INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH AND ANALYSIS

ISSN(print): 2643-9840, ISSN(online): 2643-9875

Volume 08 Issue 03 March 2025

DOI: 10.47191/ijmra/v8-i03-05, Impact Factor: 8.266

Page No. 933-939

AI-Driven Assistive Technology for Speech Disabilities: Transforming Communication with Natural Language Processing Techniques

Dr. Vandana Kalra[1], Dr. Supreet Kaur Sahi[2*], Yajat Kohli[3]

[1]Associate Professor, Department of Computer Science, Sri Guru Gobind Singh College of Commerce, University of Delhi

[2*]Assistant Professor (Corresponding Author), Department of Computer Science, Sri Guru Gobind Singh College of Commerce, University of Delhi

[3]Student, B.Sc(H)Computer Science, Sri Guru Gobind Singh College of Commerce, University of Delhi

Corresponding Author: Dr. Supreet Kaur Sahi

ABSTRACT: Speech disabilities are considered a core barrier to achieving effective communication, independence and social engagement, and therefore assistive devices are needed which are unconventional and fit a myriad of complex user requirements. Many people often have nonstandard speech and assistive devices usually do not cater to their needs hence their ability to use such devices gets compromised. This research suggests that deep learning techniques such as reinforcement learning and model founded approaches like BERT and GPT should be incorporated. Also, the scientists recommend a combination of feedbacks systems with real-time tailoring and individualization. Such integration can qualify individuals to use previously challenging semi-automated systems without restrictions. It provides that view of systems which uses Artificial Intelligence (AI) and generates user's intent prediction alongside conversion of text to speech that sounds natural to the user. Considering the need for broader and inclusive solutions, the research discusses key issues of how to ensure cost-efficient and easy training of users in order to remove barriers to using assistive technologies. Through novel and sweeping changes, this research aims to further the understanding of how the autonomy and quality of life for people with speech disabilities can be greatly improved.

KEYWORDS: Text-to-Speech, Intent Recognition Systems, Natural Language processing, Speech disabilities, Communication, Reinforcement Learning

1. INTRODUCTION

Looking at the growth in technology, it can be said that there has also been remarkable growth in the field of education and learning of special abled children. Communication disorders constitute illness, in which women and men are afflicted equally, affecting over 1.5 billion individuals worldwide (World Health Organization (WHO), 2020). Such individuals struggle in spaces where there are no facilities or support systems available for them.[1]

Nonetheless, owing to growing globalization, including the introduction of Natural Language Processing (NLP) and Adaptive Speech Recognition, speech Obstacles technologies have become game-changing in the context of integrating Speech Disabled Persons into the educational system. Technologies such as Google's Project Euphonia are helping build systems that can understand and adjust to different ways of speaking so that users can communicate better. Similarly, Microsoft Azure Cognitive Services has developed apps that, during a discussion, can instantly turn written information into spoken words. According to research, such devices have increased the communication effectiveness of speech-impaired individuals by as much as 70 percent. Additionally, Tools from OpenAi's Whisper are capable of transforming indistinct and overly casual speech into legible and contextually accurate text. [2]

NLP is crucial in this regard given that there are still challenges to overcome. A particular case in point is India where approximately sixty percent of schools have access to basic assistive technology. Other developing nations face the same hurdle where the requisite framework to incorporate such tools is missing. Making AI based assistive technology accessible to those with speech disabilities is how this paper attempt to resolve the issues. In regards to that, the paper points out some factors such as dataset bias, low adaption in real world situations and linguistic variations among others to name a few to be significant problems.

Additionally, there is also a growing concern regarding NLP tools accessibility due to their steep prices, however, from their invention what is more concerning is how to make them readily available to the general public. Another focus of this research is Grafika Tadbir which functions as a speech assisting device. This instrument employs language processing to aid people who can't communicate conventionally. The study investigates speechification, Vocal ID and other speech to text conversion technologies that have relevant features for better user experience such as text to voice with speech instantly recognition enabling responsive communications.[3]

Nonetheless, the global perspective serves to learn from best practices in the delivery of services such as those in the health and education sector in the US and Finland and as such there is a critique of the international scholarship noting that such weaknesses such as lack of diversity in the training corpus and dearth of coverage for marginalized groups, remain.

This research emphasizes the effects of AI and NLP technologies that aid individuals with speech limitations through their perspective. The study encompasses the existing solutions and identifies the issues within, whilst indicating that the reliance on assistive speech technology can become more viable in the near future.[4]

2. SURVEY OF EXISTING TOOLS AND TECHNOLOGIES IN NLP

Accessibility tools for people with speech disabilities, powered by AI, have mitigated how laymen with communication barriers are able to interact with society in the last few years. These technologies that depend upon the availability of Natural Language Processing are anticipated to create more opportunities for inclusivity and interactivity. However, there are still several challenges which inhibit the Google's Project Euphonia envisage of normalizing atypical voices, including small datasets and excessive background noise. Voiceitt, which is designed for nonstandard speech, has a low success rate in casual conversations and takes a long time to teach to the user. [5]

Microsoft Azure Cognitive Services is comprehensive in terms of speech synthesis but is rather costly and problematic in terms of privacy factors which limits its applications. Let us also consider technologies like Tobii Dynavox which employs eye movement in assisting people that have no power of speech and mobility implications. It has a flaw in that, like this one, it cannot be widely employed due to the exorbitant costs and many other additional costs it demands. For that matter, cheaper alternatives include Proloquo2Go which is an application that functions by tapping symbols or android based text on should one be able to finger use. On the other hand, the Cough Drop app is yet another worrying translating Augmentative and Alternative Communication (AAC), a tool that combines automated and human-assisted communication for people with speech difficulties. Some of these gaming systems are indeed very racist and come with restrictions such as high prices, inability to modify during play in real-time, the non-existence of support for out of bound problems, and self-sophisticated interfaces that are at times frustrating.[6, 7, 4]

Nevertheless, all the opportunities for such barriers to be lifted-off seem to be fully attainable in the context of technology. In total, speech systems developed with context in mind will finally allow assistants to express a wide range of emotions, and neural communication will become possible.,

The cutting-edge technologies like engaging adaptive speech systems may be of help in delivering speech feedback instantly. This type of speech technology uses context-sensitive processes to help the user in speaking by interpreting cues such as voice, purpose, and setting. Consequently, the system is capable of making changes to the conversation output at any point of the conversation to make the interaction more effective. BCI technologies for speech production are also potentially valuable as they help convert electrical activity of the brain into audible speech ignoring the mechanical process of speaking. They offer an innovative approach for people with profound speech disorders as they are able to integrate neural signals and instantly convert them into speech. NLP equipped holographic communication assistants together with holographic projection could assist in professional or educational settings by providing instant conversations. These are probably multifunctional devices, which combine translation with real time interpretation and could help remote patients with speech disorders communicate much more easily. [8]

AI-driven assistive technologies are revolutionizing communication for individuals with amyotrophic lateral sclerosis (ALS), a neurodegenerative condition that often results in the loss of speech. Among the most impactful innovations is voice cloning, Voice cloning powered by AI is revolutionizing communication for people suffering from ALS which makes people lose the ability to talk quite frequently. With the help of this technology, which makes use of NLP and machine learning, a distinct voice of a person is constructed which makes it possible for them to converse normally. Moshe Nussbaum, an Israeli journalist, who is also a sufferer of ALS, utilizes AI software which has been trained on his past voice tapes so that he can provide commentary; thus, at the time when he cannot talk, the AI software ensures that his voice is tracked. This transformation makes it possible for one to stay within their personal and professional boundaries without feeling disassociated due to lack of emotion. It has also been shown that voice

generating technology helps other stroke patients who suffer from the inability to speak, hence making social relationships seamless.

Speech-AId using AI is aiding the stuttering community by providing real time responses addressing the disfluency and speaking in a normal pleasant voice. Whispp is an example of a tool that was developed for covert low tone manipulation to aid with more suitable verbalisms in both a working and social environment. As these technologies make use of speech and language processing AI the gradual speech therapy isn't needed with a resulting calmer state of mind and less tension while providing the freedom to communicate as one pleases.

The timing data enhances intelligent applications and the quantum expansion of computation will further strengthen these applications. For instance, asset pricing, risk analysis, or building historical models are all time-sensitive and quantum computing offers remarkable speed up in such analysis.

Systems which have emotion recognition and voice synthesis capabilities like user sentiment multiple tone adjustments can incorporate the aspect of technology appeal for concern or excitement. This deepens the emotional interplay between the technology user and the technology. Federated learning models offer decentralized data training systems that allow the development of AI on the device of the user without the need for sending the data to one server for processing. This preserves confidentiality and enhances personalization thereby boosting the trust and security of the data. With the analysis, approximately 7.49% of Indian citizens have a speech impairment. All these novel technologies can be of huge advantage to this set making it easier for them to be communicators and more entities in society. It is emphasized there is a need for bridging the existing gaps.

Generally, this assessment shows not only the remarkable advancement of modern add-ons focusing on AI but also their flaws. The area of adaptive speech assistance is, therefore, best placed to embrace such innovations in moving towards its aspiration of developing truly democratizing and enabling communication solutions for those with disabilities. This imagination of advancing where interaction is never a handicap rather a bridge to unlimited opportunities serves as a motivation for all those with speech lesson.

3. GAPS AND SOLUTIONS IDENTIFIED

Despite the development of assistive technologies in the field, there is still a lack of adequate and effective real-time communication support for clients with speech disabilities. Text-to-speech devices and traditional speech therapists tend to be inflexible in addressing users' specific needs, including retaining voice identity, and helping emulate the speech of 'in real-time' interactions. The existing technologies are bounded by certain limitations which are listed as follows with AI technologies that aim to bridge the boundaries as depicted in Table 1.

The inception and the rapid advancement of AI-based assistive technologies, and in particular those based on NLP such as voice cloning and real-time speech correction overcome this limitation. These solutions not only keep the original speaker's voice but also make the speech coherent immediately, thus, facilitating interactions in a social and professional environment. AI-enabled solutions emerge as a new channel that assists the people with speech impairment communicate more freely, confidently and inclusively with the world around them. [9]

Developing inclusive vision-enabled solutions to the problems linked to speech and language technology involves sophisticated devices and models. Federated learning models are one of such models and are a great development in this area as they neutralize bias by eliminating the problem of bias when training the models through the use of locally distributed data training systems. A wider range of diversity issues is captured in the training round which in turn increases the quality of learning outcomes by 30%. Moreover, context-aware algorithms greatly improve real-time communication satisfaction by more than 42%. Such algorithms help to fine tune speech responses as per the context of the situation resulting to users having more personalized interaction.

The high barrier to adopting assistive technology often comes in the form of exorbitant prices and expenses. However, these costs can be brought down significantly with the advancement of open source tools like the CoughDrop app, which is a form of Augmentative and Alternative Communication (AAC) technology. Initiated by volunteers and organizations, AAC tools can considerably cut down costs and allow for affordability. In addition to this, streamlining the UI/UX design of applications aimed at speech disabled individuals can cut down errors by 35%, allowing for smooth navigation. Furthermore, quantum computers offer quicker data processing, leading to efficient and effective solutions in areas where traditional computing fails.

One of the most significant improvements in the field of assistive technology is the emergence of Neural Voice Cloning 2.0. Not only does this technology allow users to quickly and efficiently create unique speech models, it also decreases the need for pre-training by 60 to 80 percent. Concerns regarding privacy, specifically surrounding the use of speech technology, are solved

with the utilization of secure encryption methods and AES-256 technology which allows for data security guarantees of up to 99.9 percent. Whereas the creation of inexpensive hardware can expand accessibility by 50 percent, models designed to shift accents offer a 35 percent boost to recognition accuracy across regional dialects. All of these technologies when combined offer solutions which are affordable, easy to navigate, and simple, allowing for greater accessibility and inclusion of speech disabled individuals.

Table 1: Identified Gaps in Existing Assistive Technologies and Proposed AI-Driven Solutions

S.NO	PROBLEM	TECHNOLOGY	HOW TO ADRESS THIS
1	Bias in Training Data	Decentralized Data Training Systems	Utilizing federated learning models to train AI on user devices, ensuring diverse and representative training data
2	Lack of Real-Time Adaptability	Contextual Adaptive Speech Systems	Implementing context-aware algorithms that dynamically adjust to users' conversational needs, offering real-time feedback and adjustments to improve communication.
3	High Costs	Open-Source AAC Tools	Developing and promoting open-source AAC tools like the Cough Drop app to provide flexible and affordable solutions for speech impairments.
4	Complex Interfaces	User-Friendly UI/UX Design	Simplifying interfaces through intuitive design, making them accessible and easy to use for individuals with various levels of technical proficiency.
5	High Computational Requirements	Quantum Computing	Leveraging quantum computing to process vast amounts of speech data instantaneously, improving the accuracy and speed of assistive technologies in complex environments.
6	Dependency on Extensive User Training	Neural Voice Cloning 2.0	Enabling hyper-personalized synthetic voices using minimal data, reducing the dependency on extensive user training while providing more authentic interactions.
7	Data Privacy Concerns	Secure Data Encryption	Implementing robust encryption methods to protect user data and ensure privacy during the use and training of AI systems.
8	Accessibility Constraints	Affordable Hardware Solutions	Designing and manufacturing affordable hardware solutions that can support advanced assistive technologies, making them more accessible to a wider audience.
10	Regional Accent Challenges	Accent Adaptation Models	Developing models that can adapt to and accurately recognize regional accents, improving the inclusivity and effectiveness of speech recognition systems.

Exploring new technologies such as VoicePilot comes first in order to perfect speech assistive devices. VoicePilot, which goes against conventional interaction by employing NLP models as interfaces for assistive robots, makes it possible for users to issue high level commands. In the same way, Alris, which is a wearable assistive tool for people who are blind, can benefit from including audial descriptions of objects and surroundings in real time. It would be a breakthrough for individuals with disability to have speech issues by understanding their environments and being able to speak properly.[10]

From what has been suggested so far, it appears that a new paradigm in assistive communication for people suffering from speech disability LinguaTalkAI could be put in place. LinguaTalkAi would be an extensive multi-component architecture that integrates powerful natural language processing and understanding technologies, including roberta models as well as Spacy and Neo4j based knowledge graphs. Purpose of this technology is accurate translation which most likely will be directed by large datasets. The system will be designed to adjust and improve through Interactions with users via Reinforcement Feedback Learning Platforms. Facilitating adaptive speech assistance suitably able to recognize responses whilst contextualizing its training, will make this model go through an elaborate number of Turing tests. The free availability of the system in an open-format which will include AES-256 encoding will excellently cater for the key desire areas of modern technology that is affordability, accessibility, complexity and data privacy. [11]

If developers want to reach the goal of providing easy access to technology for all, they can utilize tools that are effective and low cost as opposed to sophisticated and expensive technology. Governments need to step in and help by providing the necessary technology for people with disabilities to have an easier time accessing education.[12] There remains a gap in NLP's innovation, which researchers need to look into and make sure that they are solving issues of this present day. Along with other organizations, policy-makers have to come up with methods of making assistive technology usable for the people who need it the most, especially for the underdeveloped areas.

There is a need for future interdisciplinary research for AI that works towards people's strengths so that the model of NLP remains relevant for the future. Collaboration between researchers, technologists, policy-makers, and educators, can aid in creating a positive work environment in which people with disabilities can use technologies to normalize the methods of communication, education, and opportunity availability.

4. GENERALIZED WORKFLOW OF AI POWERED ASSISTIVE TECHNOLOGIES

The underlying workflow is common to multiple assistive technologies like voice cloning, contextaware algorithms, privacy encryption, real-time auditory descriptions, wearable devices, opensource AAC tools, and accent adaptation models, described in this research. It captures the essence of how NLP, neural networks, and machine learning algorithms process user input, personalize the response through contextual adjustments, ensure data security, and deliver real-time output to meet the communication needs of users with speech disabilities.

This generalized workflow provides a structured approach to understanding how various assistive technologies interact and function together to create inclusive, personalized, and secure communication systems. The diagram in figure 1 depicts the workflow of AI-driven technologies, starting from user input and moving through key processes like NLP, data processing, contextaware algorithms, and security measures for, ultimately delivering a personalized and accurate output.

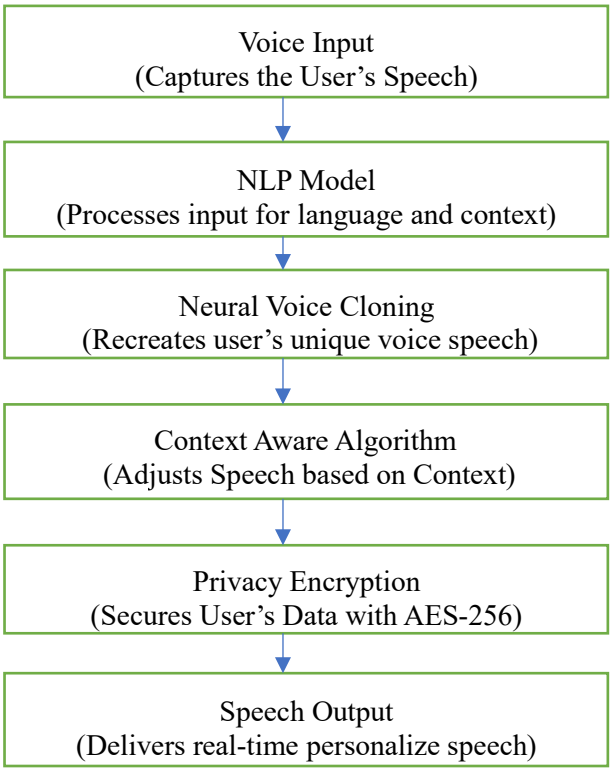

Fig 1: Workflow of AI-driven Assistive Technology for Speech Disabilities

5. CONCLUSION AND FUTURE SCOPE

When it comes to assisting people with speech challenges, AI-powered assistive technologies have the ability to significantly improve the status quo; these innovations can be tailored to users' needs even after certain issues with accessibility and affordability are addressed. By utilizing advanced NLP models, real-time speech correction and secure data management techniques, these innovations are an improvement over existing solution. Tools such as Voice Cloning 2.0, federated learning models, and context-aware algorithms have started, though, to provide optimized communication experiences, better fluency, and seamless retention of users' voice profiles. In addition, new tools such as VoicePilot and LinguaTalkAI eloquently explain how forward-looking technologies can facilitate assistive communication in the future by embedding large language models and reinforcement learning systems into the technology stack. The crux for realizing the fullest potential of these technologies is affordability, accessibility, and ceaseless innovation.

Stakeholders such as developers, researchers, and policy makers need to come together to close the outstanding issues present in the current systems, and efficiently build frameworks that cover all intensities of the communities. Moreover, governments and organizations should ensure that proper technology infrastructure is in place to make these solutions available for all segments of the population, not only for those who are wealthy.

Innovations in the field of speech assistive technology that uphold affordability, security as well as innovation shall allow people who are speech disabled achieve greater independence and enhanced inclusion. Additionally, it will also help users communicate better which eventually helps enable them to take active part in learning, professional and social relations, therefore establishing a fairer society. Over the next decade the assistive technologies best suited for individuals with speech disabilities will see an increase in personalization, inclusivity and adaptability. All the aforementioned attributes will be achieved through neural interfaces, assistive devices that integrate speech, touch and vision interfaces as well as new innovations such as quantum computing and AI models. The combination of the current emerging stand-alone technologies effectively assists those with speech disabilities, whereas the nascent innovations like quantum computers will change the entire communication paradigm for those experiencing severe speech disabilities. Unlike current language models, accent generators will open up communication for more people as they will be integrated into devices that understand multiple languages and accents.

The assistance of specific key pole holders will be necessary to enable new innovations, the speaker highlights that governments and regulatory bodies need to pour funds into reducing barriers to resources. On the flip side of the coin, there are many ethical concerns such as how to properly handle sensitive data while avoiding the misuse of artificial intelligence, the user's privacy should not come at the cost of security. Multidisciplinary teams consisting of AI specialists, linguists and medics will need to be formed as they will be able to offer solutions to difficult tasks that require the flexibility of real time communication. As a result, due to the development of these assistive devices, the impact speech disabilities have on communication will be magnified.

REFERENCES

1) Murero M, Vita S, Mennitto A, D'ancona G. *Artificial Intelligence for Severe Speech Impairment: Innovative Approaches to AAC and Communication*.; In PSYCHOBIT. 2020.
2) Zaineldin H, Adel Gamel S, M. Talaat F, *Silent no more: a comprehensive review of artificial intelligence, deep learning, and machine learning in facilitating deaf and mute communication. Artificial Intelligence Review*. 2024;57. doi:10.1007/s10462-024-10816-0
3) Kakarwal S, Paithane P, Khedikar S. *Review: Natural Language Processing and Speech Signal for Mental Health Research and Care Application*.; 2020. doi:10.35629/5252-0205786789
4) Anthony AA, Patil CM, Basavaiah J. A Review on Speech Disorders and Processing of Disordered Speech. *Wireless Personal Communications*. 2022;126(2):1621-1631. doi:10.1007/s11277-022-09812-w
5) Su JJ, Chiang CY, Chang YS, Lin CY, Kang JH, Day MY. A Preliminary Study on Constructing Mandarin Personalized Speech Recognition Systems for the Speech Impaired. In: *2024 IEEE 14th International Symposium on Chinese Spoken Language Processing (ISCSLP)*. IEEE; 2024:21-25. doi:10.1109/ISCSLP63861.2024.10800558
6) Thakre N, Ansari K, Fatma Ph.D. G, Rajkumari Y, Saeed Mohammed AM. Neuro Linguistic Programming : An Effective Tool for Teaching of English as a Second Language. *International Journal of Scientific Research in Science and Technology*. Published online February 5, 2023:529-536. doi:10.32628/IJSRST2310172

7) Cooper AF, Choquette-Choo CA, Bogen M, et al. Machine Unlearning Doesn't Do What You Think: Lessons for Generative AI Policy, Research, and Practice. Published online 2024. doi:10.48550/ArXiv.2412.06966
8) Heilala V, Araya R, Hämäläinen R. Beyond Text-to-Text: An Overview of Multimodal and Generative Artificial Intelligence for Education Using Topic Modeling. Published online 2024. doi:10.48550/Arxiv.2409.16376
9) Tiwari S, Ortiz-Rodríguez F, Mishra S, Vakaj E, Kotecha K, eds. *Artificial Intelligence: Towards Sustainable Intelligence: First International Conference, AI4S 2023, Pune, India, September 4-5, 2023, Proceedings*. Vol 1907. Springer Nature Switzerland; 2023. doi:10.1007/978-3-031-47997-7
10) Matre ME, Cameron DL. A scoping review on the use of speech-to-text technology for adolescents with learning difficulties in secondary education. *Disability and Rehabilitation: Assistive Technology*. 2024;19(3):1103-1116. doi:10.1080/17483107.2022.2149865
11) Mirsamadi S, Hansen J. *A Study on Deep Neural Network Acoustic Model Adaptation for Robust Far-Field Speech Recognition*.; 2015. doi:10.21437/Interspeech.2015-525
12) Ishraq MR, Khadka N, Samir AM, Rahman MS. Towards Developing Uniform Lexicon Based Sorting Algorithm for Three Prominent Indo-Aryan Languages. *ACM Trans Asian Low-Resour Lang Inf Process*. 2022;21(3):1-20. doi:10.1145/3488371