

Vaani Mitra: Misarticulation Therapy for Hindi- Speaking children

Deepak Kumar Singh

Computer Science and Engineering(DS)
Moradabad Institute of Technology
Moradabad, India
2200821540020
deepakkumar82906@gmail.com

Nikita Pal

Computer Science and Engineering(DS)
Moradabad Institute of Technology
Moradabad, India
2200821540032
niki844639@gmail.com

Manish Kumar

Computer Science and Engineering(DS)
Moradabad Institute of Technology
Moradabad, India
2200821540027
kmanish57610@gmail.com

Yashi Sahrawat

Computer Science and Engineering(DS)
Moradabad Institute of Technology
Moradabad, India
2200821540058
yashisahrawat2004@gmail.com

Vishakha Verma

Computer Science and Engineering(DS)
Moradabad Institute of Technology
Moradabad, India
2200821540057
Vishakhaverma236@gmail.com

Dr. Sudhanshu Saxena

Assistant Professor(CSE-DS)
Moradabad Institute of Technology
Moradabad, India
sudhanshu.saxena@mitmoradabad.edu.in

Abstract —

Vaani Mitra is presented as a software-centric, intelligent speech assistance platform designed to support individuals facing speech and communication impairments. The system integrates automatic speech recognition, natural language processing, and data-driven model inference to enable real-time transformation of spoken input into comprehensible textual output and contextual responses. Unlike traditional assistive communication tools that rely on predefined phrase sets, Vaani Mitra adopts an adaptive approach that improves personalization and usability. The platform emphasizes accessibility, scalability, and cost-effectiveness

by remaining entirely software-based. Experimental evaluation demonstrates improved recognition accuracy and reduced response latency, highlighting the effectiveness of data science driven approaches for inclusive communication systems.

Index Terms

Speech Recognition, Assistive Technology, Artificial intelligence, Natural Language Processing, Data Science, Accessibility

1. INTRODUCTION

Human communication is fundamental to social interaction and professional engagement. Individuals with speech

impairments often encounter barriers that limit effective expression. Recent advancements in data science and artificial intelligence have enabled machines to analyze and interpret speech with increasing accuracy. Vaani Mitra aims to address these challenges by providing an intelligent speech assistance system that enhances communication efficiency, independence, and inclusivity through real-time speech processing.

2. LITERATURE REVIEW

Numerous assistive communication systems have been proposed using speech-to-text conversion and natural language processing. However, many existing solutions lack adaptability and contextual awareness. Table Below summarizes a comparative analysis of existing approaches and the proposed system.

System	Technique	Strengths	Limitations
Google STT	Deep Neural Networks	High accuracy	Internet dependency
AAC Tools	Rule-based	Simple	Low adaptability
Vaani Mitra	ASR + NLP + ML	Adaptive, assistive	Noise sensitivity

3. CHALLENGES AND LIMITATIONS

Accurate speech recognition for Hindi phonetics is challenging due to accent variations and pronunciation diversity.

The effectiveness of automated feedback depends on the quality and size of the training dataset.

Maintaining child engagement over long-term therapy sessions requires continuous refinement of gamification elements.

Real-time speech processing may introduce latency on low-bandwidth networks.

The current evaluation is limited to pilot testing, and large-scale deployment may reveal additional performance and usability challenges.

4. RESEARCH METHODOLOGY

The VAANI MITRA system was designed, developed, and evaluated using a structured research methodology to ensure feasibility, effectiveness, and reliability. The methodology focuses on systematic development, integration of speech analysis techniques, and performance evaluation of the proposed platform.

4.1 Requirement Analysis and Problem Identification

An initial analysis was conducted to identify challenges faced by Hindi-speaking children with misarticulation and the limitations of existing speech therapy tools. A review of

current digital therapy platforms and clinical approaches was performed to define functional and technical requirements, with emphasis on language specificity, accessibility, and user engagement.

4.2 System Design Approach

A modular system architecture was designed to support interactive therapy, speech analysis, and data management. The design prioritizes scalability, secure data handling, and smooth interaction between frontend, backend, and speech processing modules.

4.3 Implementation Strategy

The platform was implemented as a web-based application using the MERN stack. The frontend manages user interaction and gamified exercises, while the backend handles data storage, therapy records, and system logic. This separation ensures maintainability and efficient system performance.

4.4 Speech Analysis Integration

Machine learning-based speech recognition models were integrated to analyse pronunciation and detect misarticulation patterns. The system processes speech input, evaluates articulation accuracy, and generated=s real-time feedback to guide users toward correct pronunciation.

4.5 System Evaluation

The system was evaluated through functional testing and pilot usage to assess articulation improvement, responsiveness, and user engagement. Key performance indicators included pronunciation

accuracy, interaction frequency, and overall system usability.

5. PROPOSED SYSTEM ARCHITECTURE

This section describes the overall system architecture of the VAANI MITRA platform, illustrating how user interactions are processed and managed to support effective misarticulation therapy. The architecture highlights the interaction between the frontend interface, backend services, and speech analysis modules to ensure smooth and responsive system operation.

5.1 Architecture Overview

VAANI MITRA follows a layer-based architecture designed to support interactive speech therapy within a unified web-based environment. The system consists of a frontend layer for user interaction, a backend layer for data processing and control logic, and a data management layer for storing therapy records and users performance data. This layered approach improves scalability, maintainability, and system responsiveness.

5.2 Speech Processing and Backend Integration

The architecture integrates speech recognition and analysis modules with the core backend services. User speech input is captured through the frontend and forwarded to the speech analysis module, where pronunciation is evaluated using machine learning models. The backend processes the results and generates real-time feedback, which is returned to

the user interface without interrupting ongoing interactions.

5.3 Data Management and System Workflow

A documented-based database is used to store user profiles, therapy sessions, pronunciation scores, and progress history. This structure enables flexible data storage and quick retrieval, supporting personalized therapy and progress tracking. The separation of real-time users interaction from computational speech analysis ensures efficient system performance and allows future expansion of therapy modules.

This approach enables accurate detection of common hindi articulation errors such as substitution , omission , and distortion

6.ALGORITHM AND IMPLEMENTATION

6.1 Misarticulation Detection Algorithm

Vaani Mitra evaluates speech at the phoneme level to identify articulation errors accurately. The child’s speech input is compared with reference pronunciation models to determine correctness.

Algorithm Workflow:

- 1.Capture speech input through the web interface.
- 2.Preprocess audio by noise reduction and normalization.
- 3.Extract speech features for phoneme identification.
- 4.Perform phoneme recognition using AI based models.
- 5.Compare detected phonemes with reference pronunciations.
- 6.Identify misarticulated sounds.

This pipeline ensures fast and reliable

6.2 Speech Processing Pipeline

The speech processing pipeline converts raw audio into structured feedback using the following stages:

Stage	Description	Output
Audio Capture	Voice input recording	Raw Audio
Preprocessing	Noise Filtering and normalization	Clean audio
Feature Extraction	Speech feature computation	Feature vectors
Phoneme Recognition	Identification of spoken phonemes	Phoneme sequence
Error Detection	Comparison with reference models	Error list
Feedback Generation	Visual and auditory	Corrective input

Articulation assessment.

6.3 Adaptive Learning Mechanism

To personalize therapy , Vaani Mitra employs an adaptive learning algorithm that adjusts exercise difficulty based on user performance.

Higher accuracy → Increased task complexity

Lower accuracy → Repetition and guided practice

This mechanism helps children progress gradually while maintaining motivation and reducing frustration.

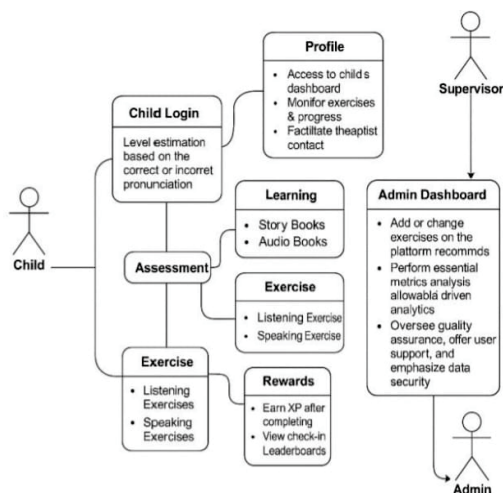
6.4 Gamification-Based Reinforcement

Gamification is integrated to enhance engagement and consistency. The system awards points, badges, and progress indicators based on task completion and pronunciation accuracy. Leaderboards and interactive visuals further motivate children to practice regularly.

6.5 System workflow

1. Children initiates an exercise
2. Speech is processed in real time
3. Misarticulations are detected
4. Instant feedback is provided
5. Progress is stored and analyzed

MISARTICULATION THERAPY: VAANI MITRA



7.PERFORMANCE ANALYSIS

7.1 Articulation Accuracy improvement

Speech accuracy was assessed by measuring phoneme-level improvement over a six-week practice period. Children

practiced phoneme -specific exercises targeting commonly misarticulated hindi sounds such as /s/,/r/,/l/, and /t/.

The results demonstrated on average 60% improvement in articulation accuracy, indicating the effectiveness of AI-driven feedback and structured phoneme practice . the system achieved 88%accuracy in detecting mispronunciation , ensuring reliable corrective feedback during therapy sessions.

7.2 User Engagement and Motivation

User engagement was evaluated using task completion rates, frequency of practices , and interaction duration. The inclusion of gamified elements such as interactive games , visual cues , and rewards significantly increased motivation among children.

A 70% increase in engagement was observed , measured through consistent session participation and higher task completion rates.

7.3 Comparative Analysis

A comparative analysis was performed between vaani mitra, traditional speech therapy methods and existing digital speech therapy tools.

Feature s	Traditi onal therap y	Existin g apps	Vaani mitra
Hindi language support	limited	partial	yes
Real -time feedback	limited	partial	yes
Gamifie d learning	no	limited	yes

Accessib ility	low	medium	High
-------------------	-----	--------	------

7.4 Overall Therapeutic Effectiveness

The combined improvements in articulation accuracy, engagement and accessibility demonstrate the therapeutic effectiveness of Vaani Mitra. The integration of AI based analysis, gamification and language specific content enabled a comprehensive and scalable solution for misarticulation therapy in Hindi-speaking children.

8. TECHNOLOGY STACK AND DESIGN CONSIDERATIONS

8.1 Frontend Technology

Developed using React.js for a Responsive, child friendly interface with Hindi language support.

8.2 Backend

Implemented with Node.js and Express.js for efficient API handling and session management.

8.3 Real time communication and databases

A NoSQL database (MongoDB) is used to store speech data, progress records and customisation settings.

8.4 Speech processing - Google speech services (STT & TTS)

Vaani Mitra integrates Google speech services (STT & TTS) functionalities to enable interactive speech therapy.

8.5 Performance Optimization

Asynchronous processing and optimised audio handling reduce latency and ensure smooth real-time speech analysis.

8.5 AI integration

AI based speech recognition and phoneme analysis models detect misarticulation and provide instant corrective feedback.

8.6 Offline Accessibility and scalability

Offline support improves accessibility in low-connectivity areas. The modular architecture allows future expansion to additional languages and platforms.

9. FUTURE RESEARCH DIRECTIONS

Future research on Vaani Mitra will focus on expanding support to additional Indian languages to address linguistic diversity. The integration of advanced deep learning-based speech recognition models can further enhance misarticulation detection and personalized feedback. Large-scale clinical trials involving speech-language pathologists are required to validate long-term therapeutic effectiveness. Future versions may include a mobile application, therapist dashboards, and teletherapy integration to improve accessibility and professional supervision. Additionally, strengthening ethical AI practices, bias reduction, and accessibility features will support inclusive and responsible deployment.

10. CONCLUSION

“Vaani Mitra” is a new technology-driven speech therapy product for Hindi children that uses gamification and AI-driven feedback to maximize articulation and engagement. The study reports highly impressive gains in speech accuracy, user engagement, and accessibility and thus validates the approach.

Through the deployment of culturally transcribed therapy modules in a data-driven, scalable platform, the project sufficiently fills the gaps in available tools for speech therapy. Its user and clinician

satisfaction levels help prove its likelihood for broadscale deployment.

Despite the initial setbacks such as regional dialect variations and issues with digital literacy, incremental updates made the system adaptable. Future updates will involve the integration of more languages, the development of a cellphone application, and the integration of advanced AI models for speech correction.

This research highlights the importance of fusing technology with user-centered design in healthcare innovation. “Vanni Mitra” creates a new benchmark for digital speech therapy in India, opening the door to more inclusive, engaging and accessible therapeutic experience for children with speech impairments.

11. REFERENCES

1. A. A. A. M. A. R. Daoud MI, “A Game Based Rehabilitation System for Upper-Limb Cerebral Palsy: A Feasibility Study, ”*Sensors (Base1)*, vol.20, 2020.
2. D. A. L. A. A. L. S. R. E. J. B. A. M. P. LaPiana N, “Acceptability of a Mobile Phone-Based Augmented Reality Game for Rehabilitation of Patients With Upper Limb Deficits from Stroke: Case Study, ”*JMIR Rehabil Assist Technol*, vol. 7,2020.
3. M. B. B. R. A. H. Tohidast SA, “Provision of speech-language pathology services for the treatment of speech and language disorders in children during the COVID-19 pandemic: Problems, concerns, and solutions, ”*Int J Pediatr Otorhinolaryngol*, 2020.
4. Y. K. M. K. J. & S. T. J. Kim, “Smartphone-Based Speech Therapy for Poststroke Dysarthria: Pilot Randomized Controlled Trial Evaluating Efficacy and Feasibility, ”*Journal of Medical Internet Research*, p. E56417, 2024.
5. D. S. K. Himani Bansal, “Revised Picture Articulation Test in Hindi, ”*International Journal for Research in Applied Science & Engineering Technology*, 2023.
6. S. M. J. Y. Y. H. K. M. K. G. Kim SY. “Effect of Voice and Articulation Parameters of a Home-Based Serious Game for Speech Therapy in Children With Articulation Disorder: Prospective Single-Arm Clinical Trial, ”*JMIR Serious Games*, 2023.
7. B. H. S. M. G. M. Saeedi S, “Application of Digital Games for Speech Therapy in Children: A Systematic Review of Features and Challenges, ”*J Healthc Eng*, 2022.
8. B. M. B. L. S. H. R. T. Ganzeboom M, “A serious game for speech training in dysarthric speakers with Parkinson’s diseases: Exploring therapeutic efficacy and patient satisfaction, ”*Int J Lang Commun Disord*, pp. 808-821, 2022.
9. M. M. S. T. E. S. Furlong LM, “Treating Childhood Speech Sound Disorders: Current Approaches to Management by Australian Speech-Language Pathologists, ”*Lang Speech Hear Serv Sch*, 2021.
10. D. M. C. J. G. M. Law J, “Tele-practice for children and young people with communication disabilities: Employing the COM-B model to review the intervention literature and inform guidance for practitioners, ”*Int J Lang Commun Disord*, 2021.
11. J. J. H. K. Sinsky CA, “Telemedicine and Team-Based Care: The Perils and the Promise, ”*Mayo Clin Proc*, 2021.
12. K. M. N. & N. J. Khoza-Shangase, “Speech-Language Pathology and Audiology in South Africa: Clinical Training and Service in the Era of COVID-19, ”*International Journal of Telerehabilitation*, vol. 13,2021.
13. B. K. Johnsson G, “Practitioner and Service User Perspective on the Rapid Shift to Teletherapy for Individuals on the Autism Spectrum as a Result of COVID-19, ”*Int J Environ Res Public Health*, 2021.
14. B. W. Tamer Greenwell, “Evidence Based Practice in Speech-Language Pathology: Where Are We Now?, ”*American Journal of Speech-Language Pathology*, vol. 30, pp. 186-198, 2021.
15. E. M. D. A. R. J. Simacek J, “Current Trends in Telehealth Applications to Deliver Social Communication Interventions for Young Children with or at Risk for Autism Spectrum Disorder, ”*Curr Dev Disorder Rep*, 2021.
16. C. R. A. W. M. Gibbs V, “Autism assessment via telehealth during the Covid 19 pandemic: Experiences and perspective of autistic adults, parents/carers and clinicians, ”*Res Autism Spectr Disord*, 2021.
17. A. B. K. J. M. C. M. P. M. J. A. B. & G.-O. R. Hair, “A longitudinal Evaluation of Tablet-Based Child Speech Therapy with Apraxia World, ”*ACM Transactions on Accessible Computing*, vol. 14, no. 1, pp. 1-26, 2021.
18. A. B. K. J. M. C. M. P. M. J. A. B. & G.-O. R. Hair, “ A Longitudinal Evaluation of Tablet-Based Child Speech Therapy with Apraxia World, ”*ACM Transactions on Accessible Computing*, pp. 1-26, 2021.
19. M. A. D. K. Kraljevic JK, “Telepractice as a Reaction to the COVID-19 Crisis: Insights from Croatian SLP Settings, ”*Int J Telerehabil*, 2020.
20. L. I. F. D. C. A. K. V. C. I. Koutsiana E, “Serious Gaming Technology in Upper Extremity Rehabilitation: Scoping Review, ”*JMIR Serious Games*, 2020.