LANGUAGE, TECHNOLOGY, AND SOCIAL MEDIA

WISE Pendidikan Indonesia

Advances and Challenges in Artificial Intelligence and Machine Learning for Multilingual Speech Disorder Diagnosis and Therapy: An Integrative Review

Husnain Mukhdoom*✉ and S.R. Rajkumar✉

To link to this article:

📅	Published online: 31 August 2025
↗	Submit an article to this journal
🔖	View crossmark data
▶	Watch the video on YouTube

Check for updates

Research Article

Advances and Challenges in Artificial Intelligence and Machine Learning for Multilingual Speech Disorder Diagnosis and Therapy: An Integrative Review

Husnain Mukhdoom* and S.R. Rajkumar

Abstract

This study examines the role of Artificial Intelligence (AI) and Machine Learning (ML) in the diagnosis, assessment, and management of speech disorders, focusing on enhancing clinical practices. A comprehensive review of existing studies was conducted, highlighting the application of AI technologies such as Automated Speech Recognition (ASR), Natural Language Processing (NLP), and deep learning models, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). The findings reveal that AI and ML techniques have significantly improved diagnostic precision, therapeutic interventions, and clinical efficiency, especially in underserved populations. However, challenges related to limited multilingual datasets, model generalizability, and the interpretability of deep learning models were identified. Despite the promising advancements, challenges such as data privacy, model bias, and ethical concerns need to be addressed for broader clinical integration. This study contributes by synthesizing current AI and ML applications in speech disorder management, identifying key challenges, and proposing future directions, including the development of multilingual datasets, Explainable AI (XAI), and the integration of multimodal data to further enhance diagnostic and therapeutic outcomes.

Keywords: Artificial Intelligence; Machine Learning; Speech Disorders; Speech Therapy; Natural Language Processing; Deep Learning; Diagnostic Precision.

Publisher's Note:

Language, Technology, and Social Media

VOL.3 NO.2 2025

WISE Pendidikan Indonesia
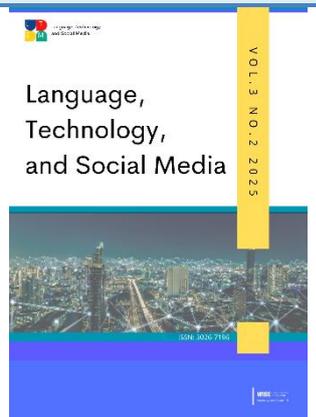
INTRODUCTION

Speech and language disorders affect millions of individuals globally, creating significant barriers to communication, education, and social inclusion [1], [2]. These conditions encompass articulation impairments, fluency disruptions, voice quality issues, and difficulties in language comprehension and production, which may result from neurological damage, developmental delays, or structural abnormalities in the speech apparatus [3]. Conventional speech-language pathology (SLP) relies on clinician-administered assessments, including perceptual analyses, standardized testing, and behavioral observations [4], [5]. While effective, such approaches are often time-intensive, subjective, and limited in accessibility particularly in rural or low-resource contexts [6], [7], [8].

Recent advances in Artificial Intelligence (AI) and Machine Learning (ML) have introduced transformative possibilities for augmenting SLP practices. Technologies such as Automated Speech Recognition (ASR), Natural Language Processing (NLP), and acoustic signal processing enable large-scale analysis of disordered speech to detect subtle phonetic, prosodic, and semantic anomalies [9]. Deep learning architectures, including Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, have demonstrated high diagnostic accuracy for conditions such as dysarthria, stuttering, apraxia of speech, and aphasia [10], [11]. AI-powered therapeutic tools ranging from intelligent tutoring systems to mobile-based assistive applications offer adaptive interventions, real-time feedback, and remote monitoring capabilities, thereby addressing service delivery gaps for underserved populations [12], [13].

Despite these advancements, several challenges hinder the effective integration of AI into clinical SLP workflows. These include the scarcity of large, annotated, disorder-specific speech datasets [14], difficulties in ensuring model generalizability across diverse linguistic and cultural contexts [15], and ethical concerns related to privacy, bias, and transparency [16]. Most AI models in this domain are trained predominantly on English-language corpora, which restricts applicability in multilingual and low-resource language settings [17]. Furthermore, the "black-box" nature of many deep learning systems limits interpretability, complicating clinical adoption and regulatory compliance [18].

However, existing review studies have largely focused on either the technical aspects of AI models or specific disorder categories, without providing an integrated perspective that connects computational methods, linguistic diversity, and clinical applicability. There remains a need for a systematic synthesis that critically examines current AI and ML approaches across disorder types, languages, and application contexts, while also addressing ethical, technical, and linguistic considerations in a unified manner. This study seeks to provide such a synthesis by mapping recent technological advancements, evaluating their methodological strengths and limitations, and proposing strategic directions for developing AI tools that are inclusive, accurate, and ethically aligned for speech and language disorder applications.

METHODS

This study adopts a literature review research design to systematically examine the application of AI and ML techniques in diagnosing, assessing, and assisting individuals with speech and language disorders. The methodological framework was developed to ensure comprehensive coverage of the field while maintaining rigorous inclusion standards to enhance the validity and reliability of the findings [19].

Literature Search Strategy

A comprehensive search was conducted across major academic databases, including PubMed, IEEE Xplore, Scopus, Web of Science, and Google Scholar, to identify peer-reviewed journal articles, conference proceedings, and review papers published between 2013 and 2025. Search terms were applied in Boolean combinations such as *"Artificial Intelligence" AND "Machine Learning" AND "Speech Disorders"*, *"Speech Impairment"*, *"Stuttering"*, *"Dysarthria"*, *"Apraxia"*, *"Natural Language Processing"*, *"Speech Recognition"*, *"Linguistic Analysis"*, *"Speech Therapy"*, *"Assistive Technology"*, and *"Clinical Linguistics"*.

Inclusion and Exclusion Criteria

The inclusion criteria encompassed studies that investigated the application of AI or ML in identifying, assessing, or treating speech and language disorders, incorporated linguistic features such as phonetic, prosodic, or semantic elements into AI models, were published in English between 2013 and 2025, and were based on clinical datasets or realistic experimental conditions. Exclusion criteria were applied to studies focusing solely on speech synthesis or text-to-speech systems unrelated to disorders, studies that did not employ AI or ML techniques, and studies without methodological clarity or accessible full text.

Data Extraction and Categorization

Data extraction involved recording details from each selected study, including the type of speech disorder under investigation, the AI or ML methods implemented, the linguistic features analyzed, the intended application type, and the population and language context. The collected studies were subsequently organized into thematic areas that covered diagnosis and classification systems, automatic speech recognition for disordered speech, natural language processing for linguistic profiling, multimodal and cross-lingual approaches, and integration into clinical workflows with considerations of usability.

Data Analysis

A thematic analysis approach was employed to identify recurring patterns, emerging technological trends, and the methodological strengths and weaknesses of the studies reviewed. The process involved iterative reading, coding of the extracted data, and synthesis into overarching themes that reflect the current state of AI and ML in speech-language pathology. The quality of each study was appraised using established tools, including QUADAS-2 for diagnostic accuracy studies, CONSORT-AI for clinical trials involving AI, and the Mixed Methods Appraisal Tool (MMAT) for mixed-methods research.

Ethical Considerations

As this review draws exclusively on previously published studies, there was no direct involvement of human participants or collection of new data. Nonetheless, all ethical guidelines for academic research were observed, including accurate attribution of sources, avoidance of plagiarism, and compliance with intellectual property rights.

RESULTS AND DISCUSSION

Diagnostic Applications of AI and ML in Speech Disorders

The integration of AI and ML in diagnosing speech disorders has substantially transformed the field of speech-language pathology (SLP). These technologies offer a more objective, scalable, and efficient diagnostic approach for conditions such as dysarthria, stuttering, apraxia, and aphasia. By analyzing various acoustic and prosodic features of speech, AI and ML models can detect subtle discrepancies that are often difficult for clinicians to identify manually, enabling earlier detection and more precise treatment interventions.

Key AI and ML Techniques for Diagnosing Speech Disorders

Several AI and ML techniques have been applied effectively to analyze speech patterns and diagnose speech disorders. Among these, CNNs, RNNs, and SVMs have shown notable success in the clinical context.

- CNNs are primarily used for analyzing acoustic features, detecting phonetic distortions, and identifying articulatory errors, which are key indicators in disorders like dysarthria. CNNs are particularly effective in distinguishing fine-grained variations in speech that result from neuromuscular impairments.
- RNNs are particularly well-suited for speech fluency disorders, such as stuttering. They excel at processing sequential data and are capable of analyzing temporal aspects of speech, such as rhythm, pauses, and rate of speech, which are disrupted in stuttering.
- SVMs are applied to analyze prosodic features in speech and have shown effectiveness in diagnosing apraxia of speech, which involves difficulty in coordinating the movements needed for speech production, often characterized by inconsistent errors in articulation.

These AI/ML techniques have demonstrated remarkable diagnostic accuracy, with some studies reporting accuracies exceeding 85% in identifying various speech disorders. The use of AI models significantly improves diagnostic efficiency, making them valuable tools in clinical settings where early and accurate diagnosis is crucial for timely intervention.

Performance and Diagnostic Accuracy

The performance of AI and ML models in speech disorder diagnosis has been promising, with studies showing high diagnostic accuracy. For instance, CNNs have been shown to effectively detect subtle articulation issues in dysarthria, while RNNs are adept at identifying the disruptions in fluency characteristic of stuttering. The high accuracy levels achieved by these models surpass those typically associated with human diagnosis, where subjectivity and time limitations often hinder precise results.

These advances are clearly demonstrated in Table 1, which summarizes various studies on the application of AI/ML in diagnosing speech disorders. The table details the AI methods used, the types of speech disorders targeted, and the sample sizes for each study, providing a clear picture of the diagnostic capabilities of these AI-driven tools.

Table 1. Studies on Speech Disorders Utilizing AI/ML Techniques with Quality Appraisal Scores Based on Tools Such as QUADAS-2, CONSORT-AI, and MMAT

Authors	Year	Speech Disorder	AI/ML Method	Sample Size	Study Design	Quality Appraisal Tool
Lin et al. [20]	2021	Dysarthria	CNN + Phonetic Posteriorgram (PPG)	120	Experimental	CONSORT-AI
Alharbi et al. [21]	2020	Stuttering	CRF + Bi-LSTM	85	Diagnostic Accuracy	QUADAS-2
Hsu et al. [22]	2024	Dysarthria	FEM + Diffusion Models	50	Experimental	MMAT
Themistocleous et al. [23]	2024	Autism Spectrum Disorder (ASD)	NLP + XGBoost 2.0	120	Diagnostic Accuracy	MMAT
Shih et al. [11]	2022	Dysarthria	CNN + GRU	160	Experimental	CONSORT-AI
Zheng et al. [24]	2022	Dysarthria	Voice Conversion + PPG	40	Experimental	MMAT
Fusaroli et al. [25]	2023	Autism Spectrum Disorder (ASD)	NLP	67	Longitudinal Study	MMAT

As demonstrated in Table 1, these studies leverage various AI/ML techniques, including CNNs, RNNs, Support Vector Machines (SVMs), and NLP, to diagnose a range of speech disorders. For example, Lin et al. [20] utilized CNNs with PPG features to diagnose dysarthria, achieving significant accuracy with 120 participants. Alharbi et al. [21] applied CRF and Bi-LSTM networks for stuttering detection, highlighting high diagnostic accuracy with 85 participants. Hsu et al. [22] introduced Fuzzy Expectation Maximization (FEM) integrated with diffusion models for dysarthria diagnosis, with 50 participants, demonstrating improvements in phoneme prediction accuracy and voice conversion. Themistocleous et al. [23] explored NLP models for diagnosing ASD, with a sample of 120 participants, while Fusaroli et al. [25] employed NLP for ASD, using a longitudinal study design with 67 participants, highlighting the predictive value of caregiver linguistic alignment in ASD. Shih et al. [11] utilized a hybrid CNN-GRU model for dysarthria diagnosis, showing high performance with 160 participants. Zheng et al. [24] focused on improving speech intelligibility in dysarthria through voice conversion combined with PPG, involving 40 participants. Collectively, these studies demonstrate the growing potential of AI/ML in providing more precise, efficient, and accessible diagnostics for speech disorders, surpassing traditional diagnostic methods.

Challenges and Limitations

Despite the promising results, there are several challenges associated with the use of AI/ML in speech disorder diagnosis. One of the primary limitations is the lack of diverse and comprehensive

datasets. Many current AI models are trained on datasets primarily consisting of English-language speech data, which limits their effectiveness in diagnosing disorders in non-English-speaking populations. Additionally, these models are often trained on relatively small sample sizes, which restricts their generalizability to larger and more diverse populations.

The studies highlighted in Table 1 reveal that the sample sizes used in these studies ranging from 50 to 200 participants are still relatively small, and often reflect a narrow demographic. This limits the ability of the models to account for the broad range of speech patterns found across different age groups, cultural contexts, and linguistic backgrounds. For instance, the study by Conti et al. [26] on childhood apraxia used an SVM model with prosodic features, but the sample was limited to a specific demographic, which may affect the model's ability to generalize to other populations.

Another challenge is the interpretability of deep learning models. While these models achieve high accuracy, they are often criticized for being "black-box" systems, making it difficult to understand how decisions are made. This lack of transparency can hinder the adoption of AI/ML tools in clinical settings, where clinicians need to trust and understand the reasoning behind diagnostic outputs. Explainable AI (XAI) frameworks are therefore crucial in making these models more interpretable and fostering trust among clinicians.

Future Directions

The future of AI and ML in speech disorder diagnosis lies in addressing these challenges. First, expanding the diversity of speech disorder datasets is essential for ensuring that AI models are applicable across various linguistic and cultural contexts. Multilingual datasets will be crucial for adapting these models to diagnose speech disorders in non-English-speaking populations.

In addition, developing XAI solutions will be important for increasing transparency and trust. Clinicians must be able to understand how an AI model arrives at its conclusions in order to integrate these tools effectively into clinical workflows. Future AI models should also focus on multimodal data integration, where visual cues such as lip movements and facial expressions are combined with acoustic data for more robust diagnostics.

AI-Powered Tools for Speech Therapy and Intervention

The integration of AI into speech therapy and intervention has not only enhanced diagnostic processes but also revolutionized therapeutic approaches. AI-powered tools are increasingly being used to deliver personalized, adaptive interventions that cater to the individual needs of patients with speech disorders. These tools have enabled real-time feedback, progress tracking, and remote therapy, addressing gaps in service delivery, particularly for underserved populations in rural or resource-constrained settings. By leveraging speech recognition, NLP, and reinforcement learning, these tools offer a more dynamic and responsive approach to therapy than traditional methods.

Adaptive Therapeutic Interventions

AI-based therapeutic tools are able to adjust the difficulty and type of exercises in response to a patient's progress, simulating the adaptive intervention strategies typically employed by human SLPs. One such method is reinforcement learning, where the system dynamically adapts exercises to match the patient's evolving needs. As illustrated in Figure 1, AI-powered systems track speech input from the patient, which is then processed using ASR technology. This process involves several stages: acoustic analysis, NLP processing, and ultimately, the provision of real-time

feedback to the patient. This feedback loop mimics the personalized and progressive nature of human-led therapy, adjusting the difficulty based on the patient's responses.
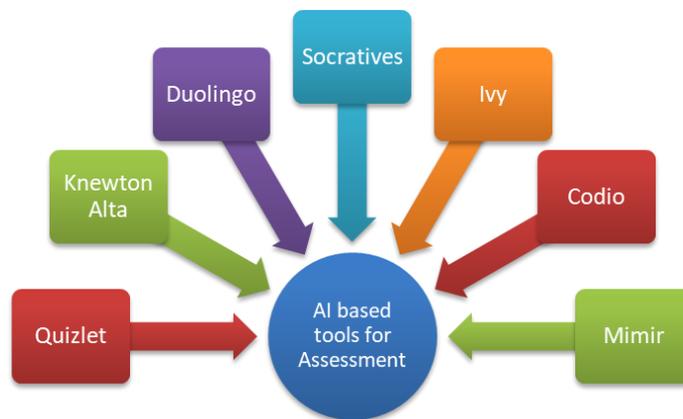

Figure 1. AI based tools for Assessment

Figure 1 offers a schematic overview of the AI-powered workflow used in speech therapy. The process begins with the patient's speech input, which is analyzed using ASR to detect phonetic, prosodic, and semantic discrepancies. The data is then processed through NLP algorithms to assess the quality of speech. Finally, the system provides real-time feedback to the patient. Additionally, the therapy system dynamically adjusts the difficulty level of subsequent exercises according to the patient's progress, ensuring personalized therapy that evolves with the patient's needs.

The feedback loop facilitated by this technology is key to providing interventions that can be personalized and adjusted based on each patient's ongoing progress. For example, a system could begin with basic phonetic articulation exercises for a patient with dysarthria, gradually increasing in complexity as the patient improves. As the patient advances, the system may introduce more complex speech tasks, such as sentence formation or exercises that target speech fluency, further tailoring the therapy to match their developing capabilities.

The ability of these systems to modify therapy dynamically based on real-time feedback is crucial for tracking patient progress and making timely adjustments to the therapy plan. This not only allows for better tracking of improvements but also ensures that the patient receives the most relevant and effective interventions as they advance through the therapy stages.

Furthermore, the incorporation of reinforcement learning ensures that therapy is not static. The system adjusts in real time, providing patients with exercises that continually challenge them, helping them to improve gradually while avoiding the frustration of tasks that are either too difficult or too easy. This personalized, adaptive approach enhances patient engagement and the effectiveness of therapy, leading to improved outcomes over time.

Applications of Natural Language Processing in Linguistic Profiling
NLP has become a critical tool in linguistic profiling within clinical linguistics. By automating the analysis of syntactic, semantic, and discourse-level features, NLP enhances the ability to diagnose and assess speech disorders, such as aphasia, ASD, and neurodegenerative conditions like

Alzheimer's disease. Techniques such as POS tagging, dependency parsing, and semantic similarity measures help identify linguistic irregularities, reduced lexical diversity, and semantic incoherence.

NLP Techniques for Linguistic Profiling

NLP techniques have become indispensable in the analysis of speech disorders, particularly in clinical linguistics. POS tagging is widely used to identify grammatical errors and omissions, which are particularly evident in conditions like aphasia, where patients often struggle with the correct use of words or sentence construction. By classifying words into categories such as nouns, verbs, adjectives, etc., POS tagging helps reveal discrepancies in speech structure, offering valuable insights into the patient's linguistic capabilities. Another important technique is dependency parsing, which analyzes the syntactic relationships between words in a sentence. This is especially useful for detecting disruptions in sentence structure, commonly seen in disorders like apraxia, where patients may have difficulty sequencing words and forming coherent sentences. Furthermore, semantic similarity measures are used to evaluate the coherence and relevance of the words used in a sentence. This technique plays a crucial role in detecting semantic incoherence, which is a hallmark of neurodegenerative conditions such as Alzheimer's disease, where patients often produce speech that lacks logical flow or contextually appropriate word choices. Together, these NLP techniques provide a detailed and automated method for assessing linguistic abilities, making them essential for diagnosing and monitoring various speech disorders.

NLP Workflow in Speech Disorder Assessment

The NLP workflow for speech disorder assessment, as illustrated in Figure 2, begins with tokenization, which breaks down continuous speech into smaller, manageable units like words and punctuation. This step is essential for organizing the speech into analyzable parts. Following tokenization, lemmatization reduces words to their root forms, ensuring consistency in the analysis by treating different word forms as a single entity. Next, syntactic parsing identifies the grammatical relationships between words in a sentence, which is particularly useful for detecting sentence construction issues in conditions like aphasia or motor speech disorders, where sentence structure may be impaired. Finally, semantic coherence evaluation assesses the logical flow and consistency of speech, helping to identify semantic inconsistencies, which are often seen in neurodegenerative diseases like Alzheimer's, where patients might exhibit disorganized or incoherent speech patterns. These steps collectively automate the process of analyzing speech, improving both the accuracy and consistency of diagnosing linguistic impairments.
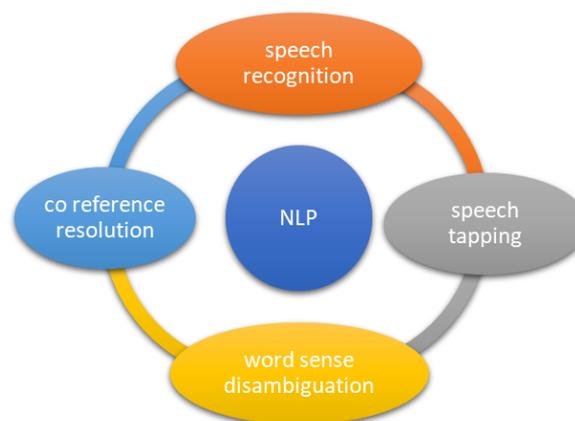

Figure 2. NLP Workflow in Speech Disorder Assessment

Challenges and Limitations

While NLP offers significant advantages, its performance can be limited when applied to disordered speech. Models trained on normative datasets may struggle with the atypical patterns found in speech disorders, such as those seen in aphasia and autism. This limitation underscores the need for domain-specific data that reflects the linguistic nuances of disordered speech. As pointed out by Fraza et al. [27], applying normative models without adaptations to disordered speech can lead to inaccurate assessments.

Future Directions

The future of NLP in clinical linguistics lies in multilingual datasets and more domain-specific training. Models need to be trained on speech from individuals with various disorders to improve their accuracy and generalizability. Additionally, multimodal data, such as incorporating facial expressions and gestures, could enhance the analysis and provide a more comprehensive assessment of speech disorders.

Multimodal and Cross-Lingual Approaches

The integration of multimodal data sources, such as visual lip movements, facial expressions, and articulatory gestures, has significantly enhanced the performance and robustness of AI-based speech disorder assessments. Visual speech recognition systems, when combined with acoustic models, provide a more comprehensive diagnostic insight, especially for individuals with severe motor impairments who may struggle to produce clear speech. Studies such as Majda-Zdancewicz et al. [28] highlight how multimodal systems that integrate both audio and visual cues can improve diagnostic accuracy, particularly in conditions like dysarthria or Parkinson's disease, where speech clarity is often compromised. These systems allow clinicians to analyze both articulatory and acoustic features, improving the diagnosis of disorders where motor control impacts speech production.

Moreover, cross-lingual adaptation has become a critical area of focus in expanding the applicability of AI systems across non-English and low-resource language populations. Recent advances in multilingual transformer models such as BERT (Bidirectional Encoder Representations from Transformers) have demonstrated the potential for fine-tuning AI models to work across multiple languages, reducing the dependency on large language-specific corpora. Dubey et al. [29] emphasized that cross-lingual transfer learning, particularly through models like multilingual BERT, can bridge the gap between English-based training data and the linguistic patterns of non-English languages, making AI systems more accessible and applicable in diverse linguistic contexts.

However, these multilingual models still face challenges, particularly in adapting to low-resource languages. Hunt et al. [30] point out that the scarcity of high-quality, multilingual data for speech disorders limits the effectiveness of these systems in accurately diagnosing speech impairments across different linguistic groups. Furthermore, cultural differences and regional dialects can complicate the training of AI models, which underscores the need for domain-specific datasets that are reflective of the diverse linguistic and cultural variations in speech disorders.

Figure 3 provides a thematic synthesis of the AI and ML techniques employed across various application domains, showcasing the integration of multimodal systems and cross-lingual models in speech disorder assessments. The figure outlines the key benefits and persistent

challenges of using these methods, particularly in the context of multilingual adaptation and the ethical governance of AI tools.
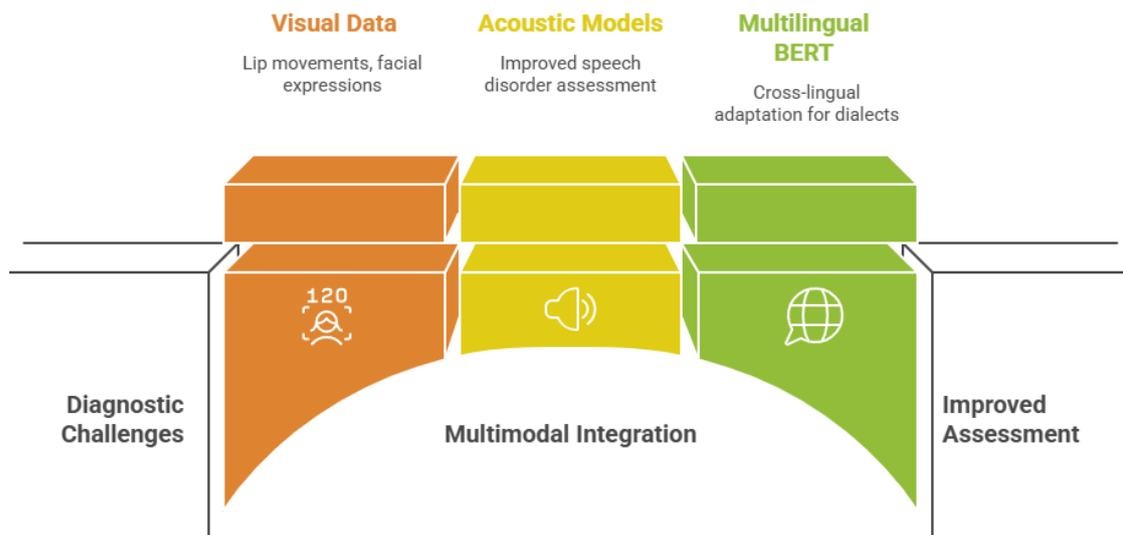

Figure 3. Multimodal and Cross-Lingual AI Integration

Figure 3 illustrates the integrated approach of combining visual data (e.g., lip movements and facial expressions) with acoustic models for improved speech disorder assessment. It highlights the various AI/ML techniques such as multilingual BERT for cross-lingual adaptation and shows how these tools can be used to address diagnostic challenges across diverse linguistic populations. The figure also provides a synthesis of application domains, the benefits of integrating these multimodal and cross-lingual approaches, and the persistent challenges such as data scarcity and cultural diversity gaps.

Despite these advancements, as summarized in Table 2, there remain significant challenges, particularly in the areas of generalizability and data availability. Multilingual AI models have been able to provide some degree of adaptation to non-English languages, but they still struggle with small and homogeneous datasets, making it difficult to develop truly universal diagnostic tools. The issue of linguistic diversity, especially in low-resource languages, persists as a major limitation, affecting the scalability of AI systems designed to assess speech disorders in diverse populations.

Table 2. Summarizing the Utilization of AI and ML in Speech Disorders

Application Area	AI/ML Techniques Used	Key Benefits	Major Challenges
Diagnosis	CNNs, RNNs, SVMs, LSTM	High accuracy in classifying speech impairments	Limited generalizability; small, homogeneous datasets
Screening & Assessment	NLP, Acoustic Feature Extraction, ASR	Early detection of disorders; real-time quantification	Lack of standardization; limited testing on diverse users
Intervention & Therapy	Reinforcement Learning, Decision Trees	Personalized therapy and feedback; remote accessibility	Limited clinical oversight; insufficient outcome studies

Application Area	AI/ML Techniques Used	Key Benefits	Major Challenges
Linguistic Analysis	POS Tagging, Parsing, Embedding Models	Detailed profiling of syntax, semantics, fluency	Poor robustness on atypical/disordered speech
Multilingual Applications	Transfer Learning, Multilingual BERT	Cross-lingual adaptation; low-resource language support	Data scarcity; cultural and linguistic diversity gaps
Ethical & Clinical Concerns	N/A (Design Principles)	Emphasizes explainability, user trust, and fairness	Lack of transparency, patient data privacy, regulatory hurdles

As seen in Table 2, the research on diagnostic tools has been prolific, yet multilingual adaptability and ethical governance remain under-explored areas. The potential of cross-lingual adaptation using methods like transfer learning with multilingual BERT is significant, but the field still faces challenges related to data scarcity, cultural gaps, and the lack of transparency in AI models.

Effectiveness of Different Machine Learning Algorithms

Performance comparisons of various ML algorithms, including Support Vector Machines (SVMs), Decision Trees, K-Nearest Neighbors (KNNs), Artificial Neural Networks (ANNs), Convolutional Neural Networks (CNNs), and Recurrent Neural Networks (RNNs), are visualized in Figures 3-8. These figures present a comparative analysis of how each algorithm performs in speech disorder diagnostics, focusing on their accuracy, computational efficiency, and interpretability. It is evident that while deep learning methods, particularly CNNs and RNNs, consistently outperform traditional approaches in terms of accuracy, they come with trade-offs in terms of interpretability and computational efficiency. This trade-off between performance and interpretability is a well-established issue in AI research, especially in clinical applications where understanding model decisions is crucial for clinicians. Hasan Saif et al. [31] and Amiri et al. [32] found that deep learning models such as CNNs and RNNs achieve higher accuracy rates in complex tasks like speech signal processing but are often criticized for being black-box models, making it challenging to explain their reasoning, which is vital for adoption in clinical settings.

Support Vector Machines (SVMs)

SVMs have been widely employed for speech classification tasks, especially in the context of diagnosing motor speech disorders like dysarthria. As shown in Figure 4, SVMs demonstrate strong generalization capabilities, particularly when the dataset is small and well-labeled. SVMs excel at finding hyperplanes that separate different classes of speech signals, offering high accuracy for linear classifications. However, SVMs can struggle with non-linear data and larger datasets, requiring more sophisticated kernel functions to maintain performance. Lever et al. [33] highlights that SVMs work best in high-dimensional spaces, where the feature set is carefully selected, which can be a limitation in dealing with unstructured speech data, which can be abundant in real-world clinical settings.
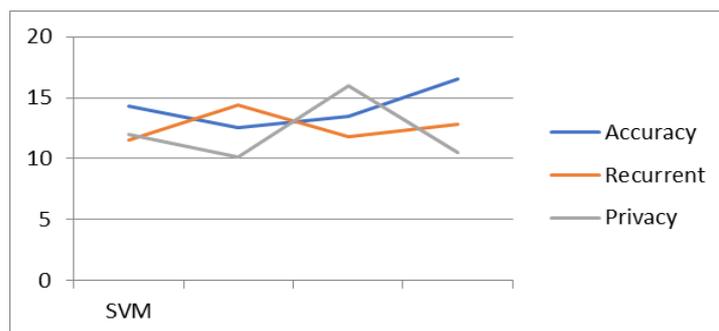
Figure 4. Effectiveness of SVMs Algorithms

K-Nearest Neighbors (KNNs)

KNNs are a simple yet effective algorithm for speech disorder classification, especially in terms of their intuitive approach to decision-making. As demonstrated in Figure 5, KNNs classify speech samples based on their proximity to labeled training data points, making them particularly well-suited for non-parametric speech recognition tasks. One advantage of KNNs is their ability to work well with small, localized datasets. However, KNNs suffer from significant computational inefficiency, especially as the dataset grows, because they require calculating the distance between test samples and all training samples at prediction time. Rasul et al. [34] found that KNNs perform relatively well in early-stage diagnosis of autism spectrum disorder (ASD) but are less effective for more complex disorders where the data structure is highly non-linear.
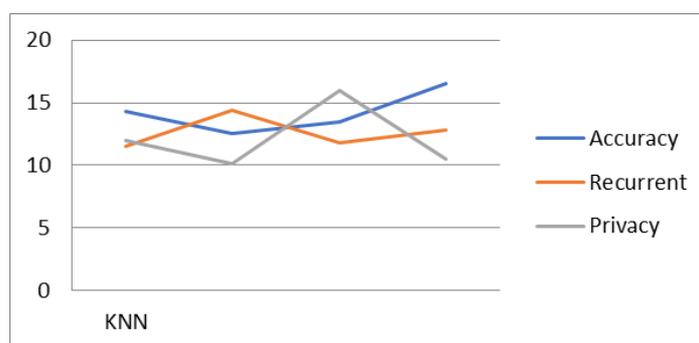

Figure 5. Effectiveness of KNNs Algorithms

Artificial Neural Networks (ANNs)

ANNs, particularly feed-forward neural networks, are foundational to many speech recognition and classification tasks. As illustrated in Figure 6, ANNs are capable of learning complex non-linear relationships between input features and output labels, making them useful for identifying patterns in speech data that traditional algorithms may miss. ANNs have demonstrated good performance in diagnosing speech disorders such as aphasia and stuttering. However, ANNs require large amounts of labeled data for training and can be prone to overfitting if not properly regularized. Abiodun et al. [35] emphasized that, although ANNs are powerful for speech disorder classification, their performance can vary significantly depending on the network architecture and training data used.
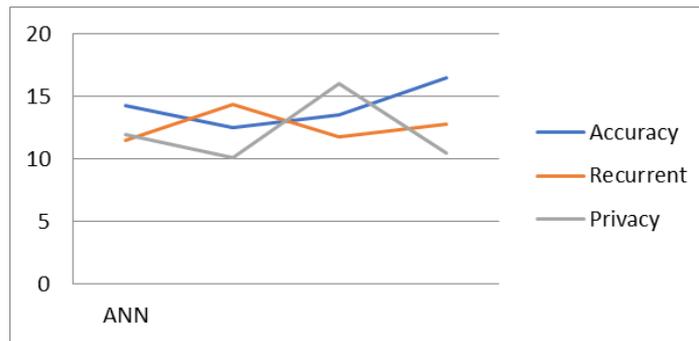
Figure 6. Effectiveness of ANNs Algorithms

Convolutional Neural Networks (CNNs)

CNNs have revolutionized image-based and speech signal processing tasks due to their ability to capture hierarchical features from raw data. As shown in Figure 7, CNNs are particularly effective in recognizing spatial hierarchies in speech spectrograms, which makes them well-suited for speech classification tasks. The deep architecture of CNNs allows them to automatically learn and extract features from speech signals, leading to high accuracy rates in diagnosing complex speech disorders like dysarthria and aphasia. Gupta et al. [36] demonstrated that CNNs outperform traditional models in terms of diagnostic accuracy in the context of speech disorder recognition but are computationally expensive, requiring substantial resources for training and inference.
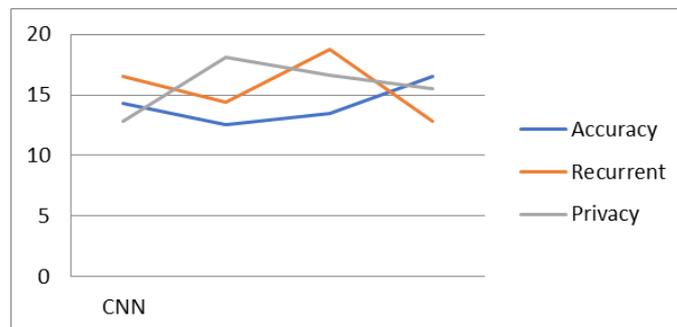

Figure 7. Effectiveness of CNNs Algorithms

Recurrent Neural Networks (RNNs)

RNNs have become the go-to model for speech disorder analysis, particularly for sequential data such as speech signals, where the timing of events plays a crucial role. Figure 8 illustrates the effectiveness of RNNs in capturing temporal dependencies in speech, making them ideal for analyzing fluency and rhythm-related disorders like stuttering. Unlike CNNs, which focus on spatial features, RNNs excel in tasks where the order and context of the data matter, such as in detecting delayed speech onset or prosodic abnormalities. However, the vanishing gradient problem remains a challenge for traditional RNNs, limiting their performance in long-term sequence predictions. Ken et al. [37] suggested that more advanced Long Short-Term Memory (LSTM) models can overcome this issue by maintaining better memory of previous inputs.
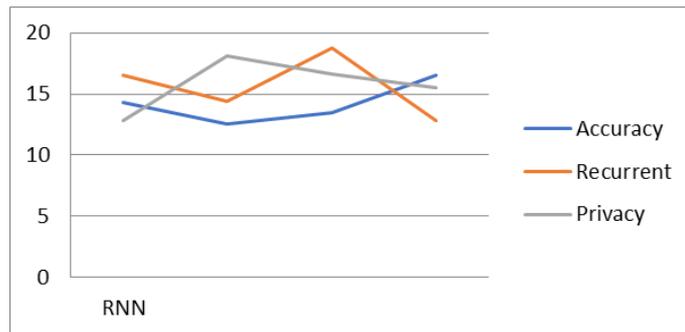
Figure 8. Effectiveness of RNNs Algorithms

Comparison and Performance Overview

Figure 9 provides a consolidated overview of the performance comparison of various machine learning algorithms in speech disorder diagnosis, showcasing the relative strengths and weaknesses of SVMs, Decision Trees, KNNs, ANNs, CNNs, and RNNs. While deep learning techniques such as CNNs and RNNs consistently outperform traditional algorithms in terms of accuracy, they come with significant trade-offs in terms of interpretability and computational efficiency. These models require large amounts of labeled data, and training them can be resource-intensive. Moreover, while CNNs and RNNs excel in accuracy, their black-box nature makes it difficult for clinicians to understand the rationale behind the model's predictions, which is a critical issue in medical applications where trust and explainability are key. Ding et al. [38] noted that CNNs and RNNs are best suited for large-scale speech disorder datasets but may not be as practical for smaller datasets or real-time applications.
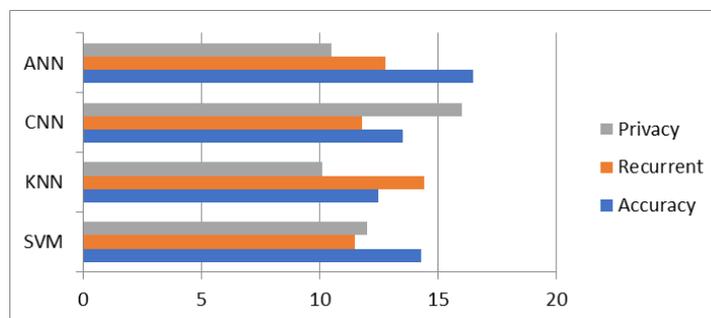

Figure 9. Effectiveness of Different Machine Learning Algorithms

Complementing these visual results, *see Table 3* provides a detailed overview of AI-powered assistive technology personalization features. These include adaptive vocabulary learning, voice customization, and context-aware suggestions, which significantly improve user engagement and accessibility. This table emphasizes that personalization is a critical determinant in the successful adoption of assistive communication technologies.

Table 3. AI-Powered Assistive Technologies: Personalization Overview

Feature	How It Supports Individual Needs	Example
Adaptive Vocabulary	Learns the user's common words and phrases	Predictive text in AAC devices suggests personalized words

Feature	How It Supports Individual Needs	Example
Voice Customization	Mimics or restores a user's natural voice	AI-generated voice clones for users with ALS
Context-Aware Suggestions	Adapts communication based on time, location, or activity	Recommends phrases like "I'm at school" during the day
Multimodal Input Support	Accepts eye gaze, touch, head movement, or voice commands	Eye-tracking systems for non-verbal users
Cultural and Language Options	Supports multiple languages and local dialects	Bilingual AAC tools with region-specific vocabulary
User-Controlled Settings	Lets users and caregivers adjust features and privacy settings	Customizing interface size, output speed, or tone

AI-Powered Assistive Technologies

The integration of AI into assistive technologies has profoundly enhanced the capabilities of tools designed for individuals with speech and language impairments. Recent advancements in personalization features such as adaptive vocabularies, voice customization, and context-aware phrase prediction have significantly improved user engagement and communication effectiveness. These personalization strategies allow assistive devices to adapt to individual needs, creating more natural and accessible interactions for individuals with neurological or speech-related disorders, such as amyotrophic lateral sclerosis (ALS), aphasia, and autism spectrum disorder (ASD).

Adaptive vocabularies are central to these advancements. They enable the system to learn the user's preferred words and phrases over time, providing suggestions based on frequent interactions. Predictive text functions in Augmentative and Alternative Communication (AAC) devices, such as those used by individuals with ALS, suggest personalized words or phrases as users communicate, reducing the effort required to find appropriate terms. This approach aligns with Elsahar et al. [39], who explored how AAC systems with predictive text improve communication speed and reduce cognitive load for individuals with motor impairments. Their research emphasizes the critical role of predictive algorithms in enhancing speech accessibility.

Voice customization is another powerful feature that has gained prominence, especially for individuals who have lost the ability to speak due to neurodegenerative diseases. AI-generated voice cloning has been successfully applied to restore or mimic a user's natural voice. For instance, Angrick et al. [40] demonstrated the efficacy of voice synthesis technology in restoring speech for patients with ALS, allowing them to communicate using a synthetic voice that closely resembles their natural one. The emotional impact of maintaining a familiar voice, rather than using a generic synthetic voice, was shown to enhance the overall well-being of patients, making voice customization a critical component in maintaining personal identity during disease progression.

Context-aware phrase prediction further improves communication by adapting suggestions based on external factors, such as time of day, location, and current activity. Nichol et al. [41] explored the potential of context-aware communication systems for individuals with aphasia, finding that systems capable of suggesting relevant phrases based on the user's context led to more fluid and efficient conversations. For example, suggesting phrases like "I am at home" or "I need

help" depending on the user's location improved the efficiency and naturalness of communication. Context-aware systems also provide emotional support by reducing frustration, which is often experienced by individuals with communication difficulties.

These personalization features are visualized in Figure 10, which summarizes their implementation and benefits. The figure highlights how these technologies adaptive vocabularies, voice customization, and context-aware phrase prediction contribute to the overall effectiveness of assistive technologies by improving accessibility and promoting natural communication. A key takeaway from Figure 10 is the importance of user-centered design in ensuring the sustained adoption and efficacy of these technologies in real-world settings. Personalization has been identified as a critical determinant for successful integration in daily life, especially for people with severe speech impairments.
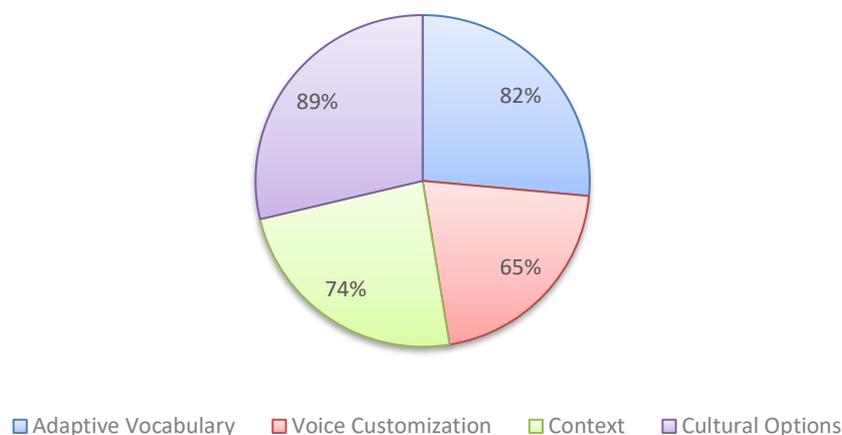

Figure 10. Personalization Features in AI-Powered Assistive Technologies

Figure 10 visualizes the core personalization features of AI-powered assistive technologies, including adaptive vocabularies, voice customization, and context-aware phrase prediction. It underscores the role of user-centered design, emphasizing how these features lead to enhanced accessibility, better engagement, and improved communication for individuals with speech and language impairments. The figure also highlights the dynamic relationship between personalization and real-world adoption.

The user-centered design approach, emphasized in Figure 10, ensures that the technology adapts to individual needs and provides a tailored experience. This approach not only boosts the effectiveness of the technology but also ensures greater user satisfaction and long-term adoption. Kumar et al. [42] argue that personalization is essential for engaging users, as it makes the technology feel more intuitive and responsive to their needs, particularly for individuals who rely on these tools for daily communication.

Ethical, Technical, and Clinical Considerations
The integration of AI into SLP introduces significant ethical, technical, and clinical challenges that must be carefully addressed to ensure effective and equitable deployment. As AI models become more prevalent in diagnosing and treating speech and language disorders, issues related to ethics, transparency, and inclusivity become increasingly important. Ethical considerations, such as bias

in AI models, data privacy, and patient consent, must be prioritized to prevent discrimination and ensure that AI tools are used fairly across diverse populations.

One of the major ethical concerns is the lack of transparency in AI decision-making processes. Many AI models, especially deep learning models like CNNs and RNNs, operate as "black boxes." These models are highly accurate but are difficult for clinicians to interpret, which can erode trust in their use. As Sadeghi et al. [43] highlighted, the inability to explain how a model reaches its conclusions makes it challenging for healthcare professionals to rely on these systems, particularly when they are making clinical decisions. XAI frameworks are emerging as a promising solution, aiming to provide more interpretability in AI systems. XAI approaches are designed to make AI decision-making more transparent, enabling clinicians to understand the rationale behind a model's output. According to Mienye et al. [44], XAI can significantly enhance clinician trust, as it allows them to validate AI outputs and incorporate them into their clinical judgment more confidently. Furthermore, human-in-the-loop (HITL) systems, where human clinicians are actively involved in the decision-making process, have been proposed to mitigate the risks associated with black-box models. HITL systems ensure that AI tools serve as assistive technologies rather than replacing human expertise.

Another crucial consideration is the need for inclusive AI models that reflect the diversity of speech disorders across ages, genders, and languages. Research by Xia et al. [45] emphasizes the need for diverse datasets that include a range of speech disorders, linguistic backgrounds, and cultural contexts. Current AI models, particularly those trained on predominantly English-language corpora, struggle with low-resource languages and cultural diversity. The lack of representation in datasets can lead to biased AI predictions, making it difficult to provide equitable care to non-English-speaking or underrepresented populations. To address this, AI tools must be developed with multilingual capabilities and designed to handle diverse speech patterns across cultures. Roumeliotis et al. [46] demonstrated that fine-tuning multilingual models like BERT to handle multiple languages significantly improves their performance in non-English settings, ensuring that AI systems can be applied globally.

User-centered design is another critical element in the development of AI-based speech-language technologies. According to Pascoe et al. [47], involving clinicians in the design process ensures that AI tools align with real-world clinical needs and workflows. Clinicians are best positioned to provide feedback on the practicality and usability of AI systems, leading to more effective tools that integrate seamlessly into clinical environments. Additionally, AI systems should feature user-friendly interfaces that simplify interactions for both clinicians and patients. The ease of use is crucial to encourage adoption and ensure that patients, especially those with speech and cognitive impairments, can effectively engage with the technology.

Multimodal data integration combining speech with non-verbal cues like facial expressions, gestures, or even brain signals has shown great promise in improving the accuracy and understanding of speech disorders. Ziegler et al. [48] demonstrated that integrating visual data from lip movements and facial expressions with acoustic signals can help overcome challenges in diagnosing motor speech disorders such as dysarthria. The use of multimodal data allows for a more holistic view of a patient's communication abilities, improving diagnosis and treatment outcomes. For instance, a patient with severe dysarthria may still be able to communicate effectively through facial cues or gestures, which can be captured and integrated by multimodal AI systems.

Finally, it is essential to test AI tools in real-world settings to ensure their effectiveness beyond controlled environments. Pilot testing in clinics, homes, and schools allows for the identification of potential challenges and barriers to adoption before AI tools are rolled out more widely. Jelassi et al. [49] conducted field trials of speech therapy AI tools and found that testing in real-world environments led to significant improvements in usability and patient compliance. These tests provide valuable insights into how AI tools perform when used by a broader, more diverse population and help fine-tune the technology for optimal performance.

Table 4. Strategies for Advancing AI and ML in SLP

Strategy	Description	Expected Benefit
Build Diverse Speech Datasets	Collect data from various speech disorders, ages, and languages	More accurate and inclusive AI models
Involve Clinicians in Design	Work with speech-language pathologists (SLPs) during development	Tools that meet real clinical needs
Improve Model Transparency	Make AI decisions easy to understand	Builds trust with users and professionals
Support Multiple Languages	Train models for non-English and underrepresented languages	Expands global accessibility
Use Multimodal Data	Combine speech with gestures, facial cues, or brain signals	Better understanding of communication abilities
Create User-Friendly Interfaces	Design tools that are easy for patients and clinicians to use	Improves adoption and daily usability
Test Tools in Real Environments	Pilot tools in clinics, homes, and schools	Ensures effectiveness in real-world settings

CONCLUSION

This review highlights the significant potential of AI and ML in diagnosing and managing speech disorders, offering improvements in diagnostic accuracy, therapeutic efficiency, and accessibility for underserved populations. While techniques such as CNNs, RNNs, and SVMs have demonstrated promising results in identifying speech impairments, challenges remain, including limited access to diverse datasets, issues with generalizability across languages, and a lack of transparency in deep learning models. To ensure broader applicability and inclusion, future research must focus on the development of multilingual models, multimodal approaches, and the integration of XAI and HITL systems, which are essential for ensuring transparency, trust, and clinical acceptance. Despite the promising advances, further efforts are needed to address these challenges and enhance the integration of AI and ML into clinical practice.

LIMITATIONS

This review has several limitations that should be considered. First, the focus is primarily on AI and ML applications in diagnosing and treating speech disorders, which may exclude other promising technological advancements in the field of speech-language pathology. Second, the studies reviewed are predominantly based on English-language datasets, which limits the generalizability of the findings to non-English-speaking populations and multilingual contexts. Furthermore, the relatively small sample sizes in many studies, as well as the homogeneity of the participant groups, may affect the robustness and applicability of the AI models across diverse clinical settings and cultural backgrounds. Finally, while the review addresses technical challenges, ethical considerations such as data privacy, model biases, and the regulatory framework for AI in healthcare are briefly discussed but require more in-depth exploration for practical implementation in clinical environments.

AUTHOR INFORMATION

Corresponding Author
Husnain Mukhdoom – Famous Forum of Worldwide Collaborators (Pakistan);
ⓘ orcid.org/0009-0005-0303-6417
Email: husnainmukhdoom06@gmail.com

Authors
Husnain Mukhdoom – Famous Forum of Worldwide Collaborators (Pakistan);
ⓘ orcid.org/0009-0005-0303-6417
S.R. Rajkumar – St.Joseph's College (India);
ⓘ orcid.org/0009-0000-0828-6574

AUTHOR CONTRIBUTION

H.M. and S.R.R. contributed equally to this study. H.M. conceptualized the research, designed the methodology, and supervised the project. S.R.R. managed data collection, performed analyses, and created visualizations. H.M. also conducted the literature review and drafted the discussion. Both authors reviewed, edited, and approved the final manuscript, ensuring accountability for all aspects of the work.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DECLARATION OF USE OF AI IN SCIENTIFIC WRITING

The authors declare that generative AI tools, including ChatGPT, were used for improving the writing and editing processes of this manuscript, while Napkin was utilized for the creation of visualizations. All other aspects of the manuscript, including data analysis and conceptual

development, were independently conducted by the authors, who assume full responsibility for the originality, accuracy, and integrity of the work.

REFERENCES

[1]   H. B. S. Knudsen et al., "Allocation and funding of Speech and Language Therapy for children with Developmental Language Disorders across Europe and beyond," *Res. Dev. Disabil.*, vol. 121, p. 104139, Feb. 2022. https://doi.org/10.1016/j.ridd.2021.104139

[2]   M. Aldè, U. Ambrosetti, S. Barozzi, and S. Aldè, "The Ongoing Challenges of Hearing Loss: Stigma, Socio-Cultural Differences, and Accessibility Barriers," *Audiol. Res.*, vol. 15, no. 3, p. 46, Apr. 2025. https://doi.org/10.3390/audiolres15030046

[3]   G. P. Usha and J. S. R. Alex, "Speech assessment tool methods for speech impaired children: a systematic literature review on the state-of-the-art in Speech impairment analysis," *Multimed. Tools Appl.*, vol. 82, no. 22, pp. 35021-35058, Sep. 2023. https://doi.org/10.1007/s11042-023-14913-0

[4]   D. Denman, R. Cordier, N. Munro, J.-H. Kim, and R. Speyer, "Standardized Measures Used Regularly by Speech-Language Pathologists when Assessing the Language Abilities of School-Aged Children: A Survey," *Folia Phoniatr. Logop.*, vol. 75, no. 5, pp. 334-349, 2023. https://doi.org/10.1159/000530718

[5]   Lohmander, K. Klintö, E. Schalling, A. Szabo Portela, K. Johansson, and A. McAllister, "Students take charge of Learning - Using e-learning in Perceptual Assessment in Speech-Language Pathology," *Scand. J. Educ. Res.*, vol. 65, no. 3, pp. 468-480, Apr. 2021. https://doi.org/10.1080/00313831.2020.1716064

[6]   Y. Yu, D. Appiah, B. Zulu, and K. A. Adu-Poku, "Integrating Rural Development, Education, and Management: Challenges and Strategies," *Sustainability*, vol. 16, no. 15, p. 6474, Jul. 2024. https://doi.org/10.3390/su16156474

[7]   X. Wang, G. W. Young, M. Z. Iqbal, and C. M. Guckin, "The potential of extended reality in Rural Education's future - perspectives from rural educators," *Educ. Inf. Technol.*, vol. 29, no. 7, pp. 8987-9011, May 2024. https://doi.org/10.1007/s10639-023-12169-7

[8]   N. Kaiser and C. K. Barstow, "Rural Transportation Infrastructure in Low- and Middle-Income Countries: A Review of Impacts, Implications, and Interventions," *Sustainability*, vol. 14, no. 4, p. 2149, Feb. 2022. https://doi.org/10.3390/su14042149

[9]   Z. Qian, K. Xiao, and C. Yu, "A survey of technologies for automatic Dysarthric speech recognition," *EURASIP J. Audio, Speech, Music Process.*, vol. 2023, no. 1, p. 48, Nov. 2023. https://doi.org/10.1186/s13636-023-00318-2

[10]   Z. Brahmi, M. Mahyoob, M. Al-Sarem, J. Algaraady, K. Bousselmi, and A. Alblwi, "Exploring the Role of Machine Learning in Diagnosing and Treating Speech Disorders: A Systematic Literature Review," *Psychol. Res. Behav. Manag.*, vol. 17, pp. 2205-2232, May 2024. https://doi.org/10.2147/PRBM.S460283

[11]   D.-H. Shih, C.-H. Liao, T.-W. Wu, X.-Y. Xu, and M.-H. Shih, "Dysarthria Speech Detection Using Convolutional Neural Networks with Gated Recurrent Unit," *Healthcare*, vol. 10, no. 10, p. 1956, Oct. 2022. https://doi.org/10.3390/healthcare10101956

[12] R. S. Calabrò and S. Mojdehdehbaher, "AI-Driven Telerehabilitation: Benefits and Challenges of a Transformative Healthcare Approach," *AI*, vol. 6, no. 3, p. 62, Mar. 2025. https://doi.org/10.3390/ai6030062

[13] C.-C. Lin, A. Y. Q. Huang, and O. H. T. Lu, "Artificial intelligence in intelligent tutoring systems toward sustainable education: a systematic review," *Smart Learn. Environ.*, vol. 10, no. 1, p. 41, Aug. 2023. https://doi.org/10.1186/s40561-023-00260-y

[14] J. R. Green, "Artificial Intelligence in Communication Sciences and Disorders: Introduction to the Forum," *J. Speech, Lang. Hear. Res.*, vol. 67, no. 11, pp. 4157-4161, Nov. 2024. https://doi.org/10.1044/2024_JSLHR-24-00594

[15] D. Deffner, J. M. Rohrer, and R. McElreath, "A Causal Framework for Cross-Cultural Generalizability," *Adv. Methods Pract. Psychol. Sci.*, vol. 5, no. 3, Jul. 2022. https://doi.org/10.1177/25152459221106366

[16] P. Radanliev, "AI Ethics: Integrating Transparency, Fairness, and Privacy in AI Development," *Appl. Artif. Intell.*, vol. 39, no. 1, Dec. 2025. https://doi.org/10.1080/08839514.2025.2463722

[17] L. Qin et al., "A survey of multilingual large language models," *Patterns*, vol. 6, no. 1, p. 101118, Jan. 2025. https://doi.org/10.1016/j.patter.2024.101118

[18] V. Hassija et al., "Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence," *Cognit. Comput.*, vol. 16, no. 1, pp. 45-74, Jan. 2024. https://doi.org/10.1007/s12559-023-10179-8

[19] S. Ebneyamini and M. R. Sadeghi Moghadam, "Toward Developing a Framework for Conducting Case Study Research," *Int. J. Qual. Methods*, vol. 17, no. 1, Dec. 2018. https://doi.org/10.1177/1609406918817954

[20] Y.-Y. Lin et al., "A Speech Command Control-Based Recognition System for Dysarthric Patients Based on Deep Learning Technology," *Appl. Sci.*, vol. 11, no. 6, p. 2477, Mar. 2021. https://doi.org/10.3390/app11062477

[21] S. Alharbi, M. Hasan, A. J. H. Simons, S. Brumfitt, and P. Green, "Sequence labeling to detect stuttering events in read speech," *Comput. Speech Lang.*, vol. 62, p. 101052, 2020. https://doi.org/10.1016/j.csl.2019.101052

[22] W.-S. Hsu, G.-T. Lin, and W.-H. Wang, "Enhancing Dysarthric Voice Conversion with Fuzzy Expectation Maximization in Diffusion Models for Phoneme Prediction," *Diagnostics*, vol. 14, no. 23, p. 2693, Nov. 2024. https://doi.org/10.3390/diagnostics14232693

[23] C. K. Themistocleous, M. Andreou, and E. Peristeri, "Autism Detection in Children: Integrating Machine Learning and Natural Language Processing in Narrative Analysis," *Behav. Sci. (Basel).*, vol. 14, no. 6, p. 459, May 2024. https://doi.org/10.3390/bs14060459

[24] W.-Z. Zheng, J.-Y. Han, C.-K. Lee, Y.-Y. Lin, S.-H. Chang, and Y.-H. Lai, "Phonetic posteriorgram-based voice conversion system to improve speech intelligibility of dysarthric patients," *Comput. Methods Programs Biomed.*, vol. 215, p. 106602, Mar. 2022. https://doi.org/10.1016/j.cmpb.2021.106602

[25] R. Fusaroli, E. Weed, R. Rocca, D. Fein, and L. Naigles, "Caregiver linguistic alignment to autistic and typically developing children: A natural language processing approach illuminates the interactive components of language development," *Cognition*, vol. 236, p. 105422, Jul. 2023. https://doi.org/10.1016/j.cognition.2023.105422

[26] E. Conti et al., "Autism Spectrum Disorder and Childhood Apraxia of Speech: Early Language-Related Hallmarks across Structural MRI Study," *J. Pers. Med.*, vol. 10, no. 4, p. 275, Dec. 2020. https://doi.org/10.3390/jpm10040275

[27] C. Fraza, S. Rutherford, B. R. Bučková, C. F. Beckmann, and A. F. Marquand, "The promise of quantifying individual risk for brain disorders through normative modeling, a narrative review," *Neurosci. Biobehav. Rev.*, vol. 176, p. 106284, Sep. 2025. https://doi.org/10.1016/j.neubiorev.2025.106284

[28] E. Majda-Zdancewicz, A. Potulska-Chromik, M. Nojszewska, and A. Kostera-Pruszczyk, "Speech Signal Analysis in Patients with Parkinson's Disease, Taking into Account Phonation, Articulation, and Prosody of Speech," *Appl. Sci.*, vol. 14, no. 23, p. 11085, Nov. 2024. https://doi.org/10.3390/app142311085

[29] P. Dubey, P. Dubey, R. Raja, and S. S. Kshatri, "Bridging language gaps: The role of NLP and speech recognition in oral english instruction," *MethodsX*, vol. 14, p. 103359, Jun. 2025. https://doi.org/10.1016/j.mex.2025.103359

[30] E. F. Hunt, C. Nang, S. J. Meldrum, and E. Armstrong, "Assessment of multilingual children by Western Australian speech-language pathologists: A survey of practices, barriers and facilitators," *Int. J. Speech. Lang. Pathol.*, vol. 27, no. 3, pp. 370-384, May 2025. https://doi.org/10.1080/17549507.2025.2499516

[31] F. Hasan Saif, M. N. Al-Andoli, and W. M. Y. W. Bejuri, "Explainable AI for Alzheimer Detection: A Review of Current Methods and Applications," *Appl. Sci.*, vol. 14, no. 22, p. 10121, Nov. 2024. https://doi.org/10.3390/app142210121

[32] Z. Amiri, A. Heidari, N. J. Navimipour, M. Esmaeilpour, and Y. Yazdani, "The deep learning applications in IoT-based bio- and medical informatics: a systematic literature review," *Neural Comput. Appl.*, vol. 36, no. 11, pp. 5757-5797, Apr. 2024. https://doi.org/10.1007/s00521-023-09366-3

[33] J. Lever et al., "Facing & mitigating common challenges when working with real-world data: The Data Learning Paradigm," *J. Comput. Sci.*, vol. 85, p. 102523, Feb. 2025. https://doi.org/10.1016/j.jocs.2024.102523

[34] R. A. Rasul, P. Saha, D. Bala, S. M. R. U. Karim, M. I. Abdullah, and B. Saha, "An evaluation of machine learning approaches for early diagnosis of autism spectrum disorder," *Healthc. Anal.*, vol. 5, p. 100293, Jun. 2024. https://doi.org/10.1016/j.health.2023.100293

[35] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," *Heliyon*, vol. 4, no. 11, p. e00938, Nov. 2018. https://doi.org/10.1016/j.heliyon.2018.e00938

[36] S. Gupta et al., "Four Transformer-Based Deep Learning Classifiers Embedded with an Attention U-Net-Based Lung Segmenter and Layer-Wise Relevance Propagation-Based Heatmaps for COVID-19 X-ray Scans," *Diagnostics*, vol. 14, no. 14, p. 1534, Jul. 2024. https://doi.org/10.3390/diagnostics14141534

[37] D. Kim, Y. Lee, K. Chin, P. J. Mago, H. Cho, and J. Zhang, "Implementation of a Long Short-Term Memory Transfer Learning (LSTM-TL)-Based Data-Driven Model for Building Energy Demand Forecasting," *Sustainability*, vol. 15, no. 3, p. 2340, Jan. 2023. https://doi.org/10.3390/su15032340

[38]   H. Ding, H. Hou, L. Wang, X. Cui, W. Yu, and D. I. Wilson, "Application of Convolutional Neural Networks and Recurrent Neural Networks in Food Safety," *Foods*, vol. 14, no. 2, p. 247, Jan. 2025. https://doi.org/10.3390/foods14020247

[39]   Y. Elsahar, S. Hu, K. Bouazza-Marouf, D. Kerr, and A. Mansor, "Augmentative and Alternative Communication (AAC) Advances: A Review of Configurations for Individuals with a Speech Disability," *Sensors*, vol. 19, no. 8, p. 1911, Apr. 2019. https://doi.org/10.3390/s19081911

[40]   M. Angrick et al., "Online speech synthesis using a chronically implanted brain-computer interface in an individual with ALS," Jul. 2023. https://doi.org/10.1101/2023.06.30.23291352

[41]   L. Nichol, S. J. Wallace, R. Pitt, A. D. Rodriguez, Z. Z. Diong, and A. J. Hill, "People with aphasia share their views on self-management and the role of technology to support self-management of aphasia," *Disabil. Rehabil.*, vol. 44, no. 24, pp. 7399-7412, Nov. 2022. https://doi.org/10.1080/09638288.2021.1989501

[42]   V. Kumar, A. R. Ashraf, and W. Nadeem, "AI-powered marketing: What, where, and how?," *Int. J. Inf. Manage.*, vol. 77, p. 102783, Aug. 2024. https://doi.org/10.1016/j.ijinfomgt.2024.102783

[43]   Z. Sadeghi et al., "A review of Explainable Artificial Intelligence in healthcare," *Comput. Electr. Eng.*, vol. 118, p. 109370, Aug. 2024. https://doi.org/10.1016/j.compeleceng.2024.109370

[44]   D. Mienye et al., "A survey of explainable artificial intelligence in healthcare: Concepts, applications, and challenges," *Informatics Med. Unlocked*, vol. 51, p. 101587, 2024. https://doi.org/10.1016/j.imu.2024.101587

[45]   Y. Xia, S.-Y. Shin, and J.-C. Kim, "Cross-Cultural Intelligent Language Learning System (CILS): Leveraging AI to Facilitate Language Learning Strategies in Cross-Cultural Communication," *Appl. Sci.*, vol. 14, no. 13, p. 5651, Jun. 2024. https://doi.org/10.3390/app14135651

[46]   I. Roumeliotis, N. D. Tselikas, and D. K. Nasiopoulos, "Leveraging Large Language Models in Tourism: A Comparative Study of the Latest GPT Omni Models and BERT NLP for Customer Review Classification and Sentiment Analysis," *Information*, vol. 15, no. 12, p. 792, Dec. 2024. https://doi.org/10.3390/info15120792

[47]   L. Pascoe et al., "Strategic Considerations for Selecting Artificial Intelligence Solutions for Institutional Integration: A Single-Center Experience," *Mayo Clin. Proc. Digit. Heal.*, vol. 2, no. 4, pp. 665-676, Dec. 2024. https://doi.org/10.1016/j.mcpdig.2024.10.004

[48]   W. Ziegler et al., "Speech and Nonspeech Parameters in the Clinical Assessment of Dysarthria: A Dimensional Analysis," *Brain Sci.*, vol. 13, no. 1, p. 113, Jan. 2023. https://doi.org/10.3390/brainsci13010113

[49]   Jelassi, K. Matteli, H. Ben Khalfallah, and J. Demongeot, "Enhancing Personalized Mental Health Support Through Artificial Intelligence: Advances in Speech and Text Analysis Within Online Therapy Platforms," *Information*, vol. 15, no. 12, p. 813, Dec. 2024. https://doi.org/10.3390/info15120813