Vidyalaya: AI - Powered Telugu Language Learning Platform

Anusha C Reddy, M Hemasri, Priyanka C Nair
Department of Computer Science & Engineering,
Amrita School of Computing, Bengaluru,
Amrita Vishwa Vidyapeetham, India
bl.en.u4aie22004@bl.students.amrita.edu, bl.en.u4aie22028@bl.students.amrita.edu, v_priyanka@blr.amrita.edu

***Abstract*—Language learning in morphologically rich, low-resource languages like Telugu presents unique challenges due to the scarcity of intelligent, adaptive learning tools that offer real-time, personalized feedback. Most existing platforms prioritize globally dominant languages and often lack the linguistic depth and cultural sensitivity required to support regional learners effectively. Addressing this gap, this research introduces Vidyālaya, an AI-powered, bilingual language learning platform designed specifically for Telugu. The system integrates Large Language Models (LLMs) for grammar analysis and Wav2Vec2-based speech recognition to assess pronunciation from spoken input. It provides immediate, bilingual feedback — including grammar corrections, pronunciation evaluations, and contextual explanations in both Telugu and English — aimed at enhancing learner comprehension and fluency.**

To improve adaptability across varying language proficiencies, we employ zero-shot and few-shot learning techniques via the Mistral API. Our evaluation reveals that few-shot prompting significantly outperforms zero-shot, offering deeper contextual understanding and more accurate corrections. The platform also features an interactive chatbot capable of handling translations, definitions, and grammar assistance in both languages. With an intuitive user interface and culturally aware design, Vidyālaya offers a holistic and scalable solution that bridges the gap between AI-driven language education and the linguistic diversity of India.

***Index Terms*—Natural Language Processing, Telugu Language Learning, Grammar Correction, Pronunciation Feedback, Speech Recognition, Large Language Models, Few-shot Learning, Bilingual AI, Regional Language Education**

I. Introduction

In today's digital age, language learning plays a pivotal role in education, communication, and cultural preservation. The domain of language learning and evaluation has grown significantly, aiming to bridge the gap between linguistic ability and real-time application for both native and non-native speakers. Particularly in multilingual countries like India, there is an increasing demand for educational tools that support regional languages and dialects. Traditional approaches often fail to deliver personalized, context-aware feedback, especially in low-resource languages, highlighting the need for intelligent, automated systems that can enhance user understanding and fluency across writing and speech dimensions.

The rise of Artificial Intelligence (AI), particularly Machine Learning (ML) and Natural Language Processing (NLP) [13], has revolutionized the way we approach language-related challenges. These technologies have enabled the development of systems capable of sentiment analysis [14] in underrepresented languages like Telugu, POS tagging in morphologically rich structures [15], and zero-shot or few-shot classification of native texts with minimal training data [16]. Furthermore, AI models are now used in tasks such as accent classification and code-mixed translation [17], [18]—applications that demand deep contextual understanding and adaptive learning. With the help of Large Language Models (LLMs), multilingual ASR, and fine-tuning techniques, NLP-based platforms are now becoming more inclusive and capable of addressing the nuances of regional languages, ultimately aiming to make language learning more accessible, accurate, and engaging.

Our research contributes to this domain by introducing a comprehensive, AI-powered Telugu language learning platform that addresses critical gaps not thoroughly explored in previous works.

- Current work focuses on delivering real-time, bilingual grammar and pronunciation analysis tailored specifically for Telugu, a morphologically rich and phonetically diverse language.
- Implementing interactive learning mechanisms that provide contextual corrections along with explanations in both English and Telugu, promoting deeper understanding.
- The approach incorporates zero-shot and few-shot learning paradigms for both speech and text input, enabling robust performance even in data-scarce scenarios.

Finally, the current research introduces a bilingual AI chatbot to support learners beyond static evaluations, offering contextual meaning, translation, and grammar assistance. Together, these innovations contribute to advancing the capabilities of AI in language education and enhancing inclusivity for regional language speakers in the digital era.

The paper is structured as follows: Section II discusses related work in Language Learning; Section IV describes Vidyalaya's methodology, the architecture of the system and NLP techniques employed, Section V includes the result of the work in this study, with an analysis of the performance of all the modules; and Section VI offers conclusions for the study.

II. Related Work

This study presents a language tutoring system using end-to-end ASR and AI-based proficiency scoring to support non-

native speakers. [1] It employs transformer-based ASR models with semi-supervised and transfer learning for improved fluency detection. Deployed in real platforms, it achieved a 55.7% error reduction, but still requires enhancement in dialogue handling and granular feedback. This work introduces a hybrid MLP-LSTM model using MFCCs to enhance English speech recognition accuracy for personalized learning. [2] It achieved 98.25% accuracy and a Word Error Rate (WER) of 0.075, outperforming traditional models. Despite strong results, the study suggests exploring lighter models for multilingual applications. A cascaded framework is proposed to assess L2 speaking proficiency through ASR-based grammar error correction. [3] Using transformer-based correction and rule-based tools like ERRANT, it improved assessment validity and explainability. However, it highlights the lack of annotated learner speech datasets. This study evaluates ASR combined with peer feedback for pronunciation improvement in EFL learners. [4] With 155 participants, the approach significantly enhanced speech and interpersonal skills. It emphasizes the need for socially interactive ASR tools. This research assesses the ELSA speech analyzer's role in boosting speaking proficiency among university students. [5] Real-time feedback on pronunciation and fluency led to significant improvements in the experimental group. The study calls for long-term impact evaluation and curriculum integration.DPCSpell, a denoising transformer-based spelling corrector, is developed for low-resource Indic languages like Telugu and Bangla. [6] Its three-phase architecture achieved 94.78% exact match accuracy, outperforming prior systems. The study recommends more open corpora and broader script support. AraSpell is a Seq2Seq deep learning model using RNNs and Transformers for Arabic spelling correction. [7] Trained with noise-injected data, it achieved a low WER (4.8%) and CER (1.11%). Future work should improve support for dialectal variations. This research develops an ASR- and phonetics-based tool for ELL pronunciation correction, using DNNs and HMMs. [8] It delivered 88% accuracy in analyzing stress and intonation and supported gamified learning. Broader linguistic adaptation is needed for global use.

This research addresses typographical errors in Malayalam social media comments using two models: a sequence-to-sequence model and a hybrid rule- and graph-based method. [9] The hybrid model achieved 91.26% accuracy but faced limitations due to language complexity and data constraints. The study suggests future improvements via richer datasets and morphosyntactic enhancements. This study focuses on parsing code-switched Telugu-English text using the UDify parser and a 300-sentence annotated dataset. [10] It achieved 56.17% POS accuracy and a low labeled attachment score due to spelling and romanization issues. The work emphasizes the need for larger datasets to improve parsing accuracy in low-resource settings. The research develops a Telugu speech-to-text parser using the Probabilistic CYK algorithm to generate parse trees and validate grammar. [11] It performs well with basic sentence structures and offers a Streamlit-based interface, though it struggles with dialects and complex grammar. Future goals include expanding rules and accommodating more dialects. This paper introduces a large Hindi-Marathi code-switched dataset and uses a transformer ASR model enhanced with Q-learning for dynamic language switching. [12] With a WER of 0.28 and CER of 0.24, it outperforms older models but faces challenges with accent variation and intra-sentence switching. The study proposes hybrid methods and synthetic data for improvement.

The literature highlights substantial progress in AI-based language learning, particularly in speech recognition, grammar correction, and spelling error detection using deep learning and transformer models. However, these advancements are predominantly focused on high-resource languages, with limited application to low-resource, morphologically complex languages like Telugu. Key gaps include inadequate support for Telugu's linguistic nuances and the absence of bilingual feedback systems. Addressing these challenges, the current work introduces Vidyālaya, a comprehensive Telugu learning platform that combines LLM-driven grammar analysis, Wav2Vec2-based speech transcription, and dual-language feedback. This solution fosters inclusive, effective learning while contributing to the preservation of regional linguistic diversity.

III. Methodology & Implementation

This section describes the methodology employed in the current work of Telugu Language Learning Platform by which grammar analysis and pronunciation analysis is being performed. The aim was to provide assistance and feedback of the user's speech so that user can learn the language in a right and better manner. The entire system is bilingual inorder to make sure that user can still know how to user the interface even though no idea about telugu language.

The system architecture of Vidyālaya has been designed to offer a modular, scalable, and user-friendly language learning experience that caters to Telugu learners. The entire system is structured into multiple layers: the User Interface Layer, the Application Logic Layer (API Gateway), and the Core AI Processing Modules, which include components for Speech Recognition, Grammar Analysis, Pronunciation Evaluation, and an Intelligent Bilingual Chatbot as shown in Fig. 1. Each layer works in harmony to provide real-time feedback to users in a way that is both pedagogically effective and culturally sensitive.

A. User Interface

At the user interface level, the platform provides a bilingual frontend built using React.js, Next.js, and Gradio, offering a seamless interaction environment. Learners can input text manually or use the integrated microphone feature to speak Telugu sentences. The interface supports bi-directional translation between Telugu and English, allowing users to choose their preferred language for receiving instructions and feedback. The design of the UI emphasizes accessibility and engagement through a clear layout, real-time interaction elements, and score visualizations, making it usable by both novice and experienced users.

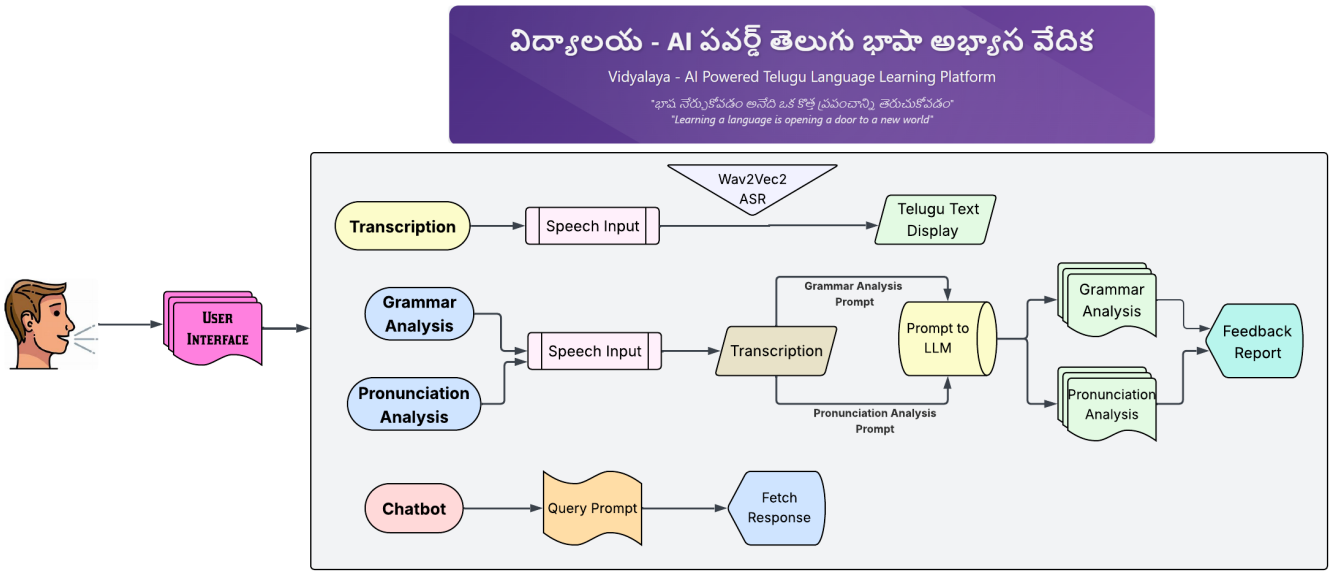


Fig. 1: System Architecture of the Vidyalaya Platform

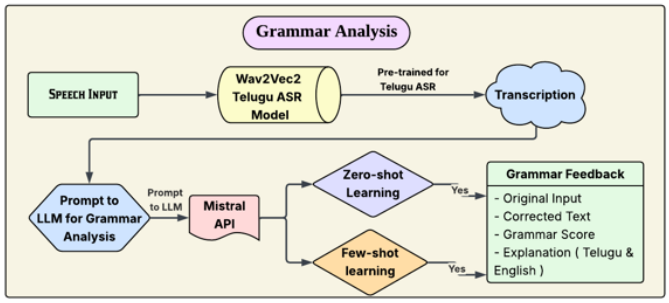


Fig. 2: Grammar Analysis Architecture

```
prompt = f"""
ద్దయచేసి ఈ తెలుగు వాక్యాన్ని విశ్లేషించండి: "{text}"

ద్దయచేసి ఈ క్రింది విధంగా JSON రెస్పాన్స్ ఇవ్వండి:
{{
    "corrected": "సరిదిద్దిన వాక్యం",
    "english_explanation": "వ్యాకరణ తప్పుల వివరణ ఇంగ్లీషులో",
    "telugu_explanation": "వ్యాకరణ తప్పుల వివరణ తెలుగులో",
    "errors": ["తప్పు 1", "తప్పు 2"]
}}
"""
```

Fig. 3: Zero-shot Prompt Used

B. ASR - Transcription

The backend logic is built using FastAPI for AI-related processing and Express.js to manage session routing and external service integration. Once a user inputs text or speech, the data is routed via the API gateway to the appropriate processing pipeline. For speech input, the backend invokes a Wav2Vec2-based Automatic Speech Recognition (ASR) [19] model, which is pretrained on multilingual datasets and fine-tuned to handle phonetic variations common in Telugu. The ASR model converts spoken input into text while preserving linguistic nuance. This transcription serves as the foundation for both grammar and pronunciation analysis.

C. Grammar Analysis & Feedback

Once the text is available—either transcribed from speech,it is passed to the Grammar Analysis Module, powered by LLMs accessed via Mistral [18]API, Fig.2. This module employs both zero-shot Fig. 3. and few-shot Fig.4,5.learning techniques using carefully crafted prompts. The prompt design ensures that the model returns not just the corrected sentence but also explains the nature of the error in both Telugu and English.

Zero-shot prompt Explanation:

- ”Please analyze this Telugu sentence: ’[User’s sentence]’
- Provide a response in the following JSON format, including:
- The corrected sentence
- Explanation of grammar mistakes in English
- Explanation of grammar mistakes in Telugu
- A list of the identified grammar errors”

Few-shot Prompt Explanation:

- Meaning in English: ”You are an expert in the Telugu language. Please analyze the following sentence and correct it.
- Examples of Telugu grammar rules: (Here, representative rule-based examples are given for few-shot learning)
- A key pattern example: Pattern: [grammar rule] Example: [sample sentence] Explanation: [explanation of the correction]
- Sentence to be analyzed: ’[User’s Telugu sentence]’ Please respond in the following JSON format, including the corrected sentence, explanations in both English and Telugu, a list of grammar errors, and their corresponding error types (e.g., sentence structure, verb form, etc.).”

```
prompt = f"""
మీరు తెలుగు భాషా నిపుణులు. దయచేసి ఈ క్రింది వాక్యాన్ని విశ్లేషించండి మరియు సరిదిద్దండి.

### తెలుగు వ్యాకరణ నియమాల ఉదాహరణలు:
{examples_str}

### ఒక ముఖ్యమైన నమూనా:
నమూనా: {self.one_shot_example['pattern']}
ఉదాహరణ: {self.one_shot_example['example']}
వివరణ: {self.one_shot_example['explanation']}

### విశ్లేషించాల్సిన వాక్యం:
"{text}"

దయచేసి ఈ క్రింది JSON ఫార్మాట్‌లో జవాబు ఇవ్వండి:
{{
    "corrected": "సరిదిద్దిన వాక్యం",
    "english_explanation": "వ్యాకరణ తప్పుల వివరణ ఇంగ్లీషులో",
    "telugu_explanation": "వ్యాకరణ తప్పుల వివరణ తెలుగులో",
    "errors": ["తప్పు 1", "తప్పు 2"],
    "error_types": ["వాక్య నిర్మాణం", "క్రియా రూపం"]
}}
"""
```

Fig. 4: Few-shot Prompt Used

```
self.few_shot_examples = [
    {
        "incorrect": "నాకు తెలుగు మాట్లాడలేను",
        "correct": "నాకు తెలుగు మాట్లాడటం రాదు",
        "explanation": "క్రియా రూపంలో తప్పు. 'మాట్లాడలేను' కాకుండా 'మాట్లాడటం రాదు' అని సరిగ్గా వాడాలి."
    },
    {
        "incorrect": "అతను బడికి వెళ్లాడు",
        "correct": "అతను బడికి వెళ్లాడు",
        "explanation": "సరైన వాక్యం. 'బడికి' అనేది సరైన సందర్భంలో ఉంది."
    },
    {
        "incorrect": "మా అమ్మ మంచి ఉడుతున్నారు",
        "correct": "మా అమ్మ మంచి వండుతున్నారు",
        "explanation": "క్రియా పదం తప్పు. 'ఉడుతున్నారు' కాకుండా 'వండుతున్నారు' అని సరిగ్గా వాడాలి."
    }
]

self.one_shot_example = {
    "pattern": "Subject + Object + Verb (కర్త + కర్మ + క్రియ)",
    "example": "రాము పుస్తకం చదివాడు",
    "explanation": "తెలుగు వాక్య నిర్మాణం సాధారణంగా కర్త, కర్మ, క్రియా క్రమంలో ఉంటుంది."
}
```

Fig. 5: Few-shot Examples Prompted for Learning

This dual-language feedback mechanism is a significant pedagogical feature that bridges the understanding gap for regional learners who may not be fluent in English. The grammar engine is dynamic and context-sensitive, handling syntactic, morphological, and semantic deviations with a high degree of precision.

D. Pronunciation Analysis

In parallel, the Pronunciation Evaluation Module works with the original speech input to assess the learner's articulation Fig. 3. It uses phoneme alignment techniques and acoustic similarity scoring to detect deviations from the ideal pronunciation that is captured by the transcription module in the transcribed text. This is not limited to simply identifying correct or incorrect speech but involves a granular phoneme-level comparison, offering precise feedback that can guide users toward more accurate speech patterns. The feedback is presented in an interpretable format, including visual indicators (score bars, badges) and textual guidance in Telugu.

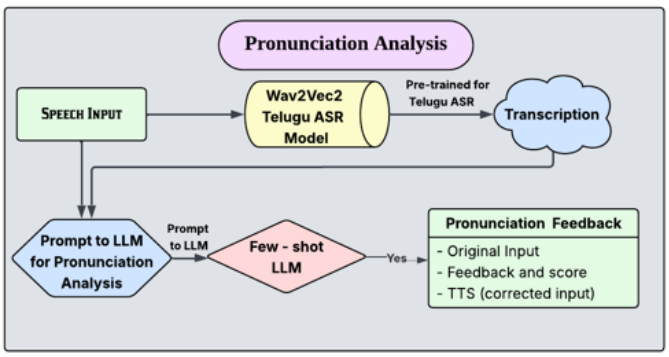


Fig. 6: Pronunciation Analysis Architecture

Another core component of the system is the AI-powered Bilingual Chatbot, designed to offer supplemental support for learners. This module is also built on LLMs and is configured through prompt engineering to handle a wide range of tasks—such as vocabulary lookup, meaning explanation, sentence translation, grammar rule explanation, and even generating custom Telugu scripts or practice prompts. The chatbot responds in both Telugu and English, based on user preference, thus reinforcing the bilingual nature of the platform and enhancing learner autonomy.

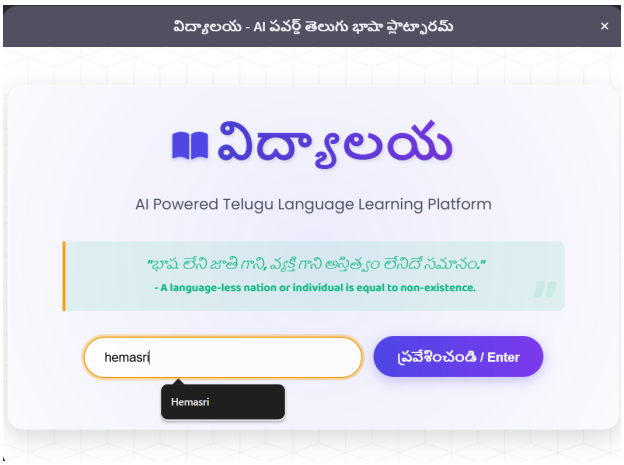
Finally, all outputs—grammar scores, pronunciation analysis, corrected text, and chatbot responses—are integrated and displayed in the frontend in an intuitive, user-friendly manner. Learners can view their results instantly, including detailed breakdowns of errors and suggestions. The system also includes progress tracking features and performance analytics that help users understand their learning trajectory over time.

This architecture ensures that Vidyālaya delivers a high-quality, AI-driven learning experience by integrating advanced NLP and speech recognition technologies within a culturally-aware, educational framework. Its modular design allows easy extension to other Indian languages in the future, and its use of bilingual, explainable feedback sets it apart from generic grammar or speech tools currently available. The system is not only a technological innovation but also a meaningful contribution to regional language education through artificial intelligence.

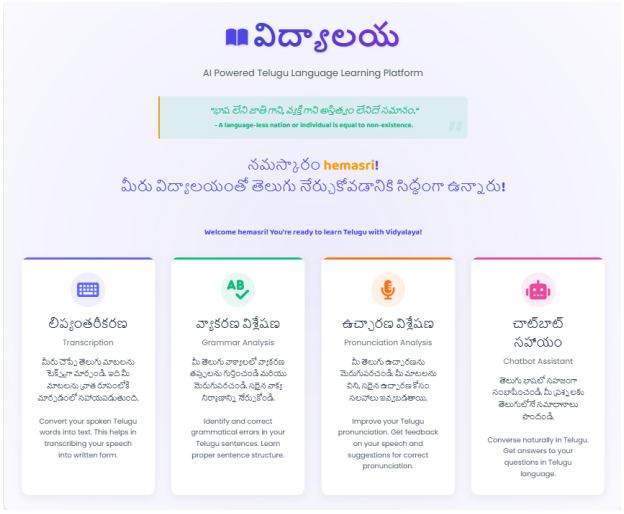
IV. Results & Analysis

The experimental evaluation of the Vidyālaya platform Fig. 7., was conducted to analyze two core components: speech-to-text transcription accuracy and grammar correction efficiency using zero-shot and few-shot learning techniques. The Automatic Speech Recognition (ASR) component, based on Wav2Vec2, was tested against real-time spoken Telugu sentences Fig. 8., while grammar analysis was evaluated using text inputs transcribed from speech and manually typed content. The system's end-to-end performance was assessed using standard metrics such as Word Error Rate (WER) [21], Character Error Rate (CER) [22], BLEU Score [23], Levenshtein Similarity [25], and BERT Score [24], allowing for a thorough examination of model accuracy and semantic alignment Fig. 9.

In the ASR component, the system recorded a WER of 0.45, indicating moderate error rates at the word level, which were primarily due to phonetic similarities in regional speech and inconsistent spacing in the Telugu script. However, the CER was significantly lower at 0.10, showing that character-level recognition was highly accurate. The BLEU score of 1 and

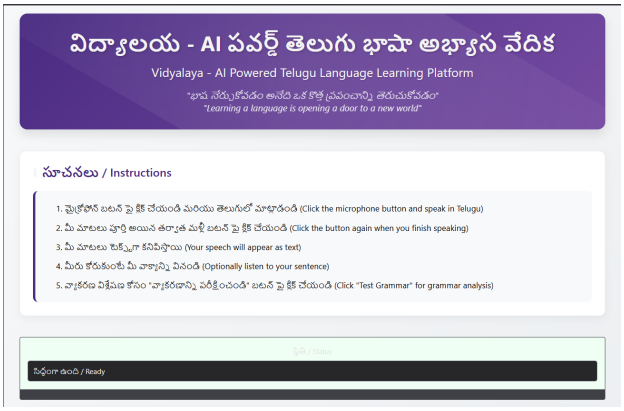


(a) Interface showing user input section.

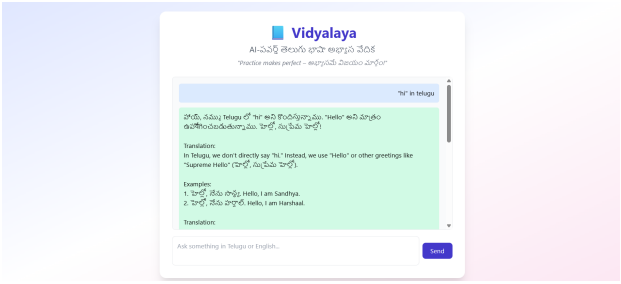


(b) Multiple choices of analysis

(c) Grammar correction output display.



(d) Pronunciation scoring feedback.



(e) Chatbot performing bilingual interactions.

Fig. 7: User Interface outputs of the Vidyālaya platform, showing various modules including input panel, speech handling, grammar correction, pronunciation analysis, and chatbot interaction.

Levenshtein similarity of 0.90 further suggest that in most cases, the transcription was semantically correct even if minor structural inconsistencies were present. These findings affirm the effectiveness of Wav2Vec2 for handling Telugu, despite the lack of extensive accent-specific tuning. As shown in Table I: Sample Telugu Sentences with Grammar Corrected Sentence by LLM, the ASR outputs were reasonably close to the spoken sentences, and even in noisy transcriptions, the downstream grammar correction module handled contextual restructuring effectively.

The grammar correction module was evaluated using zero-shot and few-shot (one-shot) prompting strategies via the Mistral [20]API. Both approaches achieved BLEU scores of 1, indicating exact matches with reference sentences. However, the BERT scores, which focus on semantic similarity, offered more insight into contextual comprehension, with both zero-shot and few-shot methods scoring 0.90. The qualitative results indicate that few-shot prompting significantly outperformed zero-shot prompting in restructuring complex sentences and resolving context-dependent errors. For instance, in the 4th sentence

TABLE I: SAMPLE TELUGU SENTENCES WITH GRAMMAR CORRECTED SENTENCE BY LLM

Sample Telugu Sentences			Grammar Correction by LLM's	
Correct Sentence	**Spoken Sentence**	**ASR - Transcription**	**Zero - shot Correction**	**One - shot Correction**
ఈ పుస్తకం నాకు చాలా నచ్చింది.	ఈ పుస్తకం నాకు చాల నచ్చినాది.	ఈ పుస్తకం నాకు చాలా నచ్చిన అది	ఈ పుస్తకం నాకు చాలా నచ్చింది	ఈ పుస్తకం నాకు చాలా నచ్చింది
ఇప్పుడు ఆరంభిద్దాం.	ఇప్పుడు ఆరంభిందాం.	ఇప్పుడు ఆరంబిందాం	ఇప్పుడు ప్రారంభిద్దాం	ఇప్పుడు ప్రారంభిద్దాం
రాముడు మంచి విద్యార్థి.	రాముడు మంచి విద్యార్థినీ.	రాముడు మంచి విద్యార్థిని	రాముడు మంచి విద్యార్థి	రాముడు మంచి విద్యార్థి
అవనిని నేను నిన్న కలిశాను.	అవనిని నేను నిన్న కలసాము.	అవని నీ నేను నిన్న కలిసాము	అవని నీకు నేను నిన్న కలిసాను	నిన్న నేను అవనిని కలిసాను
అమ్మ కూర వండుతుంది.	అమ్మ కూర ఉడుతున్న.	అమ్మా కూరా పడుతున్నా	అమ్మా కూర పడుతున్నా	అమ్మ కూర వండుతోంది

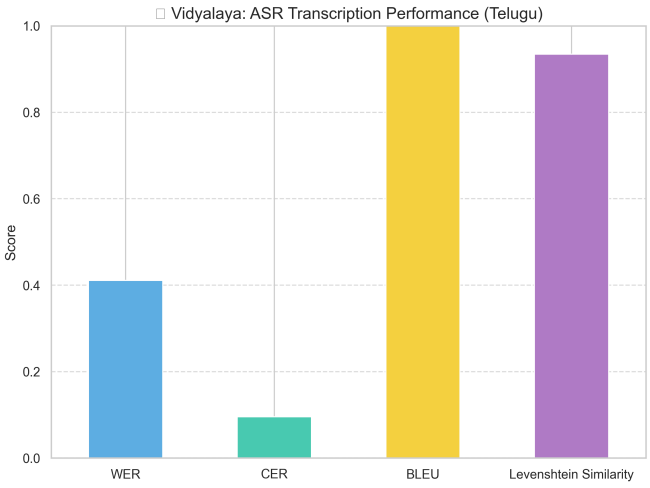


Fig. 8: ASR Telugu Transcription Performance

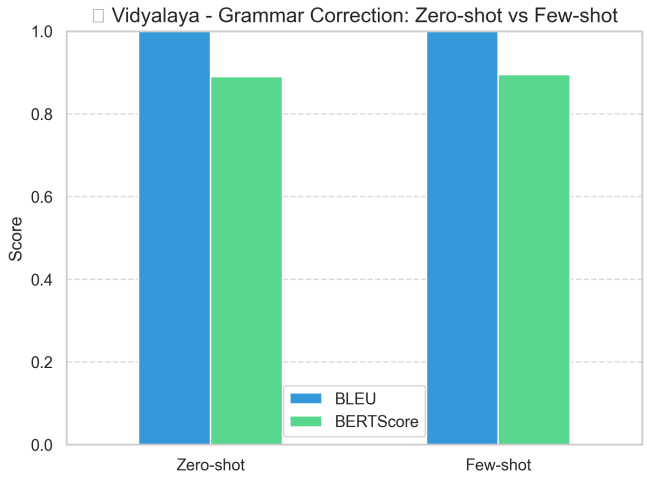


Fig. 9: Grammar Correction Metrics

of Table I., the few-shot correction accurately rephrased it demonstrating better grammatical alignment and fluency. Similarly, verb aspect correction such as sentence 5 in Table I., being transformed to right form shows the few-shot model's superior understanding of tense and voice.

In addition to individual performance, the robustness of the system lies in the combined pipeline, where the ASR output, even with slight inaccuracies, feeds into a grammar correction system that can still deliver semantically accurate and grammatically correct feedback. The consistency across corrections shown in the table and supported by the visual graphs in the results section, confirms that the integration of speech recognition and language modeling through zero-shot/few-shot prompting provides a reliable learning experience for Telugu learners. The performance, especially in few-shot mode, demonstrates that LLMs can effectively adapt to low-resource, morphologically complex languages when supported by minimal examples and well-structured prompts.

V. Conclusion & Future Enhancements

This study successfully demonstrates the design and implementation of Vidyālaya, an AI-powered, bilingual platform for Telugu language learning, that integrates large language models for grammar analysis and Wav2Vec2-based ASR for pronunciation evaluation. Through the use of zero-shot and few-shot learning strategies, the system delivers real-time, accurate feedback on grammatical correctness and spoken fluency. The evaluation results confirm that few-shot prompting not only matches but often exceeds zero-shot methods in contextual understanding and structural correction, thereby offering learners a more effective and natural learning experience. Furthermore, the integration of bilingual explanations and a supportive chatbot ensures accessibility and cultural relevance, making the platform especially valuable for underrepresented linguistic communities.

While the current system demonstrates strong performance, there remains scope for further enhancement. Improving the ASR model through fine-tuning on accent-rich Telugu datasets could significantly reduce WER and boost phoneme-level precision. Additionally, incorporating dynamic pronunciation scoring and more nuanced speech prosody analysis would offer deeper feedback to users. From a pedagogical perspective, future iterations of Vidyālaya could implement adaptive learn-

ing paths, gamified progress tracking, and offline functionality to support learners in rural or connectivity-constrained environments. Moreover, expanding the platform to support other Dravidian and Indic languages using the same modular framework could amplify its impact across India's multilingual landscape. Ultimately, Vidyālaya bridges a critical gap in regional language education by leveraging the power of AI to deliver a context-sensitive, learner-centric solution tailored for the digital age.

REFERENCES

[1] Kang, Byung Ok, Hyung-Bae Jeon, and Yun Kyung Lee. "AI-based language tutoring systems with end-to-end automatic speech recognition and proficiency evaluation." ETRI Journal 46, no. 1 (2024): 48-58.

[2] Orosoo, Myagmarsuren, Namjildagva Raash, Mark Treve, Hassan Fareed M. Lahza, Nizal Alshammry, Janjhyam Venkata Naga Ramesh, and Manikandan Rengarajan. "Transforming English language learning: Advanced speech recognition with MLP-LSTM for personalized education." Alexandria Engineering Journal 111 (2025): 21-32.

[3] Bannò, Stefano, and Marco Matassoni. "Back to grammar: Using grammatical error correction to automatically assess L2 speaking proficiency." Speech Communication 157 (2024): 103025.

[4] Rogti, Maroua. "Automatic pronunciation evaluation feedback and peer correction for shaping English pronunciation accuracy and interpersonal communication." Innovation in Language Learning and Teaching (2025): 1-17.

[5] Al-husban, Naima. "The impact of AI-assisted language learning tools on augmenting university EFL students' speaking skills in Jordan." Journal of Applied Learning and Teaching 8, no. 1 (2025).

[6] Bijoy, Mehedi Hasan, Nahid Hossain, Salekul Islam, and Swakkhar Shatabda. "A transformer-based spelling error correction framework for Bangla and resource-scarce Indic languages." Computer Speech & Language 89 (2025): 101703.

[7] Salhab, Mahmoud, and Faisal Abu-Khzam. "AraSpell: A deep learning approach for Arabic spelling correction." arXiv preprint arXiv:2405.06981 (2024).

[8] Lara, Mathusha Sam, R. Subhashini, Crispine Shiny, Joy Christy Lawrance, S. Prema, and S. Muthuperumal. "Constructing an AI-Assisted Pronunciation Correction Tool Using Speech Recognition and Phonetic Analysis for ELL." In 2024 10th International Conference on Communication and Signal Processing (ICCSP), pp. 1021-1026. IEEE, 2024.

[9] Ratnam, D. Jyothi, A. N. Karthika, K. Praveena, Rajesh Taniya, S. Thara, and Nedungadi Prema. "Phonogram-based Automatic Typo Correction in Malayalam Social Media Comments." Procedia Computer Science 233 (2024): 391-400.

[10] Putta, Akshith. "A dependency parser for code-switched Telugu-English." Journal of High School Science 8, no. 2 (2024): 143-155.

[11] Varshini, Surisetty Hima, Gottimukkala Sarayu Varma, and Meena Belwal. "A Speech2Text based Recognizer and Probabilistic Parser Application for TeluguSentences." In 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), pp. 1-6. IEEE, 2024.

[12] Hemant, P., and Meera Narvekar. "Development of a code-switched Hindi-Marathi dataset and transformer-based architecture for enhanced speech recognition using dynamic switching algorithms." Applied Acoustics 230 (2025): 110408.

[13] Karthik, A. H. N., Ch Aneesh, Gembali Saumik, K. V. V. Varun, and Kavitha CR. "Sentiment Analysis on Telugu Text Translated from English Using NLP and ML." In 2025 3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT), pp. 246-253. IEEE, 2025.

[14] Jayanth, K. Krishna, G. Bharathi Mohan, and R. Prasanna Kumar. "Indian language analysis with xlm-roberta: Enhancing parts of speech tagging for effective natural language preprocessing." In 2023 Seventh International Conference on Image Information Processing (ICIIP), pp. 850-854. IEEE, 2023.

[15] Boyina, Kamal, Gujja Manaswi Reddy, Gunturi Akshita, and Priyanka C. Nair. "Zero-Shot and Few-Shot Learning for Telugu News Classification: A Large Language Model Approach." In 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), pp. 1-7. IEEE, 2024.

[16] Jairam, R., G. Jyothish, and B. Premjith. "A Few-Shot Multi-Accented Speech Classification for Indian Languages using Transformers and LLM's Fine-Tuning Approaches." In Proceedings of the Fourth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages, pp. 1-9. 2024.

[17] Nair, Aarathi Rajagopalan, and Deepa Gupta. "Evaluating Performance and Accuracy of Large Language Models in Translating Code-Mixed Hindi to English: A Comparative Study." In 2024 IEEE 21st India Council International Conference (INDICON), pp. 1-6. IEEE, 2024.

[18] https://chat.mistral.ai

[19] telugu asr - MAvinash24, "Live Telugu Speech Recognition," GitHub repository, 2025. [Online]. Available: https://github.com/MAvinash24/ML_Project

[20] Aydin, Omer, Enis Karaarslan, Fatih Safa Erenay, and Nebojsa Bacanin. "Generative AI in Academic Writing: A Comparison of DeepSeek, Qwen, ChatGPT, Gemini, Llama, Mistral, and Gemma." arXiv preprint arXiv:2503.04765 (2025).

[21] Ali, Ahmed, and Steve Renals. "Word error rate estimation for speech recognition: e-WER." In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), pp. 20-24. Association for Computational Linguistics (ACL), 2018.

[22] Park, Chanho, Hyunsik Kang, and Thomas Hain. "Character Error Rate Estimation for Automatic Speech Recognition of Short Utterances." In 2024 32nd European Signal Processing Conference (EUSIPCO), pp. 131-135. IEEE, 2024.

[23] Saadany, Hadeel, and Constantin Orasan. "BLEU, METEOR, BERTScore: Evaluation of metrics performance in assessing critical translation errors in sentiment-oriented text." arXiv preprint arXiv:2109.14250 (2021).

[24] Cui, Yizhuo, and Maocheng Liang. "Automated Scoring of Translations with BERT Models: Chinese and English Language Case Study." Applied Sciences 14, no. 5 (2024): 1925.

[25] Antentyk, Yurii, and Serhii Havrylov. "Extracting Text Representation from Pretrained Language Models via Generalized Levenstein Distance."