



## Research Article

Volume-04|Issue-03|2024

# Trans Lingua: Unlocking Language Barriers

Palvisha Changkakati\*<sup>1</sup>, Aashish Ramana S<sup>2</sup>, Aditi Umesh Nayak<sup>3</sup>, Pranshu Sinha<sup>4</sup>, Roopashree S<sup>5</sup><sup>1,2,4,3,5</sup>Department of Computer Science and Engineering, RV Institute of Technology and Management, Bengaluru, Karnataka, India.

### Article History

Received: 20.05.2024

Accepted: 05.06.2024

Published: 30.06.2024

### Citation

Changkakati, P., Ramana, A. S., Nayak, A. U., Sinha, P., & Roopashree, S. (2024). Trans Lingua: Unlocking Language Barriers. *Indiana Journal of Multidisciplinary Research*, 4(3), 31-35.

**Abstract:** This paper presents Trans Lingua, a mobile app designed to tackle cross-lingual communication challenges. It focuses on translating and transliterating English text into various Indian languages, using advanced language models along with Speech Engines. The project prioritizes user-friendliness, inclusivity, and ethical considerations, aiming to benefit individuals with diverse linguistic backgrounds and impairments. Key findings highlight the system's effectiveness in precise translations and maintaining language subtleties, making Trans Lingua a valuable tool with broad applications.

**Keywords:** Cross-lingual Communication, Language Liberation Machine (LLM), Translation, Transliteration, Indian Languages

**Copyright** © 2024 The Author(s): This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY-NC 4.0).

## INTRODUCTION

In today's interconnected global landscape, effective communication across diverse languages is essential. Language barriers often restrict individuals' access to vital information and services, creating a need for innovative solutions. To meet this challenge, we present "Trans Lingua," a mobile application poised to dismantle language barriers and enable seamless cross-lingual communication.

The Trans Lingua project utilizes advanced technologies and models to empower users to effortlessly translate and synthesize text across various Indian languages. Designed to accommodate users with different language preferences and accessibility needs, the project offers versatility by accepting content in various formats such as .docx and .pdf files. At its core are specialized models for language translation and transliteration tailored for Indian languages, seamlessly extracting English text from documents and rendering it into a wide range of Indian languages.

Beyond traditional written translations, the application features an innovative text-to-speech function. This functionality brings translated text to life, delivering it in clear and natural-sounding speech. Beneficial for individuals with visual impairments or those preferring auditory content, this feature is made possible through integration with a text-to-speech engine. Additionally, the system incorporates an avatar video generation mechanism, enhancing the interactivity of knowledge transfer.

The project integrates elements from Natural Language Processing (NLP), Machine Learning, and Human-Computer Interaction (HCI), addressing the complexities of multilingual communication to ensure content accessibility for a broader audience, regardless of linguistic diversity. At its essence, the Trans Lingua application aims to bridge language gaps, promote inclusivity, and transform how individuals engage with textual content. It represents a fusion of AI models and innovative user interface design, offering a comprehensive solution for multilingual communication and accessibility.

## MATERIALS AND METHODS

"Adaptive Machine Translation with Large Language Models" (arXiv preprint arXiv:2301.13294v3, 2023) explores using Large Language Models (LLMs) for machine translation tasks. The paper highlights LLMs' potential to improve translation quality for various languages, even those with limited resources, compared to traditional encoder-decoder models. However, challenges like handling fuzzy or out-of-vocabulary terms and maintaining consistent quality for low-resource languages require further attention.

"INDICXNLI: Evaluating Multilingual Inference for Indian Languages" (arXiv preprint arXiv:2204.08776v1, 2022) introduces a dataset called INDICXNLI to assess how well machines can understand relationships between sentences in 11 Indian languages. The dataset leverages high-quality

translations from English and involves human verification to ensure accuracy. However, challenges include mitigating unnatural phrasing caused by translation and addressing performance gaps between languages with varying resources. The paper also explores using pre-trained models for this task, highlighting the limitations of smaller models compared to larger ones.

The paper "Machine Translation Approaches and Survey for Indian Languages" published in the *Computational Linguistics and Chinese Language Processing* Vol. 18, 2020 evaluates the performance of Statistical Machine Translation (SMT) for eight Indian languages. It finds that current SMT systems achieve low translation quality (average BLEU score of 10%) due to limited data and challenges with vocabulary size and data sparsity. The authors emphasize the need for more extensive linguistic resources and parallel corpora to improve machine translation for Indian languages.

The research paper "Transliteration for Indian Languages: A Literature Survey" published in the *International Journal of Scientific & Engineering Research* examines various techniques for machine transliteration, focusing on Indian languages. The paper explores transliteration systems for converting English text into Hindi, Tamil, Telugu, Kannada, and Malayalam, highlighting that most research centres on Hindi and Dravidian languages. It reveals that existing models primarily rely on graphemes (written units) rather than phonetics (sounds) and achieve accuracy levels ranging from 70.5% (English-Hindi) to 81.25% (English-Kannada), often using limited training datasets.

The paper "MATra: A Multilingual Attentive Transliteration System for Indian Scripts", published in 2022 (arXiv.2208.10801), presents a new system for transliteration between five Indian languages (English, Hindi, Bengali, Kannada, and Tamil). It leverages a powerful transformer-based architecture to capture complex relationships between languages, achieving high accuracy (80.7%) and strong performance in preserving phonetics, which is essential for transliteration.

The research paper "Towards Building Text-To-Speech Systems for the Next Billion Users" explores challenges in developing Text-To-Speech (TTS) systems for 13 Indian languages (2023, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)). It investigates various factors like suitable models, training approaches, and data considerations to achieve high-quality speech synthesis. The research identifies FastPitch and HiFi-GAN V1 models as optimal for monolingual TTS, emphasizing the open-source availability of the

models to improve TTS accessibility for Indian languages.

This survey paper "A Survey on Neural Speech Synthesis" (arXiv preprint arXiv:2106.15561, 2021), examines the field of neural speech synthesis, covering various aspects from text analysis to advanced techniques for different synthesis goals. It acknowledges the challenge of this broad scope, requiring expertise in linguistics, acoustics, signal processing, and machine learning. The paper provides a comprehensive overview of neural speech synthesis, making it a valuable resource for researchers in this field.

"Text Processing for Text-to-Speech Systems in Indian Languages" (Ssw, pp. 188-193, 2020) proposes methods for improving text processing in Text-to-Speech (TTS) systems for Indian languages. It focuses on handling challenges like font variations, pronunciation rules, and text normalization to ensure accurate speech generation. The paper emphasizes the need for a generic framework that can be efficiently adapted to various Indian languages despite limited linguistic resources, especially for minority languages.

The paper "JARVIS: An interpretation of AIML with the integration of gTTS and Python": This paper (2019 International Conference on Intelligent Computing, Instrumentation and Control Technologies) introduces JARVIS, a virtual assistant built using Python, AIML (a language for chatbots), and gTTS (a text-to-speech library). JARVIS leverages Python's pytsx3 library to integrate these technologies, enabling voice interaction. However, limitations include gTTS's data input restrictions and the system's inability to grasp the broader context of conversations.

"Video Transcription and Summarization using NLP" (Proceedings of the Advancement in Electronics & Communication Engineering, 2022), paper proposes a video transcription and summarization system built using NLP techniques (Natural Language Processing). The system leverages the Moviepy library to process videos and relies on a pre-trained model hub for tasks like video resampling and noise reduction, facilitating efficient video analysis. However, the paper doesn't address how the system personalizes its output based on user preferences or specific needs.

The survey paper "Survey on Face Detection Algorithms" (2021 International Journal of Innovative Science and Research Technology Volume 6, Issue 2) examines different algorithms used for face detection, comparing their strengths and weaknesses. It analyzes popular options like Cascade Classifiers, OpenCV with DNN (Deep Neural

Network), and Dlib's Histogram of Oriented Gradients (HoG) method. The paper reveals limitations such as false positives, struggles with non-frontal faces, and the inability to detect small faces, urging further development for robust face detection.

The research paper "Pre-Avatar: An Automatic Presentation Generation Framework Leveraging Talking Avatar" (arXiv preprint arXiv:2210.06877, 2022) introduces Pre-Avatar, a system that automatically generates presentations using a talking avatar. Pre-Avatar combines a user interface, a text-to-speech module, and a facial animation module to create presentations from text and a user-provided photo. The paper acknowledges that the quality of the output, both speech and video, is dependent on the quality of the input materials, such as potential noise or resolution issues in the photo and audio.

## METHODOLOGY

### User Authentication:

- Users begin by either signing up or signing in to the mobile application using their credentials.
- The application validates their credentials to ensure they are authorized to access these services.

### Service Selection:

*Once authenticated, users can select from a variety of services:*

- Text translation (from English to Indic languages)

- Text transliteration (from English to Indic languages)
- Audio file generation from text (text-to-speech)
- Video creation (involving a visual representation for the user)

### Service Processing:

#### Text Translation

- Users input English text (.docx or .txt).
- The translated text is outputted in a supported Indic language format (.docx or .txt).

#### Text Transliteration:

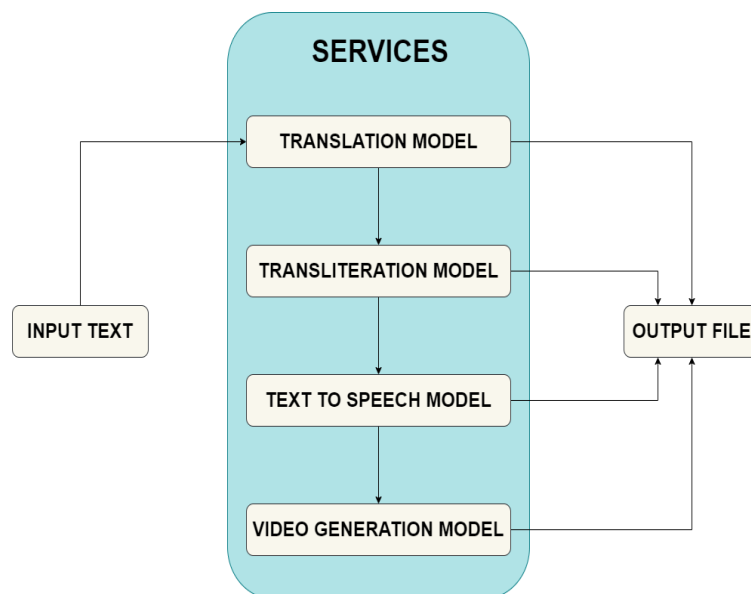
- The sentences in English (like slokas) that need to be transliterated are taken as input.
- The transliterated text is outputted in English format combined with the translated text (.docx or .txt).

#### Audio and Video File Generation:

- The translated and transliterated file is taken as input.
- The application utilizes a text to speech engine to generate an audio file (.wav or .mp3) and video (.mp4).

### Key Components:

- Admin Database: This likely stores user information, credentials, and potentially application settings.
- User Mobile Application: The interface through which users interact with the services.



**Figure 1:** Working of Trans Lingua

## RESULTS

The project's functionality heavily relies on translation and transliteration models, crucial components identified through literature review for their adeptness in handling multiple Indian languages. The recent release of the Bharat Parallel Corpus

Collection (BPCC) further enriches the project by furnishing extensive training data, benchmarks, and accessible models. Demonstrating superior performance on various benchmarks, these models underpin the project's success.

The addition of a transformative text-to-speech feature, particularly beneficial for users with visual impairments or those preferring auditory content, enhances accessibility. Integrated with the text-to-speech engine, this feature ensures clear and natural-sounding speech, enriching user experience. The inclusion of an avatar video generation mechanism further elevates the interactive aspect of knowledge transfer, catering to diverse audience preferences for consuming information.

A multidisciplinary approach, amalgamating elements from Natural Language Processing (NLP), Machine Learning, and Human-Computer Interaction (HCI), serves as a key strength. Addressing complex challenges associated with multilingual communication, the application ensures content accessibility for a diverse audience. Collaborative efforts, accessibility, and the release of models and data with permissive licenses for further research and

development are emphasized in the literature survey, supporting the project's approach.

Innovative features such as a user-friendly interface, audio capabilities, and avatar videos distinguish the project from existing studies focused primarily on translation quality, model comparisons, and dataset creation. Acknowledging challenges like data scarcity, especially for low-resource languages, aligns with limitations outlined in related papers.

To comprehensively evaluate the performance of the system, incorporating a comparison based on BLEU scores is essential. BLEU scores, a widely used metric for assessing machine-generated translation quality, provide a quantitative measure of translation accuracy by measuring similarity between the generated translation and one or more reference translations.

**Table 1:** BLEU Score of the proposed model in comparison to GPT 3.5 and Google Translate engines

Language	Google Translate	GPT 3.5	Proposed Model
Hindi	0.682	0.302	0.713
Bengali	0.107	0.102	0.318
Marathi	0.460	0.157	0.454
Gujarati	0.529	0.196	0.619
Kannada	0.769	0.129	0.765
Telugu	0.532	0.470	0.532
Tamil	0.469	0.420	0.469
Malayalam	0.321	0.311	0.320

## CONCLUSION

This project represents a significant step towards bridging linguistic gaps by integrating an advanced language translation module with a transliteration system. This integration enhances linguistic accessibility and inclusivity within India's diverse linguistic landscape. Users can now easily communicate in their preferred Indian languages, effectively overcoming traditional language barriers. Utilizing AI in Translation and Transliteration ensures not only linguistic accuracy but also cultural relevance, fostering genuine cross-lingual engagement. By incorporating automated audio and video generation modules, the project expands its accessibility features to cater to a broader audience, including individuals with sensory impairments and language learners.

This multi-modal approach to communication supports both text and auditory and visual content, increasing the technology's usability and applicability across various contexts. Looking ahead, the project aims to enhance AI algorithms to improve translation and transliteration quality, broaden language coverage, and integrate with larger platforms to reach more users. Future efforts will also concentrate on refining user feedback mechanisms to ensure adaptability and user-centeredness. This initiative strives to transform

language barriers into bridges, fostering a more inclusive and connected global digital ecosystem.

## ACKNOWLEDGEMENTS

We would like to extend our heartfelt gratitude to our Institution for providing the necessary resources, facilities, and support that enabled the successful completion of this research paper. We are also deeply indebted to our esteemed guide, for their invaluable guidance, mentorship, and encouragement throughout this project. Their expertise, insights, and unwavering support have been instrumental in shaping the direction and outcomes of this research.

## REFERENCES

- Aggarwal, D., Gupta, V., & Kunchukuttan, A. (2022, April 19). IndicXNLI: Evaluating Multilingual Inference for Indian Languages. *arXiv.org*. <https://arxiv.org/abs/2204.08776> (accessed January 5, 2024).
- Gibadullin, R. F., Perukhin, M. Y., & Ilin, A. V. (2021, May 1). Speech Recognition and Machine Translation Using Neural Networks. *IEEE Xplore*. <https://ieeexplore.ieee.org/abstract/document/9446474>
- Kandimalla, A., Lohar, P., Maji, S. K., & Way, A. (2022). Improving English-to-Indian Language

- Neural Machine Translation Systems. *Information*, 13(5), 245. <https://doi.org/10.3390/info13050245>
4. Makadiya, J., & Dave, J. (2024). NATURAL LANGUAGE PROCESSING: MACHINE TRANSLATION FOR INDIAN LANGUAGE. *Peer-Reviewed, Open Access*. Accessed: January 5, 2024.
  5. Moslem, Y., Haque, R., Kelleher, J. D., & Way, A. (2023, May 9). Adaptive Machine Translation with Large Language Models. *arXiv.org*. <https://arxiv.org/abs/2301.13294>
  6. Mujadia, V., *et al.* (2023, November). Assessing Translation Capabilities of Large Language Models involving English and Indian Languages. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2311.09216>
  7. Porwal, K., Srivastava, H., Gupta, R., Pratap Mall, S., & Gupta, N. (2022). Video Transcription and Summarization using NLP. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4157647>
  8. Raj, Y., & Laddagiri, B. (2022, August 23). MATra: A Multilingual Attentive Transliteration System for Indian Scripts. *arXiv.org*. <https://arxiv.org/abs/2208.10801> (accessed January 5, 2024).
  9. Raj, A. A., Sarkar, T., Pammi, S. C., Yuvaraj, S., Bansal, M., Prahallad, K., & Black, A. W. (2007). Text Processing for Text-to-Speech Systems in Indian Languages. *Ssw*, 188-193.
  10. Soman, D. (2011). Machine Transliteration for Indian Languages: A Literature Survey. *International Journal of Scientific & Engineering Research*, 2. Accessed: January 5, 2024.
  11. Tan, X., Qin, T., Soong, F., & Liu, T.-Y. (2021, July). A Survey on Neural Speech Synthesis. *arXiv:2106.15561 [cs, eess]*. Available: <https://arxiv.org/abs/2106.15561>
  12. Towards Building Text-to-Speech Systems for the Next Billion Users | IEEE Conference Publication | IEEE Xplore. *IEEE Xplore*. <https://ieeexplore.ieee.org/abstract/document/10096069>
  13. Pre-Avatar: An Automatic Presentation Generation Framework Leveraging Talking Avatar | IEEE Conference Publication | IEEE Xplore. *IEEE Xplore*. <https://ieeexplore.ieee.org/abstract/document/10097978> (accessed January 5, 2024).
  14. JARVIS: An interpretation of AIML with the integration of gTTS and Python | IEEE Conference Publication | IEEE Xplore. *IEEE Xplore*. <https://ieeexplore.ieee.org/abstract/document/8993344> (accessed January 5, 2024).