International Journal of Engineering Research and Development
e- ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 22, Issue 5 (May 2026), PP 97-101

AI-Powered Communication Aids for Speech-Impaired Individuals in Clinical Settings

Esther Oyindamola Oyanibi[1] , Yusuff Bolaji Ajegbile[2,]Adedoluwa Adeniyi[3] , Gbenga Gbaja[4]
[1]*SC Johnson College of Business, Cornell University Ithaca, New York, USA*
[2]*Department of Health and Human Kinetic, University of Ibadan, Nigeria*
[3]*Leeds City University Ibadan,College of Applied Science,Department of Software Engineering*
[4]*Humber College,Canada.*

Abstract
Clinical practice is based on communication, and patients with speech, language, or communication needs (SLCN) are exposed to systemic obstacles that lead to a loss of autonomy and diagnostic accuracy. The presented academic literature review will discuss the disruptive character of Artificial Intelligence (AI) to the assistive communication technology and will leave the traditional rule-based Augmentative and Alternative Communication (AAC) behind in favor of the dynamic and deep-learning-based technology. This article is based on a recent literature review (published in 2019-2025) and focuses on the possibility of implementing Natural Language Processing (NLP), gesture-sensitive Automatic Speech Recognition (ASR), and multimodal Large Language Models (LLMs) in clinical practice. The key findings demonstrate that AI-based technologies such as smart wearable badges, social assistive robots, etc. are vital towards the improvement of turn-taking interactions, as well as self-expression in non-verbal autistic people and patients with acquired disorders, such as aphasia.. Nevertheless, the review finds serious issues which involve: ability bias in the generative models, data privacy risks, and the necessity of culturally sensitive interfaces. This paper concludes that clinically validated, ethically-driven AI systems are necessary in social justice and independence of healthcare. The participatory co-design and long-term longitudinal assessment should be the priority in future research because these technologies must be based on the lived experiences of neurodiverse populations.

Date of Submission: 27-04-2026 Date of acceptance: 06-05-2026

I. Introduction

Human survival and quality of life depend on sound communication as the two-way process of sending and receiving messages, and it involves linguistic skills, emotional response, and social behaviours (de Marchena et al., 2025). In the clinical context, patient safety and effective health outcomes depend on the skill to report the symptoms correctly, present medical history, and conduct an informed consent (de Marchena et al., 2025). Nevertheless, the rates of severe speech and communication impairments are around 1 percent of the world population (Elsahar et al., 2019). Radical innovation accessibility The rise of Artificial Intelligence (AI) and interconnectedness with the human-computer interaction has given us intuitive and personalized solutions that fit specific needs (Kooli and Chakraoui, 2025).

The clinical heterogeneity of the burden of speech and communication impairment is caused by a variety of conditions, such as cerebral palsy, amyotrophic lateral sclerosis (ALS), Autism spectrum disorder (ASD), stroke-associated acquired aphasia, traumatic brain injury, and neurodegenerative diseases (Beukelman and Light, 2020). All these conditions force a particular communicative profile, although at the endpoint of all they are all uniformly disadvantaged through clinical encounter where verbal exchange is the expected form of interaction. Research has always reported that patients with communication impairments get less clinical explanations, higher rates of diagnostic error, and are generally underrepresented in the shared decision making process compared to their healthy counterparts (Blackstone et al., 2019). Such inequalities do not occur by accident; they are structural imbalance in the way healthcare systems are structured with the normative assumptions of speech and the real-world experiences of people inpatients with different speech patterns.

In the past, addressing this gap was based on low technology augmentative and alternative communication (AAC) tools like picture boards, alphabet charts, and eye gaze frames, which, although significant, were slow, lacked contextual flexibility, and scalability in busy clinical settings (Beukelman and Light, 2020). The digitally aided AAC devices that were introduced in the late twentieth century were an important but not a significant change; they were relatively static, and they required a lot of cognitive and motor effort on the part of the user who already has some physical or neurological disability. Conversion of these tools in specialised rehabilitation units to acute, emergency, or primary care environment has been especially slow, and as a result, a substantial proportion of speech impaired patients have been made to lack effective communicative assistance at the point of maximum clinical urgency.

It is against this context that the intersection of machine learning with natural language processing (NLP) with computer vision and wearable biosensing technologies has signaled a qualitatively new phase in the evolution of communication aids. On the one hand, unlike their predecessors, AI powered systems are able to learn using the residual communicative signals of individual users, be it eye movement, facial micro activity, electromyographic output, or partial vocalisation and convert it into intelligible speech or text in near real time (Sadeghian et al., 2022). Importantly, these systems are also adaptive: they are enhanced with use, can accommodate changing clinical manifestations, can be integrated into the already existing infrastructure of a hospital, such as electronic health records and telehealth programs. This not only makes them not just assistive devices but also dynamic components of a wider digital health system.

AI based assistive technologies are starting to fill this gap between intent and expression, where the individual requiring assistance has complex communication needs (CCN). Modern AI interfaces are built around the concept of ability based design, which is concerned with how a user strengths are built up instead of his faults (Choi et al., 2025). The technological development of these aids, including rule-based systems and multimodal deep learning model, and their particular applications in therapeutic and hospital settings is discussed within the frames of the given review. This review will act as a roadmap towards improving clinical outcomes of the speech impaired by examining the existing situation.

II. REVIEW OF RELATED WORKS

Kooli and Chakraoui (2025) carried out an interdisciplinary, multi-modal study of AI-based assistive technologies, including screen readers and Natural Language Processing (NLP) interfaces, in the framework of inclusive education. Their investigation showed that even though AI-based interfaces have a tremendous effect on the autonomy and academic involvement of students with cognitive and physical disabilities, there are still considerable systemic obstacles. Most importantly, they found that the success of such tools determination lies both on technological design as well as educational ecosystem preparedness, such as educator readiness and algorithmic bias mitigation.

Kambouri, Simon, and Brooks (2023) examined the quantitative effects of the speech-to-text (STT) technology, which is the Dragon system, on the writing ability of young learners with special educational needs in the UK. Through their study, they proved that the technical control and quality of text on a screen were greatly improved using AI-driven speech recognition in comparison to using the traditional handwriting. The article actually demonstrated a remarkable effect size of 1.81 on gains on screen-written text (when compared to no intervention), indicating that STT enables young writers to overcome the motor and cognitive challenge of writing manually, but they point out that the results of such uncontrolled pilot studies need to be supported by randomised control trials.

Choi et al. (2025) designed and introduced AACessTalk, an LLM-generated tablet-based mediation system that is designed to help autistic children with minimal verbal skills (the MVA) engage in reciprocal communication with their parents. Their deployment study indicated a paradigm change in the communications: the guidance mediated by AI stimulated parents to go beyond straightforward instructions to the strategies of interaction that promoted the balanced turn-taking. It is important to note that the authors explained that the core benefit of AI in this case is that it acts as an assistive partner that would enable the child agency and social engagement instead of just being an automated answer.

Alotaibi, Subahi, Alghanmi, and Rizwanullah (2025) suggested a novel technique of accessibility (ASLA-DLHGR) that helps the hearing and speech-impaired by employing deep learning-learned hand gesture recognition. The researchers combined the SqueezeNet model with a bio-inspired-tunicate swarm algorithm (TSA) to optimize each of the hyperparameters, thus resulting in a better gesture recognition accuracy of 99.98%. The technical breakthrough demonstrates that the constraints of real-time performance and system reliability in the past have been overcome, and highly optimized, lightweight deep learning models can be effective even without a powerful computer.

In the article by Curtis, Lau, and Neate (2024), the authors examined the social and clinical implication of wearable technology by creating the Breaking Badge, an eInk-based smartbadge that allows people to have complex communication requirements. The researchers, with participants living with aphasia, claimed that conventional tablet-based AACs devices were stigmatizing and tend to block the non-verbal communication channels of the user, based on a co-design approach. They highlight that discrete wearable AI support should be a priority to facilitate an experience of interdependence, or assist users in making visible an invisible disability and need in the world without causing social isolation that the high-visibility assistive hardware frequently causes.

III. THEORETHICAL FRAMEWORK

Speech Impairment and Communication Barriers in Healthcare

Healthcare speech impairments develop as a result of a wide range of neurological and physiological disorders. Aphasia is an acquired language disorder caused by stroke, which affects almost two million individuals in the USA alone and impairs reading, writing, and speech production but does not usually affect intelligence (Kim et al., 2025). Moreover, not all people with Autism Spectrum Disorder (ASD) develop a normal speech, and even a quarter of them may be unable to speak at all (de Marchena et al., 2025; Ullah et al., 2021). Additional reasons are progressive ailments such as Amyotrophic Lateral Sclerosis (ALS), in which speech is distorted or replaced, and it is almost impossible to communicate with any common voice assistant (Kim et al., 2025; Elsahar et al., 2019).

The implications of these health care disabilities are far reaching. Patients can feel vulnerable to misunderstandings and they lack the acknowledged agency (Al-Araimi et al., 2024). The communication barriers affect the relationships, employment, and overall health outcomes throughout the lifespan (de Marchena et al., 2025). In acute care or emergency departments, communication rate gap may create severe delays in reporting symptoms, and the lack of awareness of medical professionals who are not familiar with the specifics of communication disabilities often results in social exclusion (Yusufali et al., 2023; Curtis et al., 2024).

Evolution of Assistive Communication Technologies

There is a long-standing division of assistive technology into no-tech, low-tech, and advanced technology (Elsahar et al., 2019). No-tech solutions are based on human perception of movements, whereas low-tech solutions imply the use of picture boards such as Picture Exchange Communication System (PECS) (Elsahar et al., 2019; Yusuf et al., 2025). High-tech solutions of yesteryear might have compelled users to scroll through tiresome sets of symbols to type out a message, which is prohibitively slow and leads to unsatisfactory sender-receiver interaction patterns (Curtis et al., 2024).

The transition to AI-based systems is a paradigm shift compared to just repairing speech as a way to supplement overall communication (Curtis et al., 2024). Today, Deep Learning (DL) is used to detect non-verbal communication and transform it into a synthesized voice in real-time (Omoyemi, 2024; Singh and Singh, 2025). The latest developments can enable the devices to take pictures of the real world and instantly create context-based vocabulary, which can be used in spontaneous interaction in a clinical and educational context via the latest developments (Just-in-time programming) (de Marchena et al., 2025; Choi et al., 2025).

Core AI Technologies Enabling Communication Aids

The current landscape of intelligent communication aids is powered by five key technical domains:

- **Natural Language Processing (NLP):** NLP interfaces redefine comprehension for those with cognitive disabilities by simplifying dense text and providing context-specific feedback that evolves with the user (Kooli & Chakraoui, 2025).
- **Speech Recognition and Synthesis:** Advanced ASR systems, such as **Whisper**, have shown increased robustness in transcribing disfluent speech from patients with aphasia (Kim et al., 2025). Modern synthesized speech uses mathematical algorithms to generate natural-sounding voices, providing greater flexibility than pre-recorded digitized speech (Elsahar et al., 2019).
- **Computer Vision and Gesture Recognition:** Vision-based systems analyze hand landmarks (via frameworks like **MediaPipe**) and classify gestures using Convolutional Neural Networks (CNNs) (Singh & Singh, 2025; Patil et al., 2025). Platforms can now recognize complex sign language or repetitive autistic motions with accuracies exceeding 95% (Ullah et al., 2023; Manap et al., 2025).
- **Predictive Models:** ML frameworks incorporate Recurrent Neural Networks (RNNs) and **LSTMs** to manage temporal dependencies in gestures and predict user intent, reducing message composition time by up to 30% (Omoyemi, 2024; Rafiq et al., 2023).
- **Multimodal AI:** The integration of multimodal LLMs allows for the simultaneous processing of linguistic, acoustic, and gestural information, bridging semantic gaps that speech alone cannot fill (Kim et al., 2025).

Applications in Clinical Settings

AI communication aids are increasingly integrated into specialized clinical workflows:

- **Smart AAC Devices:** The **InkTalker**, an eInk smartbadge, provides patients with a discreet way to signal clinical needs like "Asking for help" or "Seat request," reducing the stigma associated with bulky tablets (Curtis et al., 2024).
- **AI-Mediated Reciprocal Platforms:** Systems like **AACessTalk** utilize LLMs to curate contextual vocabulary cards for minimally verbal children while providing real-time guidance to parents and clinicians during turn-taking (Choi et al., 2025).

- **Socially Assistive Robotics (SARs):** Humanoid robots like **NAO** are used in therapeutic sessions to improve imitation skills and social interaction through deep learning-based action recognition (Alnafjan et al., 2025; Nadeem et al., 2025).
- **Telemedicine and Regional Platforms:** Innovative platforms for Arabic-speaking populations in Oman integrate **Content-Based Image Retrieval (CBIR)** with voice feedback, filling the gap in culturally appropriate assistive tools (Al-Araimi et al., 2024).

Benefits and Clinical Impact

The clinical benefits of AI-powered aids are high. To begin with, these devices contribute to patient autonomy, meaning that individuals with motor disabilities or disorders such as cerebral palsy can control the environment and dictate directly into a voice command or gesture command when assessing a patient (Kooli & Chakraoui, 2025). Secondly, they improve the accuracy of diagnosis; AI analysis of speech patterns and repetitive behaviors can offer the clinician objective quantitative data (Singh and Singh, 2025).

Thirdly, gesture-to-speech enable social inclusion in communal space in real-time, by allowing faster rhythms of interaction (Manap et al., 2025). Lastly, data shows that speech-to-text (STT) technology can improve the quality of output and self-esteem of ineligible patients with writing difficulties because of communication challenges (Kambouri et al., 2023).

Ethical, Privacy, and Regulatory Considerations

The use of AI in healthcare requires a strict ethical code. The most obvious issue is ability bias; research indicates that LLMs and chatbots have the ability to present individuals with disabilities as more restricted than they really are, which can create or even strengthen stereotypes (Urbina et al., 2024; Kooli and Chakraoui, 2025). In addition, the interpretive reliability becomes an ethical risk, as the incorrect recognition of the intention of a patient in an AAC app may contribute to the frustration or clinical mistake (Curtis et al., 2024).

The most important concerns privacy and data security because such systems handle the communication logs and behavioural data that are sensitive (Kooli & Chakraoui, 2025). It has been legitimate to be concerned about data processing without permission and the possibility of so-called Generative AI Addiction Syndrome (GAID) which could result in social withdrawal (Kooli and Chakraoui, 2025). Regulatory protections should make these instruments inclusive to different types of socio-economic situations so that they do not replicate the systemic inequalities (Kooli & Chakraoui, 2025).

Challenges and Limitations

Although the potential is present, there still exist technical and systemic constraints. In noisy clinical environments or with processing of unusual patterns of speech, accuracy may be decreased (Kim et al., 2025; Kooli and Chakraoui, 2025). Non-verbal control is not always available with commercial voice assistants, and thus they are not as helpful to people who use gestures (Esquivel et al., 2024). Moreover, most of the available AAC systems are Western systems that do not accommodate multilingual features, e.g. Arabic voice recognition (Al-Araimi et al., 2024).

Systemic obstacles comprise the educator and clinician willingness, as most practitioners are not familiar with AI affordances (Kooli & Chakraoui, 2025). Access in underserved areas is also restricted by infrastructure requirements, including constant internet connectivity to process AI-based products in a cloud (Alotaibi et al., 2025).

Future Research Directions

To advance the field, future research should explore:

1.Brain-Computer Interfaces (BCI): Non-invasive BCIs in combination with smart orators will provide users with severe motor disabilities with sufficient computing power (Elsahar et al., 2019; Verbaarschot et al., 2021).

2.Multimodal Emotion Recognition: Improving therapeutic robots with the capacity to identify affect by evaluating facial expressions and bodily signals (Nadeem et al., 2025).

3.Customized AI Models: The storage of personal conversational patterns and topic history through long-term memory use of LLMs to support more intuitively (Choi et al., 2025).

4.AR/VR Integration: Creation of AR letterboards and virtual therapeutic spaces to simulate social real-life experiences to practice skills (Voultsiou and Moussiades, 2025; Alabood et al., 2024).

IV. Conclusion

The revolutionary change in clinical access can be made through representative AI-based communication. These tools, based on the synergies of NLP, deep learning and computer vision, convert non-verbal modalities, where unique gestures are converted to uncoded forms of vocalisations, into text and speech to understand, thereby overcoming the traditional boundaries to social and medical inclusion. Allowing inflexible AAC to be transformed into the adaptive will allow users to interact in communicative agency again and play a fair role in the society again but this is subject to the responsible innovation. The possibility of becoming successful due to solving such problems as algorithmic bias, data privacy, and institutional preparedness using certain professional education is an opportunity. The ethically oriented AI, which is grounded on the principles of participatory co-design and neurodiversity, will become a necessary stimulus of human empowerment and social justice in the healthcare eco system (Kooli and Chakraoui, 2025; Choi et al., 2025).

References

[1]. Al-Araimi, A., Zheng, Y., & Liu, H. (2024). Communication Platform for Non-verbal Autistic children in Oman using Android mobile. *Presented at Annual Conference of the Association for Learning Technology 2024*.

[2]. Alnafjan, A., Alghamdi, M., Alhakbani, N., & Al-Ohali, Y. (2025). Improving Imitation Skills in Children with Autism Spectrum Disorder Using the NAO Robot and a Human Action Recognition. *Diagnostics*, *15*(1), 60.

[3]. Annapoorna, E., Nikhil, B. J., Kashyap, B., Abhishek, J., & Vadlapatla, T. S. S. (2023). Hand Gesture Recognition and Conversion to Speech for Speech Impaired. *E3S Web of Conferences*, *391*, 01148.

[4]. Cassidy, S. A., Stenger, B., Van Dongen, L., Yanagisawa, K., Anderson, R., Wan, V., Baron-Cohen, S., & Cipolla, R. (2015). Expressive visual text-to-speech as an assistive technology for individuals with autism spectrum conditions. *Computer Vision and Image Understanding*, *148*, 193–200.

[5]. Choi, D., Park, S., Lee, K., Hong, H., & Kim, Y.-H. (2025). AACessTalk: Fostering Communication between Minimally Verbal Autistic Children and Parents with Contextual Guidance and Card Recommendation. In *CHI Conference on Human Factors in Computing Systems* (Article 222). ACM.

[6]. Curtis, H., Lau, Y. H., & Neate, T. (2024). Breaking Badge: Augmenting Communication with Wearable AAC Smartbadges and Displays. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM.

[7]. de Marchena, A., Cuneo, N., Gurbuz, E., Brown, M., Trujillo, J., & Bergstrom, J. (2025). Communication in Autistic Adults: An Action-Focused Review. *Current Psychiatry Reports*, *27*, 471–481.

[8]. Elsahar, Y., Hu, S., Bouazza-Marouf, K., Kerr, D., & Mansor, A. (2019). Augmentative and alternative communication (AAC) advances: A review of configurations for individuals with a speech disability. *Sensors*, *19*(8), 1911.

[9]. Esquivel, P., Gill, K., Goldberg, M., Sundaram, S. A., Morris, L., & Ding, D. (2024). Voice assistant utilization among the disability community for Independent Living: a rapid review of recent evidence. *Human Behavior and Emerging Technologies*.

[10]. Kambouri, M., Simon, H., & Brooks, G. (2023). Using speech-to-text technology to empower young writers with special educational needs. *Research in Developmental Disabilities*, *135*, 104466.

[11]. Kim, S., Lee, D., Stark, B., & Han, J. (2025). Gesture-Aware Zero-Shot Speech Recognition for Patients with Language Disorders. *GenAI4Health Workshop @ AAAI*.

[12]. Kooli, C., & Chakraoui, R. (2025). AI-driven assistive technologies in inclusive education: benefits, challenges, and policy recommendations. *Sustainable Futures*, *10*, 101042.

[13]. Manap, Z., Baharin, A. H., Zainuddin, S., Sultan, J. M., & Yusof, A. L. (2025). Enhancing Communication Accessibility: A Deep Learning Approach for Assistive Hand Gesture Recognition in Speech Disability Communities. *IJRISS*, *9*(9).

[14]. Nadeem, M., Barakat, J. M. H., Daas, D., & Potams, A. (2025). A Review of Socially Assistive Robotics in Supporting Children with Autism Spectrum Disorder. *Multimodal Technologies and Interaction*, *9*(98).

[15]. Omoyemi, O. E. (2024). Machine learning for predictive AAC: Improving speech and gesture-based communication systems. *World Journal of Advanced Research and Reviews*, *24*(01), 2569–2575.

[16]. Patil, S. S., Chougule, O. P., Kadam, M. R., Kate, A. M., & Bhivse, V. R. (2025). Real-Time Sign Language & Gesture Recognition for Speech-Impaired Individuals. *International Journal for Research in Applied Science & Engineering Technology*, *13*(3).

[17]. Rafiq, R. B., Karim, S. A., & Albert, M. V. (2023). An LSTM-based Gesture-to-Speech Recognition System. *Proceedings (IEEE International Conference on Healthcare Informatics)*.

[18]. Rashidan, M. A., Sidek, S. N., Yusof, H. M., Khalid, M., Dzulkarnain, A. A. A., & Ghazali, A. S. (2021). Technology-Assisted Emotion Recognition for Autism Spectrum Disorder (ASD) Children: A Systematic Literature Review. *IEEE Access*, *9*, 3060753.

[19]. Singh, A. K., & Singh, V. (2025). Advanced Gesture Recognition for Autism Spectrum Disorder Detection: Integrating YOLOv7, Video Augmentation, and VideoMAE for Naturalistic Video Analysis. *ArXiv*.

[20]. Ullah, F., AbuAli, N. A., Ullah, A., Ullah, R., Siddiqui, U. A., & Siddiqui, A. A. (2023). Fusion-Based Body-Worn IoT Sensor Platform for Gesture Recognition of Autism Spectrum Disorder Children. *Sensors*, *23*(3), 1672.

[21]. Urbina, J. T., Vu, P. D., & Nguyen, M. V. (2024). Disability Ethics and Education in the Age of Artificial Intelligence: Identifying Ability Bias in ChatGPT and Gemini. *Archives of Physical Medicine and Rehabilitation*.

[22]. Voultsiou, E., & Moussiades, L. (2025). A systematic review of AI, VR, and LLM applications in special education: Opportunities, challenges, and future directions. *Education and Information Technologies*, *30*, 19141–19181.

[23]. Yusuf, R., Md.Salleh, N., Taisin, J. N., Md Adam, N. F., & Uda @ Longgok, Z. a. (2025). AI-Assisted Communication Tools for Non-Verbal Students in Special Education. *International Journal of Academic Research in Business and Social Sciences*, *15*(3), 249–256.