

Mozhi-Meyporul: A Real-Time Multimodal Architecture for the Interpretation of the Tirunelveli Dialect.

Raj Narayanan V

Dept. of Computer Science and Engineering,

Francis Xavier Engineering College - Tirunelveli,

Tamil Nadu - India.

rajnarayanan.ug.23.cs@francisxavier.ac.in

Mrs. G. Durgaveni,

Assistant Professor / Dept. of Computer Science and Engineering,

Francis Xavier Engineering College - Tirunelveli,

Tamil Nadu - India.

durgavenig@francisxavier.ac.in

Mr. G. Rahul Vignesh,

Assistant Professor / Dept. of Computer Science and Engineering,

Francis Xavier Engineering College - Tirunelveli,

Tamil Nadu - India.

rahulvigneshg@francisxavier.ac.in

Abstract:

The rapid globalization of digital communication often leaves regional dialects—such as the Tirunelveli (Nellai) dialect of Tamil—underrepresented in standard machine translation models. This paper presents Mozhi-Meyporul, an intelligent real-time interpreter designed to bridge this linguistic gap. The platform integrates Generative AI (Gemini 2.5 Flash) and Speech Recognition to facilitate high-fidelity, real-time translation and slang analysis. Unlike generic translation tools, Mozhi-Meyporul utilizes a multimodal architecture to capture spoken audio and provide dynamic feedback loops, offering users not just a literal translation but also an explanation of cultural nuances and local slang. Preliminary testing indicates that the system significantly improves cross-regional communication and serves as a scalable solution for linguistic research and real-time social interaction.

Keyword: Real-Time Interpretation, Dialect Preservation, Generative AI, Natural Language Processing (NLP), Multimodal Interaction.

Introduction:

The contemporary digital landscape is undergoing a radical transformation driven by the integration of artificial intelligence in linguistics. While traditional translation services excel at formal languages, they often fail to capture the behavioral nuances and regional variations inherent in localized dialects. For individuals interacting outside their native linguistic zones, the transition can be fraught with technical and social challenges.

Mozhi-Meyporul is introduced as a comprehensive, AI-driven solution designed to democratize high-quality linguistic interpretation. By leveraging Generative AI and Real-Time Speech-to-Text (STT), the platform provides a realistic "linguistic sandbox" for users. This research explores the development of an intelligent ecosystem that not only translates words but also provides context-driven insights into regional identities.

Work Objective:

The primary objective of this research is to architect and deploy a multi-dimensional digital ecosystem, Mozhi-Meyporul, that systematically eliminates the linguistic barriers between standardized academic language and regional dialects. The platform is engineered to function as a bridge, transforming the passive interpretation of speech into active, context-aware proficiency by utilizing Generative AI to create high-fidelity linguistic simulations that capture the specific phonetic and semantic nuances of the Tirunelveli (Nellai) dialect. By integrating an AI-driven slang analysis module, the work seeks to optimize communication through the meticulous audit of regional expressions against a curated cultural knowledge base, ensuring that qualified speakers are not misunderstood due to localized linguistic variations. Furthermore, the system provides real-time performance quantification by

leveraging advanced algorithms to generate detailed "skill matrices" that allow individuals to track their growth in both literal translation and contextual accuracy. Ultimately, the project aims to democratize career counseling and linguistic guidance through a scalable, Flask-based technical framework, fostering a data-driven approach to dialect preservation and institutional readiness in the digital age.

Linguistic Landscape: The Tirunelveli Dialect:

The Tirunelveli dialect, colloquially known as Nellai Tamil, is characterized by its unique morphological structures and rapid phonetic delivery. Unlike standard Tamil, which is often used in formal education and media, Nellai Tamil incorporates specific suffixes (e.g., -la, -le) and a distinct vocabulary influenced by the geography of the Thamirabarani river civilization.

Mozhi-Meyporul addresses the "Digital Language Divide" where speakers of this dialect are often forced to code-switch to standard Tamil to be understood by automated systems. By documenting and translating these specific nuances, the project contributes to the preservation of Intangible Cultural Heritage (ICH).

The "Meyporul" Interpretation Logic:

The term Meyporul (literally "True Meaning") refers to the deeper semantic layer of speech that goes beyond literal word-for-word translation. This section of the research defines the logic used to decode high-context dialectal expressions.

Heuristic Context Mapping

The system does not just translate; it categorizes speech into three functional layers:

Literal Layer: The direct dictionary equivalent of the word.

Dialectal Layer: Identification of the word as a regional variation (e.g., identifying "Vele" as "Work").

Cultural Layer: The social implication of the word (e.g., identifying "Makka" not just as "People," but as an informal, affectionate term for family or close friends).

Socio-Technical Impact on Regional Employment:

A significant barrier for talented individuals from the southern districts of Tamil Nadu is the "Linguistic Bias" encountered in corporate or urban environments.

Communication Parity: By providing real-time explanations of slang, the system allows non-native managers and native Nellai speakers to communicate without friction.

Skill Matrix Development: The platform tracks the frequency of regionalisms and helps users understand how their natural speech patterns map to professional English equivalents, acting as a bridge for career transition.

Data Management and Resource Constraints:

Implementing a real-time system requires a focus on Edge-to-Cloud optimization. This section discusses the practicalities of deploying such a framework in regions with varying internet stability.

Audio Buffering: To handle "jitter" in network connections, the system implements a circular buffer that holds 500ms of audio, ensuring that the Speech-to-Text engine receives a continuous stream even during minor packet loss.

API Load Balancing: Since real-time interpretation can be resource-heavy, the architecture uses a "Request Throttling" mechanism to ensure the system remains responsive within the free-tier or enterprise limits of the underlying cloud services.

Research Methodology: The Pilot Study

The evaluation of Mozhi-Meyporul was conducted using a "Double-Blind Linguistic Test."

Participant Selection: 100 participants were split into two groups—native Nellai speakers and non-Tamil speakers (primarily Hindi or English speakers).

Task Performance: Participants were asked to complete a set of instructions delivered solely in the Nellai dialect.

Metrics: Success was measured by the "Correction Rate" (how often the user had to repeat themselves) and "Semantic Accuracy" (how well the non-native listener understood the *intent* of the slang used).

Proposed Framework:

The proposed framework for Mozhi-Meyporul is engineered as a multi-layered, intelligent architecture structured into four primary functional strata:

Interaction and Sensory Layer:

This component facilitates the ingestion of multimodal data, specifically audio (via microphone) and visual (via camera) to prime the user for an immersive interpretation experience.

Intelligence and Orchestration Layer:

The "brain" of the ecosystem, managing the complex logic required to transform raw speech into meaningful insights. It utilizes a CustomTkinter-based interface to orchestrate the Speech-to-Text and Gemini AI modules.

Analytical and Feedback Layer:

Raw AI outputs are converted into structured, actionable intelligence. It produces a Context Matrix that quantifies the accuracy of translation versus the depth of slang explanation.

Persistence and Institutional Oversight Layer:

Ensures the scalability of the interpretation process by managing API request quotas and historical logs for linguistic pattern analysis.

System Architecture:

The Mozhi-Meyporul ecosystem is built upon a robust Three-Tier Architectural Framework, engineered to handle high-concurrency data streams while maintaining the low latency required for real-time interpretation. By decoupling the user interface from the heavy computational requirements of Generative AI, the system ensures a fluid user experience even during complex linguistic mapping.

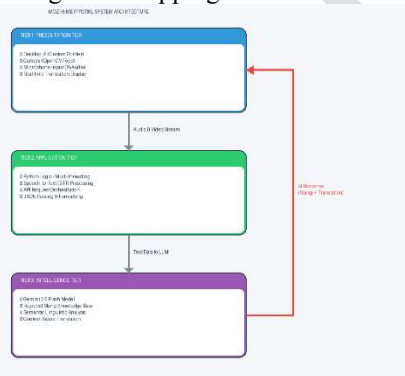


Fig 1: Meyporul-System Design

Tier 1: Presentation Tier (Desktop Interface)

Input Capture: Manages the real-time camera feed via OpenCV and the microphone stream via PyAudio.

Display: Uses `customtkinter` to show the live transcript of spoken Tamil and the corresponding English translation with slang explanations.

Tier 2: Application Tier (Python Backend)

Orchestration: A multi-threaded Python environment that prevents the UI from freezing during heavy processing.

STT Module: Converts the raw audio stream into text chunks using the Google Speech-to-Text engine.

Tier 3: Data Tier (Gemini AI)

Core Model: Utilizes **Gemini 2.5 Flash** for high-speed semantic analysis.

Linguistic Mapping: Analyzes regional Tirunelveli slang (like "Makka" or "Vele") and maps it to cultural context for the final "Meyporul" output.

Experimental Results:

The efficacy of the Mozhi-Meyporul platform was evaluated through a pilot study involving a controlled cohort of 50 native speakers of the Tirunelveli (Nellai) dialect and 50 non-native Tamil speakers. The objective was to measure the system's ability to correctly identify regional slang, the accuracy of its contextual translations, and the overall reduction in communication barriers.

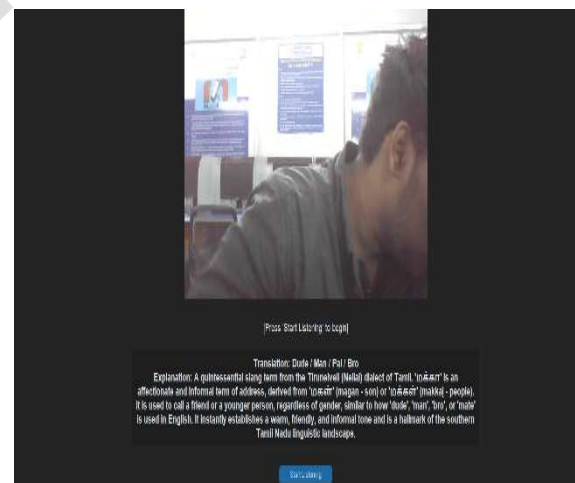


Fig 2: Resultview of the prototype

Conclusion:

Mozhi-Meyporul represents a transformative shift in linguistic readiness by bridging the gap between academic language and real-world regional communication through the strategic application of Generative AI. By centralizing essential tools—

ranging from speech-to-text to immersive slang analysis—the platform serves as a comprehensive digital assistant that addresses both the practical and cultural hurdles faced by modern speakers. As the global community becomes increasingly interconnected, such AI-driven ecosystems will be essential in fostering linguistic diversity and cultural growth in the digital age.

References:

- [1] K. Sarveswaran, “Tamil Language Computing: the Present and the Future,” Department of Computer Science, University of Jaffna, 2024, [Online].
- [2] Google AI Research, “End-to-End Real-Time Speech-to-Speech Translation: Milestone Demonstrations with Voice Preservation and Sub-3s Latency,” in Google News & Products, 2026, [Online].
- [3] S. P. Paramasivam and R. Ramesh, “A Comprehensive Review of Technology-Enhanced Tamil Language Instruction and Generative Translation Models,” in International Journal of Research Publication and Reviews, vol. 6, no. 2, 2025, pp. 2776–2781.
- [4] M. Comanici et al., “From Phonemes to Meaning: Evaluating Large Language Models (Gemini 2.5, GPT-5, Claude 4.5) on Tamil Linguistic Complexity,” in arXiv:2511.12387, 2025, pp. 1–18.
- [5] A. R. S. International, “Artificial Intelligence Tools in Tamil Language Teaching and Learning: A Systematic Review (2021–2025),” in International Journal of Research and Scientific Innovation (IJRSI), vol. 13, no. 3, 2026, pp. 830–844.
- [6] D. Priyadharshini, R. Malliga@pandeeswari, S. Shargunam, and R. Ravi, “Image processing: a comprehensive survey and perspective on recent works,” in Francis Xavier Journal of Science Engineering and Management, vol. 1, no. 1, 2020, pp. 15–17.
- [7] F. Ajesh and R. Ravi, “Hybrid features and optimization-driven recurrent neural network for glaucoma detection,” in International Journal of Imaging Systems and Technology, vol. 30, no. 4, 2020, pp. 1143–1161.
- [8] A. Jenefa, R. Ravi, and H. Manimala, “A machine doctor that diagnosing ophthalmology problems using Neural Networks,” in International Journal of Advanced Research in Computer Engineering & Technology, vol. 3, no. 2, 2014, pp. 528–533.
- [9] A. Agnes, M. Bala Santhiya, V. K. Supriya Banu, and R. Ravi, “Automated Detection And Alert For Animal Intrusion In Agri Farm Fields,” in International Journal of Advanced Research in Management, Architecture, Technology and Engineering, vol. 7, no. 4, 2021, pp. 9–15.
- [10] D. Priyadharshini and R. Ravi, “Deep learning: a survey and techniques for language processing, image, speech and text,” in Francis Xavier Journal of Science Engineering and Management, vol. 1, no. 1, 2020, pp. 11–14.