International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4171- 4175
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0540

# Audio Based Screening Tool to Identify Speech Sound Disorder – Kannada

*Dr. Guru R [1], Dr. Anusuya M A [2], Sheethal K V [3], Lakshmi P [4], Thejaswini U S [5], Bharath K M [6]*
*[1,2]Professor, Computer Science and Engineering, JSS Science and Technology University, Mysuru, Karnataka, 570006, India*
*[3,4,5,6] UG - Computer Science and Engineering, JSS Science and Technology University, Mysuru, Karnataka, 570006, India*
***Emails:*** *guruirg@sjce.ac.in[1], anusuya_ma@sjce.ac.in[2], sheethalkv24@gmail.com[3], lakshmijpb@gmail.com[4], thejaswinius67@gmail.com[5], bharathkm1312@gmail.com[6]*

## Abstract
*India's linguistic diversity highlights the need for language-specific tools for the accurate identification of communication disorders. Speech Sound Disorders (SSD) are common in early childhood and, if not detected early, can lead to long-term academic and social difficulties. However, standardized screening systems for regional languages such as Kannada are limited. This study proposes a web-based Kannada speech screening system for early identification of SSD among children aged 2–7 years in Karnataka. The system is based on the SODA (Substitution, Omission, Distortion, Addition) framework to systematically analyze articulation errors. It incorporates browser-based audio recording, speech-to-text processing, and phonetic transcription using the International Phonetic Alphabet (IPA), followed by rule-based phoneme comparison to detect deviations in speech production. The system generates structured and interpret-able reports highlighting error patterns to support clinical assessment and early intervention. Experimental findings indicate that substitution and omission are the most frequent error types, aligning with typical speech development patterns. The proposed approach improves screening efficiency, objectivity, and accessibility, providing a scalable and practical solution for early detection of SSD in Kannada-speaking children.*

***Keywords:*** *Early detection; Kannada language; Speech sound disorders; Speech screening; Web-based system*

## 1. Introduction

Speech sound disorder (SSD) is a communication disorder characterized by difficulty in the accurate production and use of speech sounds, resulting in reduced speech intelligibility. Early identification and timely intervention are important to minimize these effects and support normal language growth. Many children face difficulties in pronouncing certain sounds during early childhood. If these speech challenges are unaddressed, they can affect communication, learning, and self-confidence. Families often struggle to access professional speech therapy because of limited specialists, high costs, and geographical barriers. Conventional assessment methods are predominantly manual, time-intensive, and dependent on expert clinicians, making them less accessible in rural and resource-limited settings. This challenge is particularly evident in the context of regional languages such as Kannada, where there is a lack of standardized, technology-driven screening tools tailored to native speakers. Existing approaches often do not incorporate automated speech analysis or structured error classification frameworks, limiting their scalability and objectivity. Mobile health tools such as apps can improve access profession guidance by cost-effective ways to support therapy at home. Children often practice more interactive with computer-based tools than traditional methods, allowing families from all backgrounds to participate, regardless of location or income. Moreover, It supports early identification and active participation of younger children in learning speech sounds playfully at home or school. But most existing speech screening and assessment tools are developed for high-resource languages such as English and are not directly applicable to Kannada which is a basic and versatile south Indian regional language. Several studies have attempted to address SSD assessment using both clinical and

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4171- 4175
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0540

computational approaches. The Computer-Based Assessment of Phonological Processes in Kannada (CAPP-K) demonstrates structured phonological analysis with high sensitivity but relies on clinician intervention, making it semi-automated (Sreedevi et al., 2014; Sreedevi et al., 2024). Multilingual articulation screening methods provide phonetically rich word lists but remain manual and lack computational support (Prabhu et al., 2018). Computational architectures for SSD screening have been proposed to improve efficiency; however, many remain conceptual without practical validation or language-specific adaptation (Rajan, P, 2023). Additionally, phoneme-based speech recognition research contributes to automatic processing but does not address clinical error classification or disorder detection (Birari, H et al., 2023). To address these limitations, this study focuses on Kannada speaking children and proposes a web-based speech articulation screening tool at the phoneme level for the small set of words that leverages the SODA (Substitution, Omission, Distortion, Addition) error classification method for systematic analysis of articulatory errors [1-6]. The proposed tool aims to facilitate early screening, automated report generation and support data-driven speech assessment for Kannada regional language contexts.

### 1.1. Data Set

The data set used in this study consists of age-specific Kannada word lists designed for speech screening across three age groups: 2–4 years, 5–6 years, and 6–7 years. Each group includes carefully selected words mapped to specific target phonemes to ensure coverage of developmentally appropriate speech sounds. The progression across these age groups reflects increasing phonetic complexity, enabling systematic assessment of articulatory development. The selection of words is based on commonly observed Speech Sound Disorders (SSD) in children. From observations and analysis, it was identified that although most children were able to correctly articulate several phonemes, they exhibited difficulty in producing certain specific phonemes. These partially misarticulated phonemes are included in the data set to facilitate effective screening and early identification of speech sound errors. The data set, consisting of both words and their corresponding phonemes, is presented in Table 1.

#### 1.1.1. Region Information

The data for this study was collected from K.R. Pete (Mandya), K.R. Nagar (Mysuru), and Agrahara (Chamarajanagar) in Karnataka. The participants included both male and female children across all age groups (presented in Table II), ensuring balanced gender representation. These regions represent semi-urban and rural populations, providing diverse linguistic exposure and improving the robustness and generalizability of the proposed screening tool.

**Table 1 Data set considered for the screening tool consist of phoneme**

| Phonemes | | |
|---|---|---|
| Age group (2-4) | Age group (5-6) | Age group (6-7) |
| ಅ | ಕ | ರ |
| ಆ | ಗ | ಣ |
| ಇ | ತ | ಸ |
| ಈ | ದ | ಡ |
| ಉ | ನ | ಳ |
| ಊ | ಬ | ಶ |
| ಎ | ಪ | |
| ಒ | ಮ | |
| ಓ | ವ | |
| | ಲ | |
| | ಚ | |
| | ಜ | |
| | ಟ | |

## 2. Method

A total of 45 children were included in the study, equally distributed across three age groups (2–4, 5–6, and 6–7 years), with 15 children in each group and a balanced representation of male (21) and female (24) participants. These age groups were selected as this developmental stage is critical for the emergence of speech sound disorders (SSD). Participants were chosen based on specific selection criteria, including no reported difficulties related to speech, communication, hearing, neurological conditions, or behavioral concerns.

**Table 2 Demographic Distribution of Participants by Age and Gender**

| Age group | Number of children tested | Female count | Male count |
|---|---|---|---|
| (2-4) | 15 | 9 | 6 |
| (5-6) | 15 | 8 | 7 |
| (6-7) | 15 | 7 | 8 |

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4171- 4175
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0540

The formulas used for age-wise analysis are as follows:

$$P=\frac{C}{N} * 100 \qquad \text{Equation 1}$$

- P = Percentage of correctness of word
- C = Number of children who pronounced the word correctly (varies for each word and age group)
- N = Total number of children tested (constant within each age group; here N=15)

Formula used to calculate percentage of error occurred for the phoneme (age wise analysis) as below

$$Percentage\ of\ Error\ (Phoneme) = \frac{\sum(100-Ci)}{N} * 100$$

Equation 2

- Ci = Percentage of correctness of each word containing the phoneme (within the age group)
- N = Total number of words containing that phoneme in the given age group
- ∑(100−Ci) = Total error contribution from all such words

## 2.1. Implementation Details

The complete system architecture for articulation and SODA error detection is processed at three levels: Frontend, Backend, and Analysis (see Fig. 1).

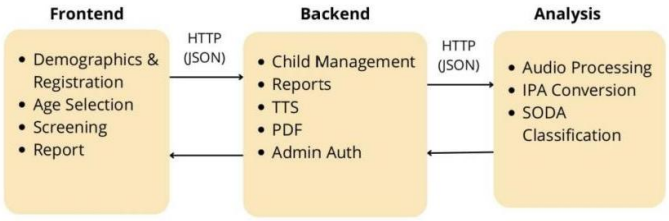


**Figure 1 System Architecture of the Proposed Kannada Speech Screening Tool (TTS - Text to Speech)**

### 2.1.1. Technology Stack used:

- Frontend: HTML/CSS/JavaScript (Kannada UI)
- Server: Node.js + Express REST API
- Analysis: Python + Flask modules
- Database: MongoDB

## 2.2. Frontend Framework

The system follows a structured workflow to support early identification. Initially, it collects demographic details and generates a unique Child ID. Based on the child's age, the system presents age-specific screening tasks using images and audio using Text-to-Speech (TTS) prompts. The child's speech is recorded and processed, and results are displayed via a dashboard with pie charts and improvement suggestions. An admin module allows authorized users to manage data efficiently. The complete workflow is shown in Fig. 2.

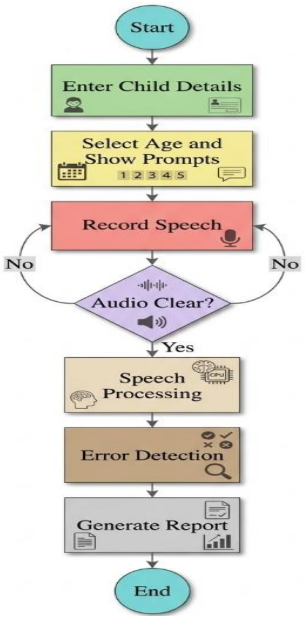


**Figure 2 Overall Workflow of the Proposed Speech Screening System**

## 2.3. Backend Framework

The backend architecture (include python analysis) serves as the central processing engine for the speech screening tool. It is designed to handle high-concurrency requests and perform resource-intensive phonetic analysis, structured across ten distinct functional areas:

- Audio Handling: Audio files are received via HTTP POST requests, with Multer

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4171- 4175
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0540

middleware enabling efficient in-memory multipart/form-data handling.

- Speech-to-Text: Uploaded audio is processed via Python based engines, using speech recognition and FFmpeg for format conversion.
- Phonetic Transcription: Text is converted to IPA using deterministic mapping rules [6], ensuring accurate phonetic representation of target and spoken words.
- Phonetic Comparison: Strings are compared using sequence alignment algorithms to determine pronunciation accuracy.
- Syllable Segmentation: Spoken and target IPA strings are segmented into syllables (aksharas) using language specific rules.
- Error Classification (SODA): Errors are categorized into Substitution, Omission, Distortion, and Addition using custom scripts. The procedure is visualized in Fig. 5.
- Data Management: MongoDB/Mongoose stores child records, while JWT handles secure session management.
- Report Generation: PDFKit is utilized to generate professional PDF summaries of SODA analysis.
- Cloud /Local Storage: Supports flexible deployment via Azure Blob Storage or local filesystem.
- API Communication: Axios facilitates RESTful communication between the Node.js server and Python analysis modules.

## 3. Results and Discussion

A total of 45 children were evaluated across three age groups (2–4, 5–6, and 6–7 years). The system computed word-level accuracy and phoneme-level errors using the SODA framework. For the 2–4 years age group, the results indicate a high level of word-level pronunciation accuracy with most words achieving above 80% correctness (see in Fig 6). For the 5–6 years age group, the results show a moderate to high level of pronunciation accuracy, with several words achieving above 80% correctness, indicating improved articulatory control compared to younger children. However, variability across words suggests that phonological development is still in progress, particularly for more complex or less frequently used sounds. For the 6–7 years age group, the results indicate a moderate to high level of speech accuracy, with most words achieving above 80% correctness, reflecting near-mature speech production. Suggesting that some phonemes still pose challenges. Across all age groups, substitution and omission errors were the most frequently observed. Some phonemes, particularly those requiring complex tongue movements (such as retroflex and fricative sounds), remained challenging even in older children.

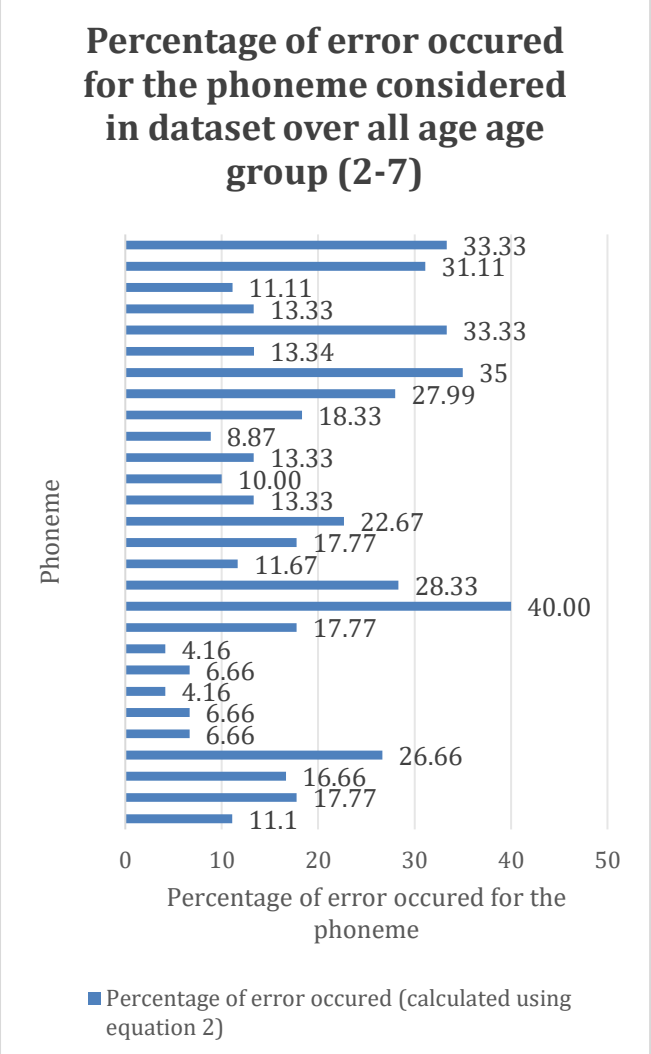


**Figure 6 Percentage of error occurred for the phoneme consider in age group (2-7)**

### 3.1. Discussion

- For age group 2-4 : Phoneme /i:/ considered difficult in this age.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4171- 4175
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0540

- For age group 5-6 : Phoneme-level analysis reveals that higher error rates are observed for consonants such as /ga/ (40%), /ʈa/ (35%), /ta/ (28.33%), /ja/ (27.99%) and /ba/ (22.67%) which typically require greater articulatory precision involving velar and retroflex placements.
- For age group 6-7 : Phoneme-level analysis highlights that /ɳʌ/ (33.33%), /ɭʌ/ (31.11%), and /ʃʌ/ (33.33%) exhibit higher error rates, which can be attributed to the articulatory complexity of retroflex and fricative sounds in Kannada.

## Conclusion

This study presents a web-based Kannada Speech Screening System designed for the early identification of Speech Sound Disorders (SSD) in children using a structured SODA-based classification approach. The system integrates speech acquisition, speech-to-text conversion, IPA-based phonetic analysis, and phoneme- and syllable-level comparison to automatically detect and classify pronunciation errors. Experimental evaluation across different age groups demonstrates that the system effectively captures variations in articulation accuracy and highlights phonemes that are developmentally challenging, particularly those requiring complex articulatory control. The generated reports, along with visual summaries, provide meaningful insights that can assist clinicians and caregivers in early intervention and monitoring. Overall, the proposed system offers a scalable, user-friendly, and cost-effective solution for preliminary speech screening in regional languages like Kannada.

## References

[1]. Sreedevi, N., & Chandran, S. (2014). Computer based assessment of phonological processes in Kannada (CAPP-K). All India Institute of Speech and Hearing, Mysore.

[2]. Soundararajan, A. (2018). A computer-based speech sound disorder screening system architecture. Retrieved from https://pubmed.ncbi.nlm.nih.gov/29968596/

[3]. Prabhu, S., Meghana, A. K., & Subhash, D. (2024). Development of a screening articulation test common to multiple languages. Mangalore University.

[4]. Sreedevi, N., Mathew, A., & Reshma, O. (2024). Articulatory error analysis in Kannada speaking children with hearing impairment: An exploratory study. Research & Reviews: A Journal of Health Professions, 14(1), 47–57.

[5]. Praveen, P. N., & Kini, S. Phoneme-based Kannada speech corpus for automatic speech recognition system. St. Aloysius College and Srinivas Institute of Technology, Mangaluru.

[6]. Arulalan, A. Kannada text to IPA converter [GitHub repository]. Retrieved from https://github.com/arulalant/txt2ipa