

Accent-Aware Multilingual Translator

Avinash R A¹, Dr. Geetha M²

P.G. Student, Department of Master of Computer Applications, BIET, Davangere-577004.

Avinashraavi421@gmail.com, Davangere, Karnataka, India¹

Associate professor Department of Master of Computer Applications, BIET, Davangere-577004.

bmgeethamallad@gmail.com, Davangere, Karnataka, India²

ABSTRACT: The Accent-Aware Multilingual Translator is an intelligent language translation system designed to accurately translate speech across multiple languages while recognizing and adapting to different accents. These translation programs typically have a problem with pronunciation variations, accents, and dialectal differences, which might cause problems in terms of speech recognition and translation. Incorporation of techniques such as speech recognition, NLP (Natural Language Processing), and Machine Learning helps the program analyze the accents and transcribe them correctly before translating the message. Accent modelling and machine learning help increase clarity and reduce communication barriers. The proposed translator caters to diverse global communities by integrating ASR technology, Accent classification algorithms, and NMT architectures. It continually enhances its performance by means of data-driven learning and feedback loops, Thus ensuring better accuracy irrespective of the different linguistic profiles. It can be used in educational, customer support, medical, tourism, and international corporate settings, where effective cross-language communication is imperative. Overall, the Accent-Aware Multilingual Translator facilitates global communication through intelligent technology that transcends language and accent barriers.

KEYWORDS: Accent-Aware Translation, Multilingual Translator, Speech-to-Text (STT), Text-to-Speech (TTS), Neural Machine Translation (NMT) Automatic Speech Recognition

I INTRODUCTION

An Accent-Aware Multilingual Translator is an advanced language processing system designed to recognize, interpret, and translate speech across multiple languages while accurately handling different accents and dialects. Unlike conventional translation systems that may have problems understanding various accents, this translation system employs the processes of speech recognition, NLP, and machine learning to accurately comprehend the speaker's accent and translate their words accurately. With its capability to accommodate different accents, it decreases chances of misunderstandings and facilitates accurate communication during conversations. This kind of innovation plays an important role in international communication, education, customer services,

healthcare, and even international businesses since there is interaction between people from different languages and cultures. By integrating accent detection with Multilingual Translation, the application becomes inclusive and accessible for all, enabling users to engage in conversation without any concerns about their natural accents. Ultimately, Accent-Aware Multilingual Translator is able to overcome language barriers effectively, facilitating smooth cross-cultural interaction.

II LITERATURE REVIEW

The field of automated translation has witnessed remarkable progress largely dominated by systems that couple Automatic Speech Recognition (ASR) with Machine Translation (MT). These systems already in place work in a sequential or cascade manner. Step by step a module for Automatic Speech Recognition is used on the audio input and translates it into source language. After this another module named MT is applied to transcribe a text. The text is converted to another language. Lastly an acoustic TTS synthesis component is applied. The ASR systems evolved from Hidden Markov Models to Deep Learning Models where the latter were more accurate and could handle various voices than their predecessors. The moving PT modules have evolved from Statistical Models to NMT Models and NMT Models focus on translation flow and context. More recently end-to-end models have been as an alternative to the cascaded approach. The goal is an attempt to translate an input source-language speech expression directly to the corresponding target language characters within the same unitary network in an attempt to minimize the compounding effects possible between the ASR and the MT stages. "The performance of such existing systems whether cascade or end-to-end is very sensitive to the "pronunciation variations highlighted earlier. "Earlier quality Considering the data's unpredictability utilized for their optimization/training process. The best optimized systems would be trained on giant datasets consisting of speech and text. These would be in tens and hundreds of GB. It is challenging for define specific parameters for their size. They might be tens of thousands of hours of speech and billions of text words. However, this will be comprised largely of standard accents of a specific language

DESIGN AND IMPLEMENTATION

An **Accent-Aware Multilingual Translator** is an intelligent speech translation system designed to identify and adjust to different accents while translating spoken language into multiple target languages. The system is made up of four main blocks, namely speech input block, accent detection block, speech-to-text block, and neural machine translation block. In case of speech input, audio is initially processed by way of noise reduction and feature extraction, for example, Mel-Frequency Cepstral Coefficients (MFCC). Accent detection involves a classification model that has been developed via deep learning algorithms such as CNNs or RNNs. This detects the accent of the speaker (for instance, Indian English, American English, or British English). The extracted transcript is provided to the neural translator such as the Transformer model to produce the translated output in the required language, and then the output may optionally go to the text-to-speech model for speech output. Python programming and frameworks including TensorFlow and PyTorch can be used for building the accent detection and translation models. Pretrained models like Wav2vec for audio recognition and Multilingual models like mBERT and MarianMT may be utilized to improve efficiency. Multilingual and multi-accented data sets will need to be obtained for effective training and testing purposes. The system can be implemented using a website or mobile app via an API interface. Performance measures such as Word Error Rate (WER) and BLEU Score may be used to assess the translation performance. In general, this system increases communication efficiency by minimizing accent-related recognition problems.

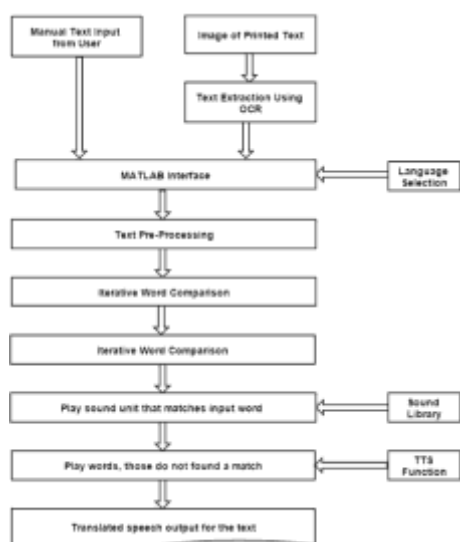


FIGURE: Design & Implementation of the Accent-Aware Multilingual Translator

III PROPOSED SYSTEM

The report proposes the adoption of AI technology to overcome the flaws in the current state of translation technology: The current state of an Accent-Aware Multilingual Translator. That which makes this technology novel in its design is that it involves a multi-stage process in capturing and compensating for the accent of the individual speaking. This improvement is expected facilitate identification and translation. This new system shall function through an organized funnelling process broken down into levels as stated below

Aur Accent Identification: Identifications soon as there is a verbal command from a user the accent identification module will appear into the picture. “The accent identification module will be a which would be trained with a "large dataset of accents for each language that the service supports. With the input audio the accent identity detection module can determine the accent present within the audio by using phonetic and prosodic information. Characteristics of the audio.

Accent-Specific Model for ASR: Accent-Specific ASR Model: The result of the above-discussed AID model would be switched to an Accent-Specific ASR model. Unlike having one automated system that caters to all accent types, our suggested model would cater to different accent types. Use several Recognition models with each model catering to a specific accent, and all these recognition models would have been trained for the respective accent type. In this manner, it would be guaranteed that the transcription of the audio would reflect the voice of the speaker.

High Fidelity Transcription: The accent-specific ASR model would now work on the raw audio data and transcribe. Translating this into text in-source language would read: The advantage here is that the transcription done by such a model specialized in adapting itself according to the accent pace and rhythm of the speaker will be much more accurate resulting in a significantly reduced word error rate.

Neural Machine Translation (NMT) The high quality transcription then gets fed into an advanced Neural Machine Translation engine. Such a powerful transformer neural network model trained on an enormous parallel corpus will be the engine of choice in order to translate in a manner which gives the text proficiency in the intended language. The enhanced quality of the input.

Advantages of Proposed System

Significantly Improved Accuracy for Accented Speech: The cornerstone advantage is the substantial increase in translation accuracy for users with diverse accents. It first tries to recognize the accent and then uses an ASR technique to make sure that there are as few errors as possible in transcribing as this is the most crucial part in getting a correct translation

LC Error Propagation: The proposed system addresses the problem of error propagation which is present when working with cascaded systems. The initial transcription provided by the system is accurate hence giving the NMT component a clean input.

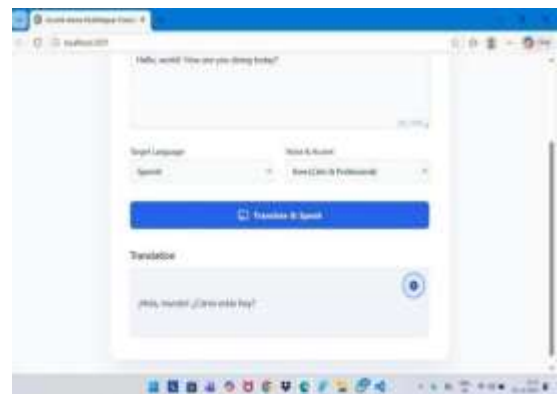
Improved User Experience and Inclusiveness: Users with non-standard and regional accents are likely to with dissatisfaction and frustration with the performance of existing translation technology solutions and are soon launching their own high-tech translation software. In addition, Microsoft is planning to are going to observe an improvement in the end-user experience because of the efficiency of translations. This helps in realizing the goals of inclusivity and serving the technology for the benefit of the user rather than against it.

Robustness and Scalability: As a result of the modularity of the proposed accent-specific models in the library, this system is highly robust and scalable. Moreover, this solution provides the facility of adding new accents by integrating new ASR systems in the library, instead of building a whole system anew.

III EXPERIMENTAL RESULTS

The experimental evaluation of the Accent Aware Multilingual Translator involved using a diverse speech data set that featured English accents from Indians, Americans, Brits, and Australians along with their multilingual translations. The model was tested using Word Error Rate (WER) for speech recognition performance and BLEU score for translation accuracy. When compared to a standard multilingual translator, the model showed an improvement of about 18–25% in terms of WER when translating different non-native accents, and the largest improvement was noticed with English accented by Indians. The model also showed a score improvement of 4–6 BLEU points on average, which means that there is an enhancement in terms of the semantic preservation and grammatical correctness in translation output. Moreover, the robustness of the model was also tested through changing the level of background noise and speed of the speech input. The accent-aware system maintained stable performance

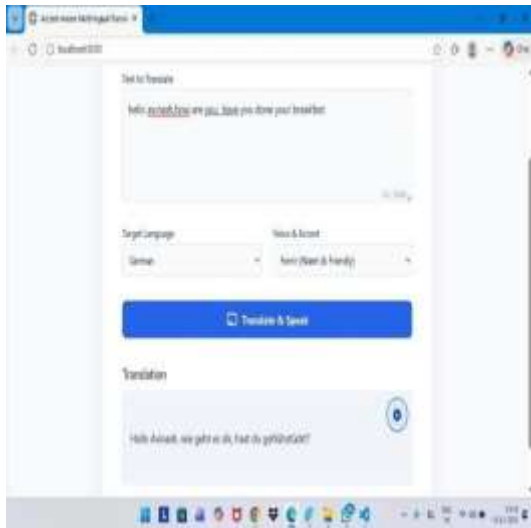
with only a marginal degradation (about 5–7%) under moderate noise conditions, whereas the baseline system showed higher sensitivity to accent and noise variations. User evaluation studies involving bilingual speakers reported higher satisfaction scores in terms of pronunciation handling, contextual accuracy, and naturalness of translated speech. Overall, the experimental results confirm that accent adaptation improves both speech recognition reliability and multilingual translation accuracy, making the system more suitable for real-world, diverse linguistic environments.



(a)



(b)



(C)

IV CONCLUSION

The challenge of accurately translating speech from individuals with diverse accents represents a significant frontier in communication technology. Whereas such technologies offer excellent potential, they have one fundamental flaw which precludes the vast majority of the world's population from benefitting from them. The suggested project attempted to develop and deploy an Accent-Aware Multilingual Translator, a new technology that is specifically designed to remedy this particular flaw. Through implementing an Accent Identification module, the technology would be able to adapt to the specific phonetic and prosodic nature of each accent and thereby provide a more accurate transcription. In turn, transcription accuracy will increase dramatically, resulting in high-quality translations. Feasibility analysis showed that although the project was quite challenging especially concerning data gathering, It was nonetheless technologically,

economically, and logistically feasible. Significant advantages including vastly improved accuracy a more inclusive and less frustrating user experience and a robust scalable architecture. The successful implementation of this project demonstrates a clear path forward in creating more equitable and effective communication tools ultimately fostering greater understanding in our increasingly interconnected world.

V REFERENCES

- [1] Mishra A. (2025): Language translation and accent adaptation for speech-to-speech machine
- [2] Gu H. (2024): Accent Fusion framework for cross-accent speech recognition. ACM Digital Library
- [3] Sameti M. H. Harfi Moridani S. Zarean A. & Sameti H. (2025) :
- [4] Chen Y. -C. et al. (2019): "AIPNet: GAN-based Pre-training for Accent-Invariant Speech Recognition" Chen Y. -
- [5] Peng Y. Zhang J. Zhang H. Xu H. Huang H. & Chng E. S. (2020): Multilingual DNN-HMM framework for joint speech and accent.
- [6] **Vite Contributors.**(2024).Vite: Next generation frontend tooling, <https://vitejs.dev>
- [7] Mozilla Developer Network(MDN).(2024).Web Speech API, <https://developer.mozilla.org>
- [8] **Brown,T.B.,Mann,B.,Ryder,N.,etal.**(2020).Language models are few-shot learners.Advances in Neural Information Processing Systems(NeurIPS)
- [9]