

# **A Systematic Survey of Multilingual Speech Transcription and Translation**

**Vaibhav Ravindra, Dheeraj K N, Prof. Vinay Raj**

Department of Information Science and Engineering  
Global Academy of Technology, Bangalore, Karnataka, India

**Abstract:** *This paper presents an innovative initiative dedicated to revolutionizing multilingual communication by leveraging state-of-the-art technologies such as Artificial Intelligence (AI), NLP, or natural language processing and Machine Learning. With a primary focus on Indian languages, the research aims to develop an advanced system capable of seamlessly transcribing speech across diverse linguistic landscapes. Through the incorporation of cutting-edge algorithms and parallel processing techniques, the proposed system facilitates real-time transcription and translation of multiple languages concurrently. Rigorous experimentation and analysis demonstrate the efficiency of the developed framework in breaking down language barriers and fostering inclusive communication. Furthermore, the paper emphasizes the cultural significance of this technology in promoting global connectivity and celebrating linguistic diversity. Ultimately, this research underscores the transformative potential of technology in facilitating cross-cultural understanding and enabling meaningful interactions within a multilingual society*

**Keywords:** Natural Language Processing or NLP, Speech-to-text translation, Linguistic analysis, Unsupervised learning, Hidden Markov Models (HMM), Speech-to-speech translation.

## **REFERENCES**

- [1] Chung, Y. A., Weng, W. H., Tong, S., & Glass, J. (2019, May). Towards unsupervised speech-to-text translation. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 7170-7174). IEEE.
- [2] Elakkiya, A., Surya, K. J., Venkatesh, K., & Aakash, S. (2022, December). Implementation of Speech to Text Conversion Using Hidden Markov Model. In 2022 6th International Conference on Electronics, Communication and Aerospace Technology (pp. 359-363). IEEE.
- [3] Sung, T. W., Liu, J. Y., Lee, H. Y., & Lee, L. S. (2019, May). Towards end-to-end speech-to-text translation with two-pass decoding. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 7175-7179). IEEE.
- [4] Shi, J., Tang, Y., Lee, A., Inaguma, H., Wang, C., Pino, J., & Watanabe, S. (2023, June). Enhancing Speech-To-Speech Translation with Multiple TTS Targets. In ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1-5). IEEE.
- [5] Nikolaeva, D. (2023, September). An Elementary Emulator Based on Speech-To-Text and Text-to-Speech Technologies for Educational Purposes. In 2023 XXXII International Scientific Conference Electronics (ET) (pp. 1-6). IEEE.
- [6] Indurthi, S., Han, H., Lakumarapu, N. K., Lee, B., Chung, I., Kim, S., & Kim, C. (2020, May). End-end speech-to-text translation with modality agnostic meta-learning. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 7904-7908). IEEE.
- [7] Omar, M., Choi, S., Nyang, D., & Mohaisen, D. (2022). Robust natural language processing: Recent advances, challenges, and future directions. IEEE Access.
- [8] Sunar, A. S., & Khalid, M. S. (2023). Natural Language Processing of Student's Feedback to Instructors: A Systematic Review. IEEE Transactions on Learning Technologies.

- [9] Meyer, S., Tilli, P., Denisov, P., Lux, F., Koch, J., & Vu, N. T. (2023, January). Anonymizing Speech with Generative Adversarial Networks to Preserve Speaker Privacy. In 2022 IEEE Spoken Language Technology Workshop (SLT) (pp. 912-919). IEEE.
- [10] Lakdawala, B., Khan, F., Khan, A., Tomar, Y., Gupta, R., & Shaikh, A. (2018, April). Voice to Text transcription using CMU Sphinx A mobile application for healthcare organization. In 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT) (pp. 749-753). IEEE.
- [11] Reddy, V. M., Vaishnavi, T., & Kumar, K. P. (2023, July). Speech-to-Text and Text-to-Speech Recognition Using Deep Learning. In 2023 2nd International Conference on Edge Computing and Applications (ICECAA) (pp. 657-666). IEEE.
- [12] V. Advait, A. Shivkumar and B. S. Sowmya Lakshmi, "Parts of Speech Tagging for Kannada and Hindi Languages using ML and DL models," 2022 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, 2022, pp. 1-5, doi:10.1109/CONECCT55679.2022.9865745. keywords: {Deep learning;Machine learning algorithms;Computationalmodeling;Tagging;Linguistics;Natural language processing;Communicationstechnology;Natural Language Processing;MachineLearning;DeepLearning;Part of Speech
- [13] R. V. Malage, H. Ashish, S. Hukkeri, E. Kavya and R. Jayashree, "Low Resource Speech-to-Speech Translation of English videos to Kannada with Lip-Synchronization," 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2023, pp. 1680-1687, doi: 10.1109/ICICCS56967.2023.10142578.
- [14] S. Geeta and B. L. Muralidhara, "Syllable as the basic unit for Kannada speech synthesis," 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, India, 2017, pp. 1205-1208, doi: 10.1109/WiSPNET.2017.8299954.
- [15] S. Hebbar, A. R. B, M. S. N, M. Supriya, N. V. G and S. L, "Named Entity Recognition Using BERT Model for Kannada Language," 2023 International Conference on Recent Advances in Information Technology for Sustainable Development (ICRAIS), Manipal, India, 2023, pp. 212-216, doi: 10.1109/ICRAIS59684.2023.10367119.
- [16] S. A. A T, S. Shankaran, H. M. Thrupthi and M. H R, "Natural Language Processing based Cross Lingual Summarization," 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2022, pp. 1825-1829, doi: 10.1109/ICOEI53556.2022.9776655.
- [17] Danyang Liu, Ji Xu, Pengyuan Zhang, YonghongYan, A unified system for multilingual speech recognition and language identification, *Speech Communication*, Volume 127, 2021, Pages 17-28, ISSN 0167-6393, <https://doi.org/10.1016/j.specom.2020.12.008>.
- [18] Islam, Sharmi, Mustahid Hasan, and Md Ismail Jabiullah. "An Implementation of Advanced NLP for High-Quality Text-To-Speech Synthesis." *Advancement of Computer Technology and its Applications* 4.2, 3 (2022): 19-30.
- [19] Rustam Shadiev, Yueh-Min Huang Facilitating cross-cultural understanding with learning activities supported by speech-to-text recognition and computer-aided translation, *Computers&Education*, Volume 98, 2016, Pages 130-141, ISSN 0360-1315, <https://doi.org/10.1016/j.compedu.2016.03.013>.
- [20] N Shikha, R Pranav, Nidhi R Singh, V Umadevi and Muzammil Hussain Conference: 2023 10th International Conference on Signal Processing and Integrated Networks (SPIN), Year: 2023, Page 379, DOI: 10.1109/SPIN57001.2023.10117096