Management of Speech Impairment Disorders in Aphasia Patients using Digital Intervention with Multilingual Regional Dialects

Shiza Hassan [1], Tanzeela Rind [1], Amirita Dewani [1*], Memoona Sami[1]
[1] Mehran University of Engineering and Technology, Pakistan.
* Correspondence: Amirita Dewani, Email ID: amirita@faculty.muet.edu.pk

Citation | Hassan. S, Rind. T, Dewani. A, Sami. M, "Management of Speech Impairment Disorders in Aphasia Patients using Digital Intervention with Multilingual Regional Dialects", IJIST, Vol. 07 Issue. 01 pp 111-128, January 2025
Received | Dec 23, 2024 **Revised |** Jan 15, 2025 **Accepted |** Jan 17, 2025 **Published |** Jan 18, 2025.

Abstract.
Speech is a zestful, and intricate activity that enables people to express ideas, emotions, and thoughts. We are able to render our views because of this neural activity. It is a significant process for learning and personal development that every individual deserves to develop, including those with special needs and who are on a journey to learn how to communicate. Children frequently suffer from speech disorders. This entails them at the risk of experiencing social, intellectual, and academic challenges that may persist and affect them in their adolescence and adulthood. In this context, we present a speech therapy solution for special children, to assist kids with speech and language impairments in improving their language skills. The proposed app can act as a useful management tool and rehabilitation system for people with aphasia disorder and their caretakers including parents, guardians, and teachers. This innovative app offers a vast number of features and practice sessions to develop language skills and overcome communication impairment problems. It also supports multiple regional languages, including English, Urdu, and Sindhi allowing users to switch between native languages effortlessly through the settings. The developed app is equipped with a dynamic accuracy assessment, and progress-tracking system, notifying the parents or guardians when practice sessions are missed, ensuring that language development remains consistent and effective. The major novelty of this work is that it has considered regional aphasia patients and their language needs. In contrast to the existing developed therapeutic tools which are mainly oriented towards resource-rich languages, the proposed work aims to address regional languages. The proposed speech therapist App for children can be a powerful tool for parents, caregivers, and educators, providing a fun and interactive way for children to improve their speech and language abilities. The developed solution also offers benefits in the context of enhanced patient involvement, motivation throughout their learning journey, greater flexibility, and accessibility in contrast to in-person therapy, immediate feedback, and careful progress monitoring that makes it easier to assess and modify treatment sessions.

Abbreviations.
Speech-Language Pathologists (SLP)
World Health Organization (WHO)
Modern Standard Arabic (MSA)
Neural Machine Translation (NMT)
Large language model (LLM)

Keywords: Speech Impairment; Speech Therapy; Aphasia Disorder; Mobile Application, Digital Intervention.

Introduction:

With the escalation of elderly people in developed countries and the prevalence of age-related chronic diseases, digital health may be crucial in assisting, managing, and controlling various aspects of human health. The use of mobile phone applications, PDAs, patient monitoring devices, and other mobile wireless technologies in healthcare are stated under the concept of mHealth by the WHO, which also acknowledges that such digital technologies can significantly lower the morbidity and premature mortality from noncommunicable diseases while also improving access to necessary healthcare services [1]. The notion of digital health encompasses a variety of areas such as electronic health (eHealth), mobile health (mHealth), wearable devices and sensors, wellness applications, quantified self, telemedicine, and electronic health records. mHealth serves as a way for the provision of healthcare services via the use of mobile devices and sensor networks. Applications in mHealth utilize mobile devices to gather healthcare information, transmit data to healthcare professionals, researchers, and patients, monitor vital signs in real time, and facilitate direct healthcare services [2]. The applicability of communication technologies and mobile computing to develop digital interventions in the fields of healthcare and public health has the potential to significantly enhance healthcare delivery systems and provide advantages to people [3]. Worldwide, mHealth technology is offering remarkable benefits to the public, including a more efficient healthcare system characterized by fewer errors, lower cost, easy accessibility, and enhanced patient engagement. In the recent era, patients are increasingly relying on such mHealth solutions for healthcare assistance [4].

Digital interventions and mobile applications are also becoming significant and practical tools for the provision of speech therapy services to people with speech impairments and aphasia disorders [5]. Healthcare models have incorporated self-management approaches to facilitate increased patient responsibility for chronic condition management [6]. Aphasia is a chronic condition that usually occurs due to damage in specific regions of the human brain and is an acquired language disorder. It is usually the consequence of an injury in the left hemisphere but may also arise from the injury or damage of the right hemisphere [7]. One of the first salient symptoms of this disorder from a clinical aspect is the presence of language impairment. Persistent communication problems ultimately lead to the development of anxiety and other psychiatric conditions [8]. Speech therapy is the primary treatment for individuals with aphasia, but access to therapy can be limited by a number of factors, including distance, cost, and availability of healthcare professionals [9]. Recent technological advancements have made smartphones an integral part of our daily lives due to their ability to be mobile and smaller in size and thus, if equipped with related applications, can be used by children and caretakers as a self-management tool to improve spoken language disorder in people with aphasia [10]. As the prevalence of aphasia is a serious healthcare concern, there is an increasing amount of research aimed toward developing technology-based interventions for persons with aphasia [11].

In line with this, the current study proposes a tool that serves as a self-administered training for individuals with aphasia to complement their speech therapy and targets a diverse group of people with native language support thereby overcoming the barriers associated with in-person therapy. The app is designed and developed specifically to assist kids who struggle with speech and language development issues. Acknowledging the significance of customized interventions, the app can facilitate parents or caretakers, teachers, and most significantly kids with a vast number of useful features and supported regional dialects. It is a comprehensive strategy designed to enable children to overcome speech difficulties and develop their communication skills. Driven by the necessity to cater to the specific requirements of kids with speech impairments, the developed app is an attempt to address the gap between traditional

approaches and the rapidly changing field of tech-based solutions. Instead of depending just on conventional therapeutic methods, the app uses mobile technology to give its young users individualized support and involvement. With the help of voiceovers generated for a variety of image data, this creative approach helps kids connect words to visuals and establish their vocabulary. Additionally, to address the cultural and linguistic diversities, the app is developed to support three languages including English, Urdu, and Sindhi. Hence, users can easily switch between several languages based on their learning needs and cultural backgrounds. Using innovative speech recognition technology, the app's sophisticated word accuracy assessment assesses pronunciation and offers dynamic progress tracking in multiple languages.

To guarantee constant language development, missed practice sessions are reported to parents and guardians via notifications. In addition, the program has a writing pad with immediate voice conversion functions for practicing spelling and written communication. With the help of this self-aphasia management and therapy tool, kids can participate in interactive learning activities that effectively help them improve their language skills. Through the provision of an extensive feature set customized to meet their requirements, the proposed app acts as a primary treatment and enables kids to recover their lost communication abilities. Moreover, it empowers parents, caregivers, and educators to actively participate in the child's language development journey.

In summary, this study makes the following contributions:

- Develops a captivating user interface customized for children, featuring interactive components and a visually appealing design that captures their attention and is user-friendly.
- Develops an extensive voiceover repository to support the linking of words and images, constructing a comprehensive resource that strengthens the associations between spoken language and visual representation.
- Implements innovative recognition technology to ensure precise evaluation of word pronunciation, ensuring an advanced and accurate evaluation process.
- Establish a responsive progress-tracking system to notify guardians about any overlooked practice session, ensuring timely notifications regarding the child's participation.
- Enables multilingual support including English, Urdu, and Sindhi language to accommodate the user's needs, ensuring broad accessibility for individuals with diverse linguistic preferences.

Undeniably, to serve the aim of developing speech therapist solutions for tackling speech impairment disorders, scholars worldwide have put their efforts, as detailed ahead in the literature review section. However, the existing body of work and the developed tools are mainly oriented towards resource-rich languages, more specifically addressing the English language. To the best of our knowledge and review of the existing literature, this is the first study that has put efforts into incorporating enhanced features for regional languages. This study demonstrates the significance of local languages and how AI-enabled mHealth solutions can successfully overcome the barriers of traditional in-person therapy for native users. It also highlights the value of such cost-effective digital interventions.

The remainder of the paper is structured as follows:

Section 2 sheds light on the recent literature available on digital tools used in speech therapy worldwide and states the research gaps. Section 3 presents the significance of in-home digital inventions and states the research problem. In Section 4, we have outlined the architecture and methodology employed in the research, detailing the modules and analysis procedures. Section 5 elaborates on methods and techniques used for the implementation of the proposed idea and provides functional implementation details. Further, we have also discussed implementation results depicting significant app functions and features. Section 6 details the

major beneficiaries of the proposed multilingual Speech Therapist system. The paper concludes in Section 7 with a summary of key findings, whereas Section 8 delves into the implications of the work, and extensions for future research.

Literature Review.

Speech is certainly a skill that can be enhanced through personalized therapy and practice. SLPs are key in ameliorating speech; however, almost 70% of SLPs have waiting lists of patients and this demonstrates interrupted access to therapeutic services. Consequently, many non-professional therapists are being trained by SLPs to provide speech therapy outside of the office. This is not a promising approach as it requires SLPs to spend time and effort educating a non-professional facilitator for performing therapy. Additionally, the therapy of an individual must depend on the schedule of the non-professionals who, unlike a professional therapist, may not deliver the speech curriculum and practices in an optimal manner. This has led scholars to the digitization and development of numerous mobile speech therapy systems with various features and capabilities [12].

As stated earlier, technology-oriented solutions and tools have become commonplace in the recent era and may significantly improve the quality of patient care and facilitate caregivers. This transformation in medical and healthcare has been accelerated more rapidly after the Covid-19 pandemic [13]**.** The vast majority of recent studies suggested that virtual speech therapy is a viable way to treat speech-language impediments in both adults and children [14]. In this context, speech therapy tools and interventions for aphasia patients have shown great promise and have recently drawn the attention of researchers worldwide. Some of the recent works and developments are discussed in this section.

Articulation Station offers six engaging activity types, and over 1,000 target words with high-quality images, customizable settings, and a record/playback feature for progress tracking in speech therapy [15]. This articulation platform uses diverse tools (Flashcards, Stories, Matching Games, and Unique Sentences) and customizable settings for an effective speech sound production methodology. Speech Tutor stands out with 132 detailed animations of mouth and tongue movements, providing users with a clear understanding of sound production through side and front view options, slow-motion breakdowns, and customizable settings [16]. Speech Tutor utilizes a visual approach, offering front and side views of the mouth to elucidate precise tongue placement and throat movements for various speech sounds, providing step-by-step breakdowns and customizable settings. However, it lacks an Android version and offers limited interactive activities, which may not engage young children effectively and lacks progress tracking or sharing features. The speech therapy game application developed by study in [17] focuses on rehabilitating language skills for aphasia patients through interactive gaming, providing a multimodal neurorehabilitation experience. Eight sessions of game-based speech therapy, lasting 10–15 minutes per day once a week for eight successive weeks were applied to seven aphasic patients. The therapy aimed to supplement traditional neuromuscular rehabilitation and utilized QAB-based assessments.

Speech Flipbook contributed the work in [18] provides comprehensive coverage of English single-syllable words with 125 phonemes. Customizable word lists are developed for tailored therapy. It comprises audio feedback, self-monitoring, and visual display options. A mobile health game called SpokeIt was proposed by the study conducted in [12] to aid with speech articulation therapy for speech improvements. The functional requirements were developed through semi-structured interviews with medical speech specialists. The speech recognition system in the game was developed using these functional requirements as a baseline. The application makes use of an offline critical speech recognition system that can provide real-time feedback to the users.

The work carried out by authors in [19] proposes GWAP, a computer-based game with a purpose, as a speech therapy tool for children who speak Egyptian and have Dyslalia. The app

was aimed to detect if a certain phoneme is spoken correctly. The Egyptian corpus was used to train a baseline acoustic model. MSA acoustic model was modified using the collected Egyptian corpus in order to take advantage of the abundance of MSA resources currently available.

Authors in [20] contributed a game-based application in which players can advance through stages after correctly pronouncing words. By incorporating entertainment elements into the program, authors hoped to improve results by allowing young children to play again. To keep the children motivated to learn while using the app, it emphasizes the enjoyment of a game and a feeling of accomplishment. The development phase of the app uses C++ as a programming language, Windows 10 operating system, Unity game engine, and Julius has been incorporated or voice analysis. The application can be operated with voice commands as it features voice recognition. From simple sounds to utterance correctness, this voice recognition detects whether the child is speaking and whether the utterance is accurate.

To assist Turkish-speaking kids with speech and language impairments, a 3D gaming environment was created by a study performed in [21]. The Unity 3D game engine, which is popular for 3D games, was used for development purposes. This allows the generated object to be used without the need for recompilation. The object is only required to be added to the database and its location inside the environment needs to be adjusted on the platform. Twenty-five different objects of various shapes were used in the existing environment, including rubbish, benches, swings, teeters, slides, balls, and boats.

The comparison of existing technology-assisted applications and therapeutic tools is also presented in Table 1.

Table 1. Comparison of existing technology-assisted applications and therapeutic tools.

Reference No.	Features Description	Supported Languages
[15]	Articulation Station offers six engaging activity types, over 1,000 target words with high-quality images, customizable settings, and a record/playback feature for progress tracking in speech therapy.	English, Spanish [22]
[17]	The study proposed a speech therapy game application that focuses on rehabilitating language skills for aphasia patients through interactive gaming and providing a multimodal neurorehabilitation experience. The therapy aimed to supplement traditional neuromuscular rehabilitation and utilized QAB-based assessments.	English
[23]	The study developed an assistive tool called MobaCare that ultimately assists children in developing speech and language skills via different practice sessions. The app improves communication development and acts as a supportive tool for speech therapists and children. The implementation of the app incorporates various intuitive learning activities for kids.	English
[12]	A mobile health game called SpokeIt aids speech articulation therapy for speech improvements. The functional requirements for this app were developed through semi-structured interviews with medical speech specialists. The speech recognition system in the game was developed using these functional requirements as a baseline. The application makes use of an offline critical speech recognition system that can provide real-time feedback to the users.	English

[20]	The study contributed a game-based application in which players can advance through stages after correctly pronouncing words. By incorporating entertainment elements into the program, authors hoped to improve results by allowing young children to play again. The application can be operated with voice commands as it features voice recognition. From simple sounds to utterance correctness, this voice recognition system detects whether the child is speaking and whether the utterance is accurate.	Japanese
[21]	This study assists Turkish-speaking kids with speech and language impairments by developing a 3D gaming environment. The Unity 3D game engine was used for development purposes. This allows the generated object to be used without the need for recompilation. The object is only required to be added to the database and its location inside the environment needs to be adjusted on the platform. Twenty-five different objects of various shapes were used in the existing environment, including rubbish, benches, swings, teeters, slides, balls, and boats.	Turkish
[24]	The study designed a 3D game named "Into the Forest" for children with speech disorders and hearing problems. The authors used Unity 3D and C# programming language for development. The game incorporates two phases. The first phase in this 3D game is the training phase which enhances learning by playing the game. The second stage evaluates and assesses the learning for which the authors have used a validation score ranging from 0 to 100 percent.	English
[25]	CineVox is an easy-to-use, entertaining, and visually stimulating app that enhances speech and vocalization. CineVox is a sound-responsive game for speech impediments. Because CineVox responds to any sound, users can also create stunning and entertaining visual effects with it. Using an external microphone or the built-in microphone, CineVox can be used to encourage kids to vocalize and generate sounds.	English

Though several scholars have put efforts into addressing the challenge of speech impairment disorder in patients with different diseases via virtual therapies, the existing literature reveals that these apps are limited with respect to the features, app accessibility, availability, and supported languages. Also, most of the existing work focuses on English language and adults with such disorders. Pakistan's linguistic patterns are diverse. Pakistan is geographically divided into five provinces, each of which has its own regional language. Urdu is recognized as the country's national language, and English is its official language. In this manner, the vast majority of the people of Pakistan are bilingual and speak two languages [26]. These cultural and linguistic diversities are quite challenging as most of the work in the state of the art has been accomplished for developed countries. This study puts an effort into designing and developing a prototype for articulation therapy in order to address speech impediment problems and assist children struggling with aphasia in different regional dialects.

Problem Statement: Early intervention in individuals with language disorders will prevent many problems in the future that can cause poor academic attainment, reduced quality of life

and human well-being, reduced employment opportunities, negative social consequences, social disconnectedness, and lack of independence. Therefore, facilitating individuals with improved treatment and presentation methods to perform speech therapy is immensely important. The conducted speech therapy sessions should be frequent, intensive, and individualized for speech therapy exercises to be effective. Physical referral to SLP is not feasible for numerous reasons such as far geographical locations, availability of the number of therapists, and economic barriers specifically for people living in rural and low-income areas. Additionally, cultural diversities and lack of comprehensive features limit access to existing available mHealth speech therapy solutions. Hence, it is of utmost significance that in-home digital inventions with available technology should be developed to address these challenges. The proposed smartphone-based virtual therapist supports multiple regional dialects and will not only aid aphasia patients but will also be a great assistance tool for caretakers, parents, and guardians.

Speech Therapist Modules and Architecture:

Previous studies in the literature have indicated that a significant number of caretakers of care recipient patients with speech and communication impairments experience difficulties from the caregiver burden [27]. These caretakers may include parents, guardians, teachers, and siblings of little children. The designed prototype is an intervention tool and assists both the patients and their caretakers. The overall architecture of the speech therapist system and its workflow mechanism is illustrated in Figure 1. The designed application prototype comprises two significant aspects: the Aphasia patient's aspect and the Caretaker or guardian aspect.

The diagram represents a mobile application designed specifically for Aphasia Patients, providing accessibility to their parents or guardians for better support and progress tracking. The app allows two-way communication in the form of request and response objects to both types of supported users. The application includes a comprehensive database named "Hugging Face" that powers its functionality. It supports three languages including English, Urdu, and Sindhi thereby allowing users to select their preferred language. The application is divided into three main components. The first modules allow aphasia patients to learn Words vocabulary, fluency, and sounds and provide therapy using images and voiceovers. In this module, words are displayed with accompanying images and voices, which can be easily switched between the selected languages. This helps patients connect visual, auditory, and linguistic elements for better understanding. The second module assists patients in developing an understanding of sentence structure and grammar while speaking to build language skills. This therapy part offers sentences with voice options and corresponding translations in the chosen language, aiding in speech comprehension and formation. The third module provides a customized notepad for communication and supporting language development via text-to-speech analysis. The notepad section enhances accessibility further by converting text to speech and vice versa. Whatever the user types can be read aloud by the application, and spoken words can be transcribed into text seamlessly. This feature promotes communication and supports language development for patients with aphasia. For the purpose of speech recognition and text translation, we integrated the mBart translation model. BART applies bi-directional and auto-regressive models and is a successor of Facebook's BART model.

It supports various languages for the purpose of neural machine translation [28]. The application also uses an API, called 'dart-levenshtein' to estimate and store the fluency of children's pronunciations. The 'dart-Levenshtein' computes the measurement of string similarity or distance, indexed on word pronunciation and phonetic resemblance. It calculates the degree of match between the typed and spoken words and encourages children's constant language progress through real-time corrections.

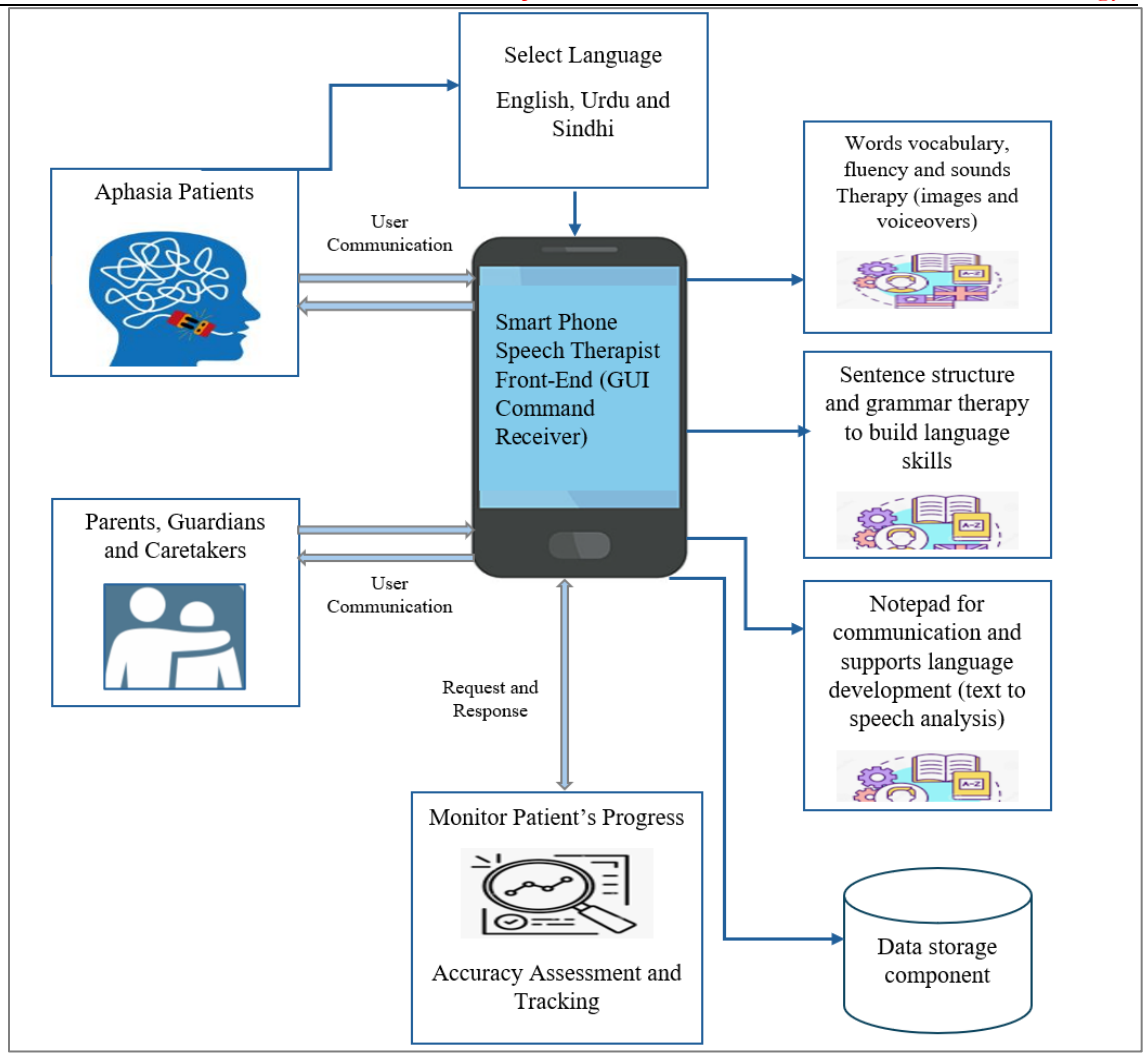


Figure 1. Virtual Speech Therapist for Aphasia Patients- Major Modules and Architecture.

The overall progress of a patient is monitored and validated based on the accuracy score of pronunciation and completion of the predefined goal of practice sessions set by the caretakers. In addition, the application includes an accuracy rate tracker, which assesses the clarity and correctness of the child's speech. This data, along with their progress, is stored in a database component and is accessible to parents or guardians, apart from patients, ensuring they can monitor improvements over time and may involve themselves in virtual speech therapy treatment of their aphasia patient. The detailed flowchart depicting the workflow mechanism and implemented backend technologies is depicted in Figure 2.

As an entrance point, the user must enter the details to complete the login process and specify his role in the articulation therapy. The caretaker of an aphasia patient is granted access rights to get involved in all the subsequent learning phases for language development. Caretakers can also monitor and track progress. For aphasia patients, it is significant to grant permission to access the microphone so that speech and text analysis tasks may be performed by the application with different implemented features. To accomplish this, the permission_handler package along with numerous functions has been used. Once permission is granted, the backend configuration of the app allows access to a vast number of features. To incorporate the support of articulation therapy in regional dialects, we have used Mbart translation model. The model supports various languages for the purpose of neural machine translation. Furthermore, to

integrate speech recognition features, we have implemented speech-to-text and text-to-speech analysis modules as shown in the flowchart in Figure 2.

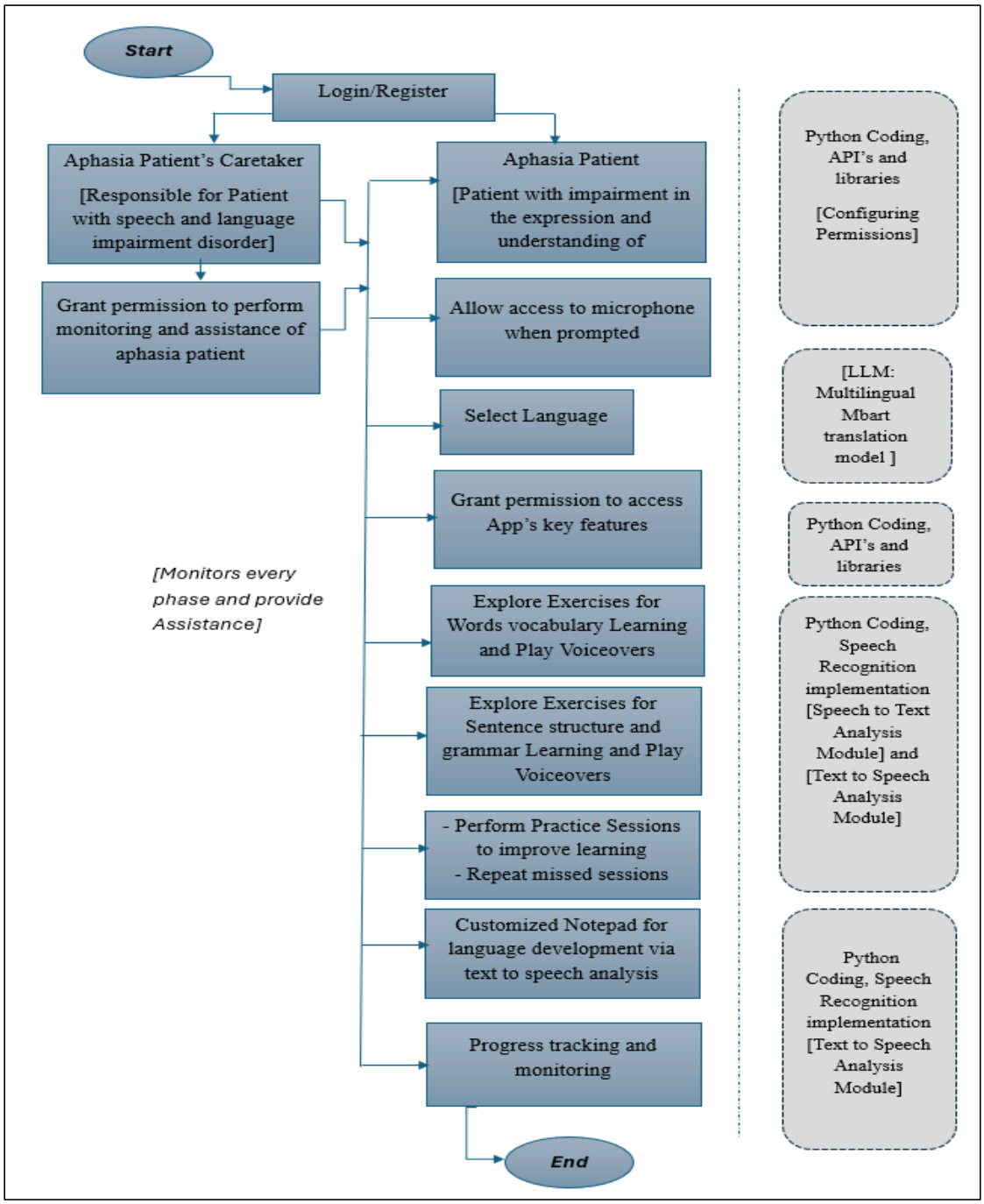


Figure 2. Detailed Flowchart for Virtual Speech Therapist.

These features are implemented via Flutter's speech-to-text and text-to-speech packages. These packages comprise several useful classes and functions that are implemented to embed speech and text recognition capabilities in the prototype. The analyzed speech recognition object is disposed of before the next input from the user arrives. Levenshtein distance implementation in Dart indicated by "dart-levenshtein" was used to support and quantify the level of match between the expected and expressed words. Before the usage of the app (pre-app usage) scenario indicates speech impairments in aphasia patients. Pronunciation of words in such cases leads to large distance values to the correct pronunciation. However, with the repetition of exercises and

practicing features supported by this app (post-app usage scenario), the overall `Levenshtein distance which is a metric for measuring the difference between two sequences is reduced thereby indicating improvement in speech and learning of patients. All the backend implementation is carried out using Python language, flutter framework, and various other methods detailed in the implementation section.

Speech Therapist Implementation:

The Speech Therapist App for Special Children is designed to help kids with speech impairments. This innovative tool connects words to visuals through voiceovers, enhancing vocabulary and language skills. The app offers multiple languages, including English, Urdu, and Sindhi, allowing easy language switching. Advanced speech recognition technology evaluates pronunciation, provides progress tracking, and sends notifications to parents about missed sessions to encourage regular practice. A built-in writing pad with voice conversion helps with spelling and written communication. The app's comprehensive features aim to bridge the gap between traditional therapy and tech-based solutions, empowering children to improve their communication skills easily and effectively. Firstly, we shed light on the techniques and tools used for the implementation of the proposed idea. Next, we provide functional implementation details.

Methods and Techniques:

Visual Studio Code:

Visual Studio Code (VS Code) [29] is a lightweight but powerful source code editor that runs on your desktop. It has built-in support for development operations like debugging, task running, and version control. In this work, we used VS Code for writing, debugging, and managing the code for both our Flutter and Python components. Its extensive extensions and support for multiple programming languages made it an ideal choice for developing the prototype of a virtual speech therapist.

Android Platform:

Android is a mobile operating system that was created by Google and is extensively utilized on a range of gadgets, such as tablets and smartphones [30]. It is available to a wide audience and provides a strong ecosystem with a variety of hardware options. Android Studio served as our main development environment for creating our speech treatment app for Android. With features like speech recognition and interactive activities to improve the therapeutic experience, the app was made with the user's experience in mind. To guarantee a flawless user experience, we concentrated on performance optimization across various Android devices and screen sizes. With Android's built-in accessibility features and notification systems, the study aimed to build a complete tool for people in need of speech therapy.

Flutter:

Flutter [31] is an open-source UI software development kit created by Google. It is used to develop cross-platform applications for Android, iOS, Linux, Mac, Windows, and the web from a single codebase. Flutter was used to build the front end of a speech therapist, providing a smooth and responsive user interface for the proposed app. Its cross-platform capabilities ensured that the app could run on multiple devices with minimal changes.

Python Language:

Python is a high-level, interpreted programming language known for its ease of use and versatility [32]. It supports multiple programming paradigms and has a vast ecosystem of libraries and frameworks. Python was chosen for the backend implementation of the proposed system

due to its simplicity and the availability of powerful libraries like Flask, transformers, and gTTS. It enables us to handle the server-side logic and integrate advanced AI models.

Flask Framework:

Flask [33] is a microweb framework written in Python. It is designed to be simple and easy to use, providing the essential tools to build web applications. Flask was used to create the backend API for this project. It permits developers to expose endpoints for translating text and converting it to speech, facilitating communication between the Flutter front-end and the Python backend.

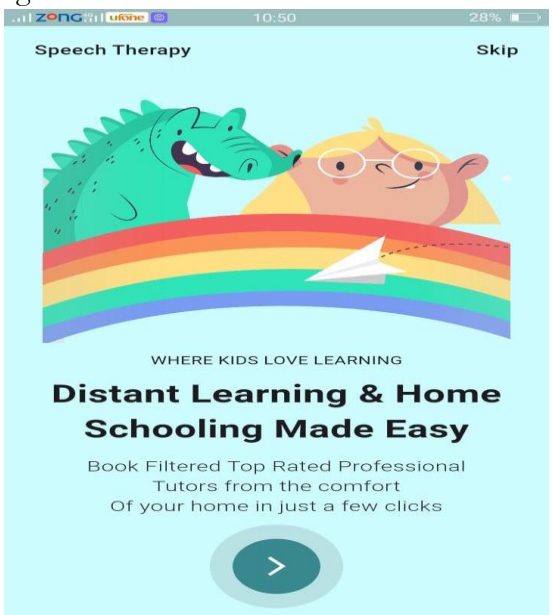
MBart Translational Model:

In today's interconnected society, the ability to efficiently process and comprehend content in several languages is becoming increasingly significant. Since natural languages are inherently complex, neural machine translation is an intricate and challenging task [28]. The majority of state-of-the-art models available are mostly trained on a limited number of commonly spoken languages, despite the recent advancements in NMT and regional languages worldwide. There has been a remarkable surge in the use of pertained models in vast practical NLP applications. These models are initially pre-trained on large datasets that include both monolingual and multilingual data and are then fine-tuned using a smaller, task-specific dataset. Recent research has shown the potential to develop architectures that can incorporate multiple regional languages. mBART50(multilingual Bidirectional Auto-Regressive Transformer) is one such model that was trained in fifty distinct languages. The Mbart Large 50 is a revolutionary model for multilingual machine translation because of its capacity to process many languages simultaneously [28]. It is a type of pre-trained LLM and is a crucial tool for many applications of artificial intelligence [34]. It is designed to handle translation tasks between multiple regional dialects [35]. The MBart model was used to perform the core functionality of this project which is translating text from English to other languages. Its pre-trained capabilities and multilingual support made it a feasible choice for the translation needs of the proposed virtual therapist.

Functional Implementation Results.

This proposed speech therapy digital prototype is expected to provide the regional users, a speech-language therapy where the progress made by the patients can be monitored throughout the sessions by caretakers. A caretaker must first register by completing the sign-up process in order to use the application. For the purpose of implementation, the Flutter Software Development Kit was used, since it is versatile enough to be used with both iOS and Android OS. The frontend of the mobile application was developed with Flutter and the IDE used was Visual Studio Code. A free UI template was adapted significantly for the needs of the proposed application. Some other incorporated elements included a picture-added voice-over with a picture-in-picture system, a dynamic progress bar, and interaction icons rated to assess the accuracy of children's pronunciation in real time. To determine fluency of pronunciation the application utilized the package called "dart-levenshtein". This package quantified the level of match between the expected and expressed words and was extremely useful in correcting the patient's language. The data for the user's progress, practice session details, and session misses were saved locally in a custom database to enable tracking, reporting, and triggering of alerts. The dataset was compiled for English, Urdu, and Sindhi, from various regional websites saturated in such content, thus covering diverse accents and dialects. This implementation setup made it possible for the app to provide the user with a constant experience of interactivity and improved therapy.

Next, for the integration of speech recognition into the system and text translation, we used the mBART translation model [36]. The model can translate directly between any pair of fifty languages. Pronunciation was handled by APIs that were used to parse and assess pronunciation in languages supported for this purpose, English, and two Pakistani dialects were tested with separate datasets to be input for Urdu and Sindhi. Several other APIs were used for accurate pronunciation analysis, and real-time feedback to the user. To reinforce usage, methods were adopted from gamification to encourage the user to continue to use the app. The significant screen interfaces of the proposed virtual speech therapist for aphasia patients are depicted in Figure 1.

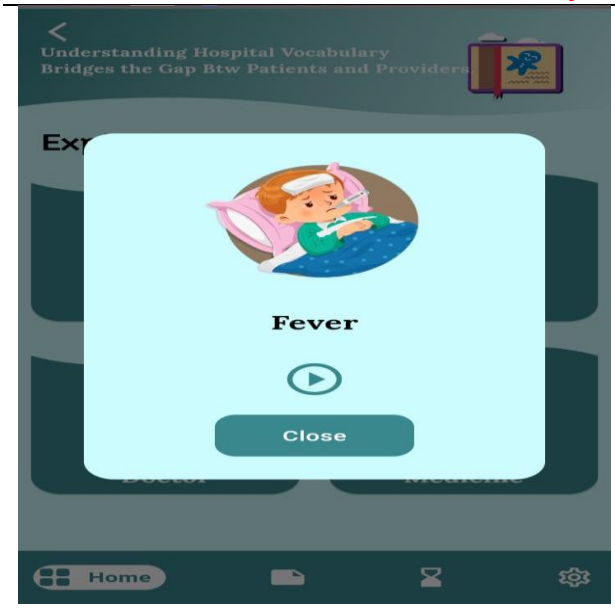


a. Onboarding Layout- App Entry Interface

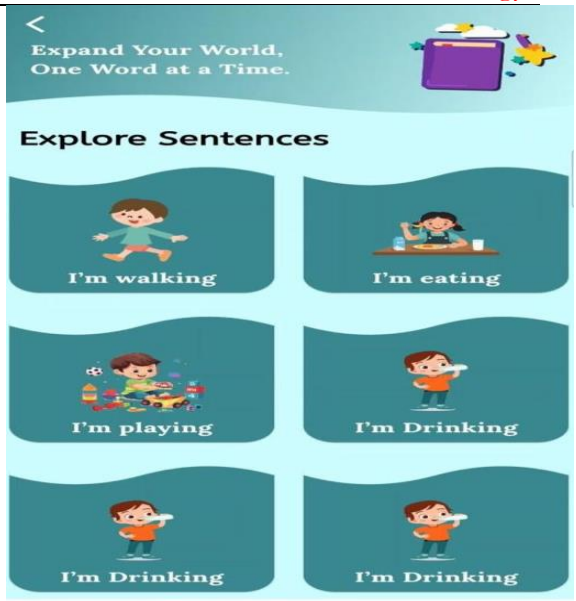
b. Learn Word vocabulary, fluency, and sounds using Speech Recognition

c. Subcategories Example for Words Vocabulary Learning- Explore Fruits

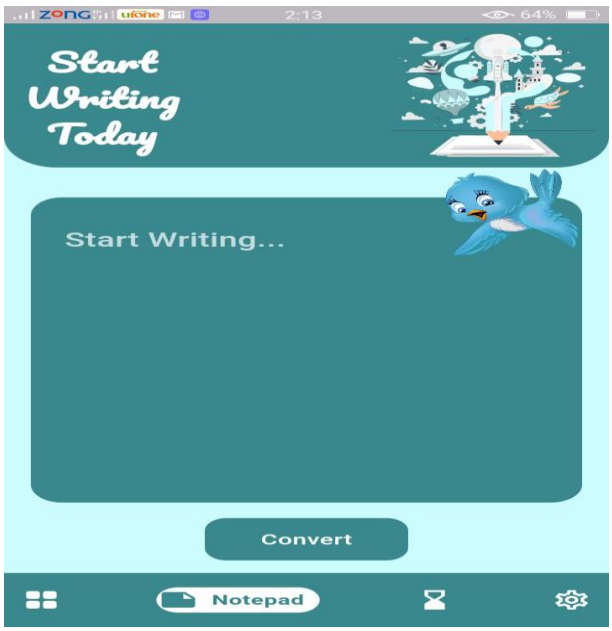
d. Subcategories Example for Words Vocabulary Learning - Explore Hospital



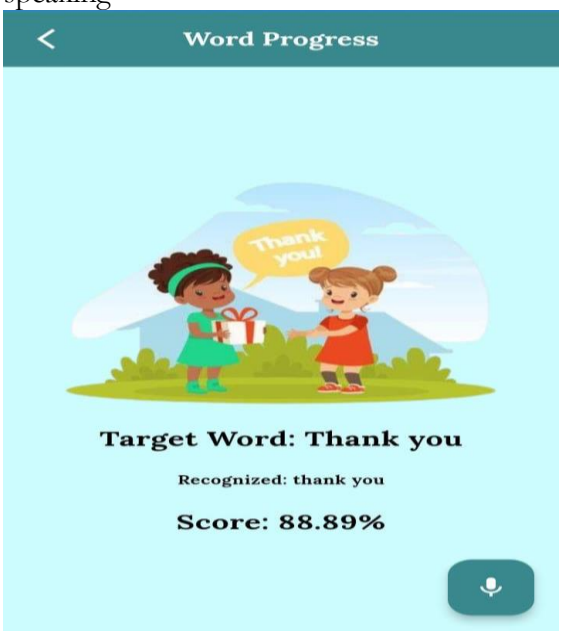
e. Voiceovers listening Interace for buidling vocabulary and producing correct sounds



f. Learn sentence structure and grammar while speaking



g. Customized Notepad for language development via text-to-speech analysis



h. Score monitoring for progress tracking

Figure 3. Significant Implementation Layouts for Speech Therapists with Multilingual Regional Dialects.

As shown in Figure 3(a), the Onboarding page of the Speech Therapy application provides a friendly and engaging interface. It introduces users to key features, emphasizing the ease of distance learning and home-schooling with access to top-rated professional tutors from home. The page is designed with child-friendly visuals to create a welcoming experience. The home page screen is displayed in Figure 3 (b) which offers a user-friendly interface designed for children with speech impairments. It greets the user by their name, creating a personalized experience. The top section features a search bar, allowing users to quickly search the specific content. Below the search bar, a "Top Categories" section provides access to various learning modules. Each category displays a daily learning goal, and words limit. The words Category page

from the first learning module is displayed in Figure 3 (c) where to excel vocabulary, fluency, and sounds there are different words available in the database and when a user clicks on any particular word, it will reply in voice over in English, Urdu, and Sindhi languages.

The vocabulary is categorized into different themes. A fruits category page is displayed in Figure 3 (d) there are multiple types of fruits available and when the user clicks on any fruit, it will respond in multilingual voiceovers in Urdu, Sindhi, and English. Similarly, the hospital Vocabulary page is displayed in Figure 3 (e) which depicts vocabulary and words associated with the hospital and when the user clicks on any particular word it will respond in multilingual voiceovers in Urdu, Sindhi, and English. The play word dialogue page is displayed in Figure 3 (f) when the user clicks on the play button it will respond in multilingual voiceovers Urdu, Sindhi, and English. These voiceovers assist patients in building vocabulary and producing correct sounds. Figure 3 (g) allows an aphasia patient to develop skills for Learning sentence structure and grammar while speaking. The customized notepad is shown in Figure 3 (h) which supports language development via text-to-speech analysis. This interface allows users to write any word, phrase, sentence, or paragraph and convert it into the corresponding voiceover. Finally, the implementation interface in Figure 3 (i) indicates the Score tracking and monitoring layout that allows for analysis of the accuracy of pronunciation and word utterance by the user. To accomplish this, we incorporated confidence level from the result class which is a part of speech to text model in the flutter framework.

Study Beneficiaries.

The Major beneficiaries of the proposed multilingual Speech Therapist system include children with speech impediments, their parents or guardians, speech therapists, educators, and caregivers involved in their language development. Further details are provided below:

- **Children with Speech Impairments:** The target audience benefiting from the proposed solution are children who have difficulties in speaking or comprehending language. Some of these children experience some difficulties in some areas of communication, and due to the distinctive features of the developed app, including voice-over, pronunciation check, and writing practice, they are able to improve their vocabulary and fluency of speech.
- **Parents and Guardians:** Parents and guardians are able to have an effective tool that supports their children's language learning at home. On the part of practice session reminders of the practices the child missed as well as progress reports and constant updates ensure that parents are in the know-how of what their child is learning and in what capacity they can assist.
- **Speech Therapists and Educators:** With this virtual speech therapist, education and speech therapy can make use of the innovative supplement to traditional speech therapy techniques. This tool increases their capacity to attend to each child and assess development without the limitations of face-to-face sessions or in-person therapies.
- **Schools and Special Education Centers:** Schools that primarily deal with children who are in special education can incorporate the app into the class so that special needs students can learn how to communicate from the app when doing lesson exercises and assignments.

Discussion and Conclusion.

As stated earlier, physical referral to speech-language pathologists is not feasible for numerous reasons. Additionally, the cultural diversities and lack of comprehensive features limit access to existing available mHealth speech therapy solutions. Hence, it is of

utmost significance that in-home digital inventions with available technology should be developed to address these challenges. To cater to the needs of real-world aphasia patients and assist them, this study was aimed at designing, developing, and implementing a specialized digital healthcare assistance tool for speech therapy. Aphasia disorder impacts the speech, comprehension, reading, and writing abilities of individuals. The app proposed in this work specifically addresses the needs of children with speech impediments, providing them with personalized support and resources to enhance their language development thereby eradicating the limitations of in-person therapy. As the major target audience for the Speech Therapist Mobile App includes children with speech impediments apart from their parents or guardians, speech therapists, and educators, hence app's design and interfaces are engaging, and easy to use.

The developed digital healthcare tool offers a voiceover feature for a diverse range of images, allowing children to associate words with visuals and improve their vocabulary. Additionally, to cater to the linguistic preferences of users, it supports multiple languages, including English, Urdu, and Sindhi. Utilizing state-of-the-art speech recognition technology, the system provides word pronunciation accuracy to provide real-time feedback. The patients can perform activities, practice sessions, and repeat the different vocabulary, which helps them to improve their accent and pronunciation. It also includes a progress tracking system to monitor users' performance and notify parents or guardians of missed practice sessions. Furthermore, a digital writing pad has also been integrated into the app, allowing children to practice spelling and written communication. It also has the provision of converting written text into speech in various languages.

The overall purpose of this app was to help educators, parents, and caregivers better assist the language development of children with special needs by providing them with useful resources. The proposed mHealth app is developed using the Flutter framework and managed into different modules, thereby ensuring easy scalability of the application to support more enhanced features including enhanced language support. In addition, such mHealth technological interventions can help lower the cost of treatment and make speech therapy more accessible to a wide range of native audiences. Despite the rising prevalence of digital devices and mHealth apps, one of the limitations of this work is that children from underprivileged areas, who lack access to the appropriate technology or do not possess basic technological literacy will not be able to use the proposed mHealth system for speech therapy.

Future Work Directions and Improvements.

In this study, we have put in detailed efforts and proposed a possible digital intervention for aphasia patients to address speech impairment disorders. Nevertheless, there is always room for improvement. One of the possible extensions of the existing work can be the incorporation of adaptive AI to enable a more customized learning experience. To guarantee a customized learning experience, this can entail modifying the level of exercises according to children's development and making unique practice session recommendations via a recommender system. A social interaction component might also be included in the proposed app, giving kids the opportunity to connect with their classmates in a secure setting and promoting social skills in addition to speech development. Moreover, additional languages and dialects may also be included in later versions to further improve accessibility and make the app more inclusive of users from a variety of linguistic backgrounds. Additionally, enabling children to have real-time video calls with speech therapists could offer them direct professional direction, ensuring that they obtain the specialized assistance they require. Young people may find learning more

engaging and inspiring if gamification components such as interactive storytelling exercises or accomplishment badges are included. Furthermore, Games and other technology interventions can also be incorporated in the future as adjunct features by SLPs to address speech impairments. As games are more appealing and entertaining in nature, in the rehabilitation process, serious games can be used to boost learning and motivation thereby making the therapy tool more engaging for children. We also suggest the integration of data analytic techniques which may allow caregivers and therapists to monitor developments more efficiently, predicting the areas that need work and modifying their teaching methods accordingly. Finally, usability testing and human-based evaluation on regional aphasia patients and their caretakers may be conducted in future extensions of this work which may provide detailed qualitative insights into the developed prototype.

Acknowledgment. We are extremely grateful to the Department of Software Engineering, Mehran University of Engineering and Technology, Jamshoro for all the technical support to accomplish this work.

Author's Contribution. Shiza Hassan. Writing – original draft, study conception, methodology, coding, implementation, and result reporting. Tanzeela Rind. Writing an original draft, methodology, and design. Amrita Dewani. Supervision, Methodology, draft writing –editing and review, validation, and result reporting. Memoona Sami. Supervision, draft writing–editing review, and validation.

Conflict of interest. The authors declare no conflict of interest for publishing this manuscript in IJIST.

Project details. This research study received no external funding. All the resources consumed were internally available.

References:

[1] A. Nguyen, V. Eschiti, T. C. Bui, Z. Nagykaldi, and K. Dwyer, "Mobile health interventions to improve health behaviors and healthcare services among Vietnamese individuals: a systematic review," in Healthcare, 2023, vol. 11, no. 9, p. 1225.

[2] M. Isakovic, J. Cijan, U. Sedlar, M. Volk, and J. Bester, "The Role of mHealth applications in societal and social challenges of the future," in 2015 12th International Conference on Information Technology-New Generations, 2015, pp. 561–566.

[3] Z. Hussein, "The advantages and disadvantages of the mHealth applications and the intention to use among smartphone users," Int. J. Mech. Eng. Technol., vol. 9, no. 12, pp. 943–947, 2018.

[4] A. P. S. Gagneja and K. K. Gagneja, "Mobile health (mHealth) technologies," in 2015 17th International Conference on E-health Networking, Application & Services (HealthCom), 2015, pp. 37–43.

[5] M. Masrom, N. M. Noor, A. I. Kamaruddin, and M. A. A. Aziz, "Speech Therapy Mobile Applications for People with Aphasia: PRISMA review and features analysis," in 2021 IEEE National Biomedical Engineering Conference (NBEC), 2021, pp. 77–81.

[6] L. Nichol, A. D. Rodriguez, R. Pitt, S. J. Wallace, and A. J. Hill, "'Self-management has to be the way of the future': Exploring the perspectives of speech-language pathologists who work with people with aphasia," Int. J. Speech. Lang. Pathol., vol. 25, no. 2, pp. 327–341, 2023.

[7] A. Rohde, L. Worrall, E. Godecke, R. O'Halloran, A. Farrell, and M. Massey, "Diagnosis of aphasia in stroke populations: A systematic review of language tests," PLoS One, vol. 13, no. 3, p. e0194143, 2018.

[8] S. Saeedi, H. Bouraghi, M.-S. Seifpanahi, and M. Ghazisaeedi, "Application of digital games for speech therapy in children: a systematic review of features and challenges," J. Healthc. Eng., vol. 2022, no. 1, p. 4814945, 2022.

[9] M. A. Aziz, N. M. Noor, S. Z. A. Jalil, and M. A. A. Aziz, "Identification of Best Classifier Method for Aphasia Classification," in 2023 IEEE 2nd National Biomedical Engineering Conference (NBEC), 2023, pp. 78–83.

[10] M. A. Aziz et al., "Development of Speech Therapy Mobile Application for Aphasia Patients," in 2021 IEEE National Biomedical Engineering Conference (NBEC), 2021, pp. 89–94.

[11] W. Riaz, G. Ali, M. Abid, I. N. Butt, A. Shahzad, and S. Shahid, "An Exploratory Study on Supporting Persons with Aphasia in Pakistan: Challenges and Opportunities," in Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility, 2020, pp. 1–4.

[12] J. Duval et al., "SpokeIt: building a mobile speech therapy experience," in Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services, 2018, pp. 1–12.

[13] C. J. T. Butcher and W. Hussain, "Digital healthcare: the future," Futur. Healthc. J., vol. 9, no. 2, pp. 113–117, 2022.

[14] S. A. S. Lee, "Virtual speech-language therapy for individuals with communication disorders: Current evidence, limitations, and benefits," Curr. Dev. Disord. Reports, vol. 6, pp. 119–125, 2019.

[15] A. Utepbayeva, N. Zhiyenbayeva, L. Assylbekova, and O. Tapalova, "Artificial Intelligence Applications (Fluency SIS, Articulation Station Pro, and Apraxia Farm) in the Psycholinguistic Development of Preschool Children with Speech Disorders," Int. J. Inf. Educ. Technol., vol. 14, no. 7, 2024.

[16] Y. Bai, "Pronunciation Tutor for Deaf Children based on ASR," Highlights Sci. Eng. Technol., vol. 24, pp. 119–124, 2022.

[17] B. Ahmed, P. Monroe, A. Hair, C. T. Tan, R. Gutierrez-Osuna, and K. J. Ballard, "Speech-driven mobile games for speech therapy: User experiences and feasibility," Int. J. Speech. Lang. Pathol., vol. 20, no. 6, pp. 644–658, 2018.

[18] M. Makhroji, R. Rahmiati, C. Chairuddin, and I. D. Isda, "Development of e-module based on flip book media to improve students' speaking skills," J. Kependidikan J. Has. Penelit. dan Kaji. Kepustakaan di Bid. Pendidikan, Pengajaran dan Pembelajaran, vol. 9, no. 4, pp. 1270–1279, 2023.

[19] R. Elhady, M. Elmahdy, I. Hamed, and S. Abdennadher, "A Game with a Purpose for Automatic Detection of Children's Speech Disabilities Using Limited Speech Resources," in Arabic Language Processing: From Theory to Practice: 6th International Conference, ICALP 2017, Fez, Morocco, October 11–12, 2017, Proceedings 6, 2018, pp. 79–89.

[20] H. Takagi et al., "Voice and Speech Training System for the Hearing-Impaired Children Using Tablet Terminal," in HCI International 2020-Posters: 22nd International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part III 22, 2020, pp. 121–127.

[21] M. Cagatay, P. Ege, G. Tokdemir, and N. E. Cagiltay, "A serious game for speech disorder children therapy," in 2012 7th international symposium on health informatics and bioinformatics, 2012, pp. 18–23.

[22] "Articulation Station Pro." https://apps.apple.com/us/app/articulation-station-pro-es/id1055553618 (accessed Jan. 02, 2025).

[23] C. O. Loyola, M. A. Diloy, and L. R. De Luna, "MobaCare: Development of Mobile Application as an Assistive Tool for Speech Therapy," in 2024 13th International Conference on Educational and Information Technology (ICEIT), 2024, pp. 433–438.

[24] N. Nasiri, S. Shirmohammadi, and A. Rashed, "A serious game for children with speech disorders and hearing problems," in 2017 IEEE 5th International Conference on Serious Games and Applications for Health (SeGAH), 2017, pp. 1–7.

[25] "Sensory CineVox - speech therapy for vocalising." https://apps.apple.com/us/app/sensory-cinevox-speech-therapy-for-vocalising/id673958511 (accessed Dec. 23, 2024).
[26] N. Ali and A. Khalid, "The Adaptation Of The Bilingual Aphasia Test In Urdu: Linguistic And Socio-Cultural Aspects," Migr. Lett., vol. 21, no. S11, pp. 642–669, 2024.
[27] M. Badesha, A. Thayanantha, S. Shune, and A. Namasivayam-MacDonald, "Caregiver burden interventions in speech–language pathology: A systematic review," Int. J. Lang. Commun. Disord., vol. 58, no. 4, pp. 1335–1356, 2023.
[28] V. Goyle, P. Krishnaswamy, K. G. Ravikumar, U. Chattopadhyay, and K. Goyle, "Neural machine translation for low resource languages," arXiv Prepr. arXiv2304.07869, 2023.
[29] "Visual Studio." https://visualstudio.microsoft.com/ (accessed Dec. 12, 2024).
[30] "Android platform." https://developer.android.com/guide/platform (accessed Nov. 13, 2024).
[31] "Flutter Framework." https://flutter.dev/ (accessed Oct. 12, 2024).
[32] "Pyhton Language." https://www.python.org/ (accessed Jul. 06, 2024).
[33] "Flask Documentation." https://flask.palletsprojects.com/en/stable/ (accessed Jul. 18, 2024).
[34] A. Navarro and F. Casacuberta, "Exploring Multilingual Pretrained Machine Translation Models for Interactive Translation," in Proceedings of Machine Translation Summit XIX, Vol. 2: Users Track, 2023, pp. 132–142.
[35] Y. Liu, "Multilingual denoising pre-training for neural machine translation," arXiv Prepr. arXiv2001.08210, 2020.
[36] "mBART-50 many to many multilingual machine translation." https://huggingface.co/facebook/mbart-large-50-many-to-many-mmt (accessed Sep. 08, 2024).