



Original Research Paper

TOWARDS INTELLIGENT THERAPY FOR SPEECH SOUND DISORDERS: A REVIEW OF ARTICULATION THERAPY APPS AND GPT-BASED ADVANCEMENTS

Keerthana M N¹, Pratiksha Barman¹, Mamatha S¹, Umme Hani¹, Vedavathi N¹ and Ayesha Taranum¹

¹Department of Computer Science and Engineering, Vidyavardhaka College of Engineering, Mysuru, India
vedavathi@vvce.ac.in

ABSTRACT

Speech Sound Disorder (SSD) is characterized by difficulties in articulation, leading to mispronunciations that impact clear communication. Early detection and customized therapy are essential for effective treatment. This paper explores multiple methodologies for SSD therapy support, with an emphasis on integrating Generative Pre-trained Transformers (GPT)-based models for phoneme-specific sentence generation. Using a GPT model, sentences containing user-specified phonemes were generated, ensuring grammatical accuracy to support targeted phonetic practice. Experiment results indicate that leveraging AI-based models, such as GPT for phoneme-focused generation of sentences, can significantly enhance therapy outcomes by providing tailored interventions, also reducing the Speech Language Pathologist (SLP's) workload. Our findings underscore the potential of the technology to support more accessible and effective SSD therapy, particularly with future expansions to include region-specific language support, to aid diverse linguistic populations.

Keywords: *Speech Sound Disorder (SSD), GPT-based Language Model, AI in Speech Therapy, Personalized Speech Intervention Tools.*

1 INTRODUCTION

Speech Sound Disorder (SSD) refers to the difficulty in articulation and production of speech sounds [1], impacting an individual's ability to communicate clearly. SSD is commonly characterized by substitution, omission, distortion, or addition (SODA) of sounds, making it challenging for individuals with SSD to express themselves. Globally, SSD affects approximately 10-15% of preschool-aged children. In India, a study conducted at Kasturba Medical College, Mangalore by Devadiga, Varghese, and Bhat (2014) found that 14% of individuals had speech disorders, with 48.4% of them related to articulation disorders [2]. People with SSD may struggle with social interactions as unclear communication can lead to misunderstandings, difficulty speaking with peers, and, in some cases, social isolation or bullying. In addition, they can suffer from reduced self-esteem because of consistent communication difficulties.

The effects of Speech Sound Disorder (SSD) extend to academic performance, especially in areas requiring phonetic awareness, such as reading [3] and writing. Studies have highlighted the importance of early

intervention to address these issues effectively. The fast progress in technologies and artificial intelligence (AI) has facilitated the creation of new computer-based solutions for treating SSD [4]. These digital tools offer structured and interactive therapy resources that support Speech-Language Pathologists (SLPs) and caregivers in providing articulation practice for targeted sounds.

This paper discusses the potential of software interventions for Speech Sound Disorder (SSD), focusing on the efficacy of AI-driven tools and computer-based therapy in improving phonetic accuracy and enhancing therapy outcomes. In addition, the paper presents an experiment designed to assess the utility of a GPT-based model [5] for sentence generation. The studies aim to demonstrate how software-based approaches can be augmented with traditional therapy methods [6-16] to make speech therapy more effective and personalized.

2 LITERATURE REVIEW

We conducted a comprehensive study to understand various software and AI-based automation tools that are developed to provide therapy for individuals affected with Speech Sound Disorder (SSD). The literature review conducted is given in the tabular format (see Table 1).

*Corresponding Author:

Received: 25th Feb, 2026; Revised: 6th June, 2026; Accepted: 8th July, 2026; Available Online: 14th Sep, 2026
(DOI): 10.70102/AEJ.2026.18.6s.31

Table 1. Summary of Literature Review.

Title	Features	Strengths	Limitations
Articulation Station[7]	1. It offers word and sentence-level practice to help individuals work on sounds in different contexts. 2. The app includes games targeting different sounds, making practice sessions more engaging and fun for young users. 3. It provides progress tracking with options to export data, helping therapists monitor individual improvements over time.	1. It provides structured practice across multiple levels. 2. The game section of the app can sustain kids engagement so that the learning process could be achieved with fun.	1. It does not support regional languages, limiting its usability to English speakers. 2. Accessing the full range of features often requires in-app purchases, which may be a barrier for some users.
Speech Tutor[8]	1. The app uses animated mouth movements to show the articulation of different sounds. 2. It also contains a recording and playback feature, where a person can test their actual speech against the model speech. 3. The app also provides visual guidance on tongue placement inside the mouth, which can improve accuracy in sound production.	1. The visual animations in Speech Tutor assist in explaining complex articulation movements. 2. The recording feature supports self-assessment by allowing users to compare their attempts with target sounds.	1. Its feedback is limited to self-assessment through recording, without interactive guidance from SLPs to correct mistakes in real time. 2. Younger users may find the app's interface challenging, as it requires navigation between several features to use effectively.
Boo Articulation Helper[9]	1. The Boo Articulation Helper provides animated speech production visuals whereby individuals may see how the sounds are pronounced. 2. The app covers basic English phonemes to encourage the development of basic articulation skills. 3. The app has an interface mostly suitable for early learners, as it can be simple and hence accessible.	1. Sound visualizations on gives visual insights into specific sound production. 2. For young children who only need basic articulation practice with lesser complex instructions, this app involves simplified interactions.	1. It is not useful for long-term usage which may limit its utility for regular practice. 2. It is useful only for the native speakers of English.
AI-Based Automated Speech Therapy Tools for Persons with Speech Sound Disorders: A Systematic Literature Review[10]	1. Database Search: The required relevant literature regarding automated speech therapy tools was gathered by performing searches through multiple databases-that is, sources like PubMed, IEEE Xplore, and Google Scholar. 2. Inclusion/Exclusion Criteria: The authors included AI-based tools targeting SSD that were published in peer reviewed journals only. 3. Data Extraction and Analysis: Relevant information was extracted and qualitatively	1. It explained how AI tools can tailor therapy sessions to meet individual patient needs, enhancing engagement and motivation during the therapeutic process.	1. Some AI systems reported accuracy issues, raising concerns about their reliability in clinical practice. Further validation and testing are required to ensure effectiveness.

	analyzed from the chosen studies based on effectiveness and usability.		
Opportunities and Challenges for AI Based Support for Speech-Language Pathologists[11]	1. The study conducted a survey with 105 SLPs and analyzed their social media discussions to understand their daily challenges, requirements, and constraints. 2. Areas where the SLPs require assistance from AI-based speech therapy tools are presented based on the survey findings.	1. AI-driven assessments provide consistent measurements and analyses, which reduce human error and subjective bias. This is especially valuable for SLPs, as AI can handle large datasets and identify patterns that might be missed in manual assessments.	1. The study is predominantly a US based participant pool, which may restrict the applicability of the findings to a more global audience. 2. AI-driven tools need huge volumes of patient data to be useful, and this is a concern for how sensitive information can be safely managed.
Articulatory Error Analysis in Kannada Speaking Children with Hearing Impairment: An Exploratory Study[12]	1. The study examined two groups of Kannada-speaking children: nine with hearing impairment (CwHI) and nine with typical hearing (CwNH), all of the same age. Findings revealed that children with hearing impairment frequently misarticulated certain sounds, particularly short vowels and consonants. In contrast, the children with normal hearing demonstrated no vowel errors. Among those with hearing impairment, fricative sounds were the most commonly affected	1. Identifies speech errors, providing a useful reference for speech-language pathologists. 2. Categorizes errors like substitutions and omissions based on where consonants appear in words.	1. There is a lack of research on speech error patterns across different Indian languages. 2. More studies are needed to understand the long-term effects of using digital hearing aids on speech development.

3 EXPERIMENT SETUP AND ANALYSIS

3.1 Building a GPT model

Generative Pre-trained Transformers (GPT) models have emerged as transformative tools in natural language processing (NLP), they can be used for text generation, summarization, and linguistic analysis tasks. By making use of large datasets and complex neural network architectures, GPT models can produce realistic, grammatically, and contextually coherent text. As such, the versatility with which GPT models can be fine-tuned to certain applications makes them highly valuable in those sectors, such as communication, education, and therapy.

To build the model, essential tools and libraries such as torch for model operations, transformers for GPT-2, and nltk for phoneme processing were imported. The CMU Pronouncing Dictionary is downloaded and loaded for phoneme analysis. A pretrained GPT-2 model and

tokenizer are then initialized which are used for sentence generation. To analyze phonemes in sentences, we convert each word in a sentence to its phonetic form using the CMU dictionary. A function is used to count the number of words containing a specified target phoneme in a sentence, helping to quantify how well the sentence meets the phoneme criterion. Sentences are generated using GPT-2. Then the generated sentence is checked for the target phoneme. If the target phoneme is present, the sentence is returned. Otherwise, the generation is repeated up to five times. For grammatical correctness, a function is defined using `language_tool_python`. This function searches for mistakes in the sentence making sure each produced sentence obeys normal grammar. The sentences are produced depending on the target phoneme and tested on its validity. When there are a greater number of words with the target phoneme and grammatically correct, accuracy will have high-end values.

Generated Sentence: Generate a sentence that includes words with the phoneme 'S', like 'she sat by the seashore'.

The sentence is then displayed in the browser.

The browser will then display the sentence in the browser.

Count of words containing target phoneme: 6

Grammatical correctness: Correct

Accuracy: 1.0

Fig. 1. Output of the GPT model.

The results from Fig.1 showed how effective the model was in producing sentences that satisfied the phoneme criteria going along with grammatical correctness. In a specific case, the model generated a sentence consisting of six words with the target phoneme, scoring hundred percent accuracy. This signifies the capability of such a methodology in generating formal linguistic outputs for specific phonetic requirements. The randomness in the context of generated sentences, however, meant that there was scope for improvement, especially in applications where higher levels of engagement were necessary, such as therapy sessions.

One of the issues with the model was that the sentences produced tended to have no meaningful or coherent context, rendering them random. In applications such as speech therapy, this randomness may make the exercises less effective and interesting. In order to make the sentences more interesting and relevant, the model needs to be enhanced to produce culturally relevant and contextually interesting content. This would not only enhance the therapeutic value but also make the practice sessions more interesting and immersive for users.

4 CONCLUSION AND FUTURE SCOPE

The paper explores the potential of AI-driven speech therapy tools in assisting individuals with Speech Sound Disorder (SSD). By leveraging advanced models, particularly GPT-based systems which generate phoneme-specific sentences, facilitates structured articulation practice. The findings indicate the potential of AI in enhancing therapy outcomes, although there are challenges in providing meaningful and contextually appropriate sentence generation.

The future scope aims at enhancing the capabilities of the platform by creating speech sound therapy specific to the Kannada language, thus catering to regional language speakers. This entails the development of an exhaustive corpus of Kannada words, which will be segmented into phonemes for targeted practice. This enhancement will act as a stepping stone for the inclusion of other regional languages, expanding the platform's scope and influence among various linguistic communities.

REFERENCES

1. Namasivayam, A.K., Coleman, D., O'Dwyer, A., van Lieshout, P.: Speech Sound Disorders in Children: An Articulatory Phonology Perspective. *Frontiers in Psychology* 10, 2998 (2020).
2. Pooja, C.K.: Development of Minimal Pair-Based Intervention Manual for Children with Speech Sound Errors in Kannada. M.S. thesis, University of Mysore, All India Institute of Speech and Hearing, Mysuru, India (2016), unpublished.
3. Tambyraja, S.R., Farquharson, K., Justice, L.: Reading Risk in Children With Speech Sound Disorder: Prevalence, Persistence, and Predictors. *Journal of Speech, Language, and Hearing Research* 63(11), 3714–3726 (2020).
4. Barletta, V., Calvano, M., Curci, A., Piccinno, A.: A Protocol to Assess Usability and Feasibility of e-SpeechT, a Web-based System Supporting Speech Therapies. In: *Proceedings of the 16th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2023) - HEALTHINF*, 546–553 (2023).
5. Zhang, Y., Li, Z., Zhang, J.: Towards the Use of Pretrained Language Model GPT-2 for Testing the Hypothesis of Communicative Efficiency in the Lexicon. In: *Proceedings of the 2021 International Conference on Asian Language Processing (IALP)*, 62–66 (2021).
6. Jesus, L.M.T., Martinez, J., Santos, J., Hall, A., Joffe, V.: Comparing Traditional and Tablet-Based Intervention for Children With Speech Sound Disorders: A Randomized Controlled Trial. *Journal of Speech, Language, and Hearing Research* 62(11), 4045–4061 (2019).
7. Articulation Station, https://littlebeespeech.com/articulation_station.php.
8. Speech Tutor, <https://www.speechtutor.org>.
9. Boo Articulation Helper App, <https://www.bildligttalat.nu/en/booen/>.

10. Suh, H., Dangol, A., Meadan, H., Miller, C.A., Kientz, J.A.: Opportunities and Challenges for AI-Based Support for Speech-Language Pathologists. CHIWORK 2024 - Proceedings of the 3rd Annual Meeting of the Symposium on Human-Computer Interaction for Work 2024, Article 14, 1-14 (2024).
11. Deka, C., Shrivastava, A., Abraham, A.K., Nautiyal, S., Chauhan, P.: AI-Based Automated Speech Therapy Tools for Persons with Speech Sound Disorders: A Systematic Literature Review. *Speech, Language, and Hearing* 28(1), (2022).
12. Sreedevi, N., Mathew, A., Reshma, O.: Articulatory Error Analysis in Kannada Speaking Children with Hearing Impairment: An Exploratory Study. *Research & Reviews: A Journal of Health Professions* 14(01), 47–57 (2024).
13. Sapsuha, Y., et al. (2024). Supplementation of chicken feed with nutmeg (*Myristica fragrans* Houtt.) extract and *Lactobacillus plantarum* improves broiler chicken performance. *International Journal of Probiotics & Prebiotics*, 19(1), 14–18. <https://doi.org/10.37290/ijpp2641-7197.19:14-18>
- [14] Pradhan, J., et al. (2024). Biosynthesis of secondary metabolites in aromatic and medicinal plants in response to abiotic stresses: A review. *Journal of Experimental Biology and Agricultural Sciences*, 12(3), 318–334. [http://doi.org/10.18006/2024.12\(3\).318.334](http://doi.org/10.18006/2024.12(3).318.334)
- [15] Souza, K. F., et al. (2026). A broad overview of the multifaceted royal jelly and its applicability in health and performance. *Current Topics in Nutraceutical Research*, 24(1). <http://www.nchpjournals.com/journal/index.php/ctnr/article/view/8>
- [16] Kumari, T., & Shukla, V. (2022). Comparative impact of fungicides on biomolecules of *Eisenia fetida*. *Bulletin of Pure and Applied Sciences-Zoology*, 41A(1), 1–13. <http://bpasjournals.com/zoology/index.php/journal/article/view/490>