba numeral. In reality, Yoruba numerals can have multiple variations depending on context, spelling, or whether they are cardinal or ordinal. This limitation may affect the model's ability to correctly translate such numbers.

The model was trained and evaluated on standard Yoruba numerals, but Yoruba has several dialects with subtle differences in pronunciation and spelling. The model's performance might vary when exposed to numerals from different Yoruba dialects, potentially affecting its generalisation to a broader population.

Incorporating data from various Yoruba dialects could enhance the model's ability to generalise across different dialects and improve its robustness. The Flan-T5 model, while effective, is relatively complex and may require substantial computational resources for training and inference, which could be a barrier for deployment in resource-constrained environments.

References

Adetomiwa, A. (2023). The Yoruba numeral system and its computational applications. *Journal of African Linguistics*, 12(3), 45-60.

Agbeyangi, A. O., Eludiora, S. I., & Popoola, O. A. (2016). Web-based Yorùbá numeral translation system. IAES International Journal of Artificial Intelligence (IJ-AI), 5(4): 127-134. http://doi.org/10.11591/ijai.v5.i4.pp127-134

Akinade, A., & Ọdẹ́jọbí, T. (2014). Computational modeling of Yoruba numerals. *African Computational Linguistics Journal*, 9(1), 30-47.

Akinkurolere, S. O., & Akinfenwa, M. O. (2018). A study on the extinction of indigenous languages in Nigeria: Causes and Possible Solutions. *Annals of Language and Literature*, 2(1), 22-26.

Akinlade, M. T., Adedayo, A. M., & Dada, S. O. (2021). Globalization, foreign culture and impacts on Yoruba land cultural heritage and identity. *International Journal of Management, Social Sciences, Peace and Conflict Studies (IJMSSPCS)*, 4(2), 451-464.

Alkhateeb, F., Barhoush, M., & Al-Abdallah, E. (2016). Developing a system for converting a numeral text into a digit number: Abacus application. *Journal of Intelligent Systems*, 25(4), 611–628. https://doi.org/10.1515/jisys-2015-0029

Babarinde, O. (2014). Linguistic analysis of the structure of Yoruba numerals. *Language Matters*. 45(1), 127-147. https://www.tandfonline.com/doi/abs/10.1080/10228195.2013.857362

Elesemoyo, I., & Ọdẹ́jọbí, Ọ. (2022). Machine translation system for numeral in English text to Yorùbá language. *Ife Journal of Information and Communication Technology*, 6 (2022): 26–37. https://www.researchgate.net/publication/369384570_Machine_Translation_System_for_Numeral_in_English_Text_to_Yoruba_Language

Eludiora, S. I., & Abubakar, M. A. (2021). A Hindu-Arabic to Hausa number transcription system. *Malaysian Journal of Computing (MJoC)*, 6(1), 727-744. https://ir.uitm.edu.my/id/eprint/47827/1/47827.pdf

Ferraris, A. F., Audrito, D., Caro, L. D., & Poncibò, C. (2025). The architecture of language: Understanding the mechanics behind LLMs. *Cambridge Forum on AI: Law and Governance*, *1*, e11. https://doi.org/10.1017/cfl.2024.16

Javed, H., El-Sappagh, S., & Abuhmed, T. (2024). Robustness in deep learning models for medical diagnostics: Security and adversarial challenges towards robust AI applications. *Artificial Intelligence Review*, 58(1), 12. https://doi.org/10.1007/s10462-024-11005-9

Jimoh, E. R., Akinlolu, A. O., & Olojede, M. B. A. and O. A. (2024). An android-based Yorùbá numeral translation system. *FUOYE Journal of Pure and Applied Sciences (FJPAS)*, 8(3), 38-46. https://doi.org/10.55518/fjpas.PIIP2126

Keutayeva, A., & Abibullaev, B. (2024). Data constraints and performance optimization for transformer-based models in EEG-based brain-computer interfaces: A survey. *IEEE Access*, 62628–62647. https://doi.org/10.1109/ACCESS.2024.3394696

Mirela. (2024). Low-resource languages: A localization challenge. POEditor Blog. https://poeditor.com/blog/low-resource-languages/

Mittal, D., Pant, A., Jajoo, P., & Yadav, V. (2024). Transformer model applications: A comprehensive survey and analysis. In V. Goar, A. Sharma, J. Shin, & M. F. Mridha (Eds.), Deep learning and visual artificial intelligence, 25–38. *Springer Nature*. https://doi.org/10.1007/978-981-97-4533-3_3

Mustapha, R., Lawal, F. I., & Ahmad, M. A. (2023). Automated conversion of numeral to words in Hausa language. *Gadau Journal of Pure and Allied Sciences*, 2(2), 140-145. https://doi.org/10.54117/gjpas.v2i2.100

Nina, E. (2022, June 22). Understanding the impact of language on cross-border ecommerce. International Trade Council. https://tradecouncil.org/understanding-the-impact-of-language-on-cross-border-ecommerce/

Ninan, O. D., Iyanda, A. R., Elesemoyo, I. Olufemi., & Obasa, E. O. (2017). Computational analysis of Igbo numerals in a number-to-text conversion system. *Journal of Computer and Education Research,* 5(10), 241-254. https://doi.org/10.18009/jcer.325804

Esan, A., Oladosu, J., Oyeleye, C., Adeyanju, I., Olaniyan, O., Okomba, N., Omodunbi, B. & Adanigbo, O. (2020). Development of a recurrent neural network model for English to yorùbá machine translation. *International Journal of Advanced Computer Science and Applications*, 11(5), 602-609.

Olakanmi, O. (2015). Yoruba language and numerals' offline interpreter using morphological and template matching. *TELKOMNIKA Indonesian Journal of Electrical Engineering*, 13(1), 166-173. https://doi.org/10.11591/telkomnika.v13i1.6782

Olayinka, F. D., & Eludiora, S. (2019). Development of a machine translation system for date referencing in Yoruba language. *International Journal of Computer Applications*, 177(20), 32–38. https://doi.org/10.5120/ijca2019919646

Osuntokun, A. (2024, July 12). Yoruba language On Decline – THISDAYLIVE. https://www.thisdaylive.com/index.php/2024/07/12/yoruba-language-on-decline/

Orife, I., Adelani, D. I., Fasubaa, T., Williamson, V., Oyewusi, W. F., Wahab, O., & Tubosun, K. (2020). Improving Yorùbá Diacritic Restoration. arXiv 2020. arXiv preprint arXiv:2003.10564.

Tang, W., & Li, G. (2023). Enhancing Competitiveness in Cross-Border E-Commerce Through Knowledge-Based Consumer Perception Theory: An Exploration of Translation Ability. *Journal of the Knowledge Economy*. https://doi.org/10.1007/s13132-023-01673-3

Tsai, M.-H. M., & Yao, Y.-C. (2011). Dealing with numbers from English to Chinese in consecutive interpretation. *In Proceeding of the Fifth International Conference on Translation, Interculturality and Translation: Less-Translated Languages,* Barcelona, Spain.

Xiao et al., 2025 Xiao, Y., Zuo, X., Lu, X., Dong, J. S., Cao, X., & Beschastnikh, I. (2025). Promises and perils of using transformer-based models for SE research. *Neural Networks*, 184, 107067. https://doi.org/10.1016/j.neunet.2024.107067