World Journal of Advanced Engineering Technology and Sciences
eISSN: 2582-8266
Cross Ref DOI: 10.30574/wjaets
Journal homepage: https://wjaets.com/

(REVIEW ARTICLE)

Check for updates

Speech-to-text and gesture recognition tools for non-verbal autistic accessibility

Esther Oyindamola Oyanibi [1] and Akinode John Lekan [2,*]

[1] *SC Johnson College of Business, Cornell University Ithaca, New York, USA.*
[2] *Computer Science Department, Federal Polytechnic Ilaro, Ogun State, Nigeria.*

World Journal of Advanced Engineering Technology and Sciences, 2025, 17(02), 522-529

Publication history: Received on 27 August 2025; revised on 19 November 2025; accepted on 26 November 2025

Article DOI: https://doi.org/10.30574/wjaets.2025.17.2.1384

Abstract

Many autistic people have long-term problems with communication, including poor proficiency in spoken language and the use of non-verbal language. These challenges are core impediments to quality interaction and independent living at clinical, educational and social levels. Although the traditional Augmentative and Alternative Communication (AAC) based, are fundamental, they do not usually have the ability to maintain real-time personalization and responsiveness that is needed to support fluent, high-fidelity interaction. The following review discusses how the Artificial Intelligence (AI) platforms, including speech-to-text (STT) and gesture recognition, have become sustainable accessibility tools to the non-verbal population with autism. A systematic review of recent literature assesses key themes that define successful deployment: precision, emphasis on algorithmic resilience to non-stereotypical vocalizations and idiosyncratic gestural communication; the requirement of sensory alignment in interface design to accommodate neurological variations; the imperative of inclusivity, particularly in relation to access and representational datasets; and overarching ethical concerns, including data privacy, possible sensory load, and algorithmic bias. This discussion finds that although AI-driven accessibility has a huge promise to promote autonomy and optimize the speed of interaction, its implementation demands a high-level of clinical alignment, interdisciplinary and commitment to neurodiversity-informed design values.

Keywords: Autism Spectrum Disorder (ASD); Augmentative and Alternative Communication (AAC); Speech-to-Text (STT); Gesture Recognition; Artificial Intelligence (AI)

1. Introduction

Communication is one of the essential human rights, and, in most cases, non-verbal or minimally-speaking individuals on the Autism Spectrum are deprived of this right due to insurmountable and persistent obstacles (de Marchena et al., 2025; Kooli and Chakraoui, 2025). A significant part of autistic individuals fail to master spoken language fluent, which requires them to use multidimensional and usually eccentric communication channels (de Marchena et al., 2025; Curtis et al., 2024). Such obstacles are especially experienced in such critical areas of life as healthcare, education, social interactions, and work (de Marchena et al., 2025; Kooli and Chakraoui, 2025; Ullah et al., 2021).

This standard of care that is established is usually dependent on the support of Augmentative and Alternative Communication (AAC) tools. Picture exchange systems (PECS; Bondy and Frost, 1994), considered to be a low-technology traditional AAC system, have been of basic speech-generating device support (Yusuf et al., 2025). Nevertheless, such traditional tools have certain limitations. Some of them are based on slow manual selection or primitive predictive features, reduced rate and responsiveness of communication, which may eventually lead to the frustration of a user and limited sense of agency (Omoyemi, 2024). Moreover, certain interventions unintentionally cannot contribute to the already existing communicative strengths of a person, including paralinguistics and non-verbal hints, which disrupt naturalistic multimodal communication (Curtis et al., 2024).

* Corresponding author: Akinode John Lekan

This is where the quick pace of the development of Artificial Intelligence (AI) offers a timely chance to fill in these shortcomings. AI-based interfaces, with the integration of such technologies as an advanced speech recognition (STT) and gesture recognition based on deep learning, are likely to provide customized, adaptive, and responsive communication solutions (Kooli and Chakraoui, 2025; Omoyemi, 2024). Such new assistive technologies will incorporate sophisticated nonverbal cues and unusual vocalizations into natural speech, making it possible to engage in learning and social activities in a like manner never before (Jaliaawala and Khan, 2020; Omoyemi, 2024).

The article describes a systematic review of research literature on AI-based STT and gesture recognition as non-verbal access on the autistic side. It then bases the discussion on a pertinent theoretical framework and then reviews significant empirical studies. The following paragraphs discover the important emerging points and deliberate the practical strengths and limitations of this technology and the ethical imperatives that must prevail to inform the responsible use of this technology in special education and clinical practice.

1.1. Theoretical Framework

An ethical exploration of AI-based accessibility should be based on the principles according to which the human user and his or her specific neurological profile should be in the center of design and implementation. The analysis is based on three primary theoretical fields: Human-Computer Interaction (HCI) and Inclusive Design, Neurodiversity-Centered Communication Theory, and Sensory Processing Considerations in Autism.

1.2. Human–Computer Interaction and Inclusive Design

The Human Activity Assistive Technology (HAAT) model is the most effective approach to the utility of any assistive technology (AT) (Cook and Polgar, 2015; Elsahhar et al., 2019). This model highlights the fact that the effectiveness of AT depends on the congruence between the activity one is trying to perform (self-care, productivity, leisure) and the capabilities of the human user (physical, cognitive, emotional), the usage scenario, and the qualities of the device (Elsahhar et al., 2019). Most importantly, the HAAT model stipulates that the choice of technology ought to be based on the communicative needs and the available skills of the individual user (Iacono et al., 2014; Omoyemi, 2024). It is anticipated that AI-powered solutions will support the principles of inclusive design, provide solutions that are intuitive, customized, and responsive to the individual preferences of the user, therefore empowering people with a disability (Kooli and Chakraoui, 2025). Moreover, a progression to the ability-based design strategy, where emphasis is placed on abilities and not on deficits, is crucial so that the user experience is inherently empowering (Wobbrock, 2017).

1.3. Neurodiversity-Centered Communication Theory

The Neurodiversity movement advocates for the acceptance of neurological differences as natural variations of the human genome (Kapp et al., 2013). Applying this lens requires researchers and clinicians to move beyond the assumption that differences in autistic communication are necessarily deficits (de Marchena et al., 2025). Communication for autistic individuals, especially those who are minimally verbal, is often inherently **multimodal**, relying on a complex mix of limited speech, paralinguistics, and non-verbal cues, such as gestures (Curtis et al., 2024; de Marchena et al., 2025).

A critical concept here is the **Double Empathy Problem**, which suggests that communication breakdowns between autistic and non-autistic individuals frequently occur because of a reciprocal difficulty in understanding differing social communication styles (Szechy et al., 2024). AI-driven tools, such as advanced gesture and voice recognition, can effectively function as a **mediator or translator**, scaffolding existing multimodal abilities and turning often misunderstood nonverbal communication into conventional linguistic output, thus mitigating the friction caused by the double empathy problem (Curtis et al., 2024; Ullah et al., 2023).

1.4. Sensory Processing Considerations in Autism

There is a sizeable amount of literature that records the peculiar sensory processing patterns of autistic people, such as distorted audio-visual information processing and increased sensitivity (Cassidy et al., 2016; Robertson and Baron-Cohen, 2017; Stevenson et al., 2018). Namely, hyper-sensitivity to pitch (Haigh et al., 2022) and audio-visual integration during the speech may become a problem. Sensory overload is a symptomatic issue that is commonly reported as a hindrance to communication and engagement in such settings as healthcare (Strömberg et al., 2022).

Such sensory realities therefore can not be left out when designing AI assistive tools. The interfaces should include options to allow users to manipulate sensory options to minimize sensation overload, including the speed, volume, voice gender, and visual properties of text-to-speech interfaces (Curtis et al., 2024; Voultsiou and Moussiades, 2025). Sensory-friendly alignment should be ensured as the basics of ensuring that these innovations are usable, acceptable, and actually empowering instead of causing distress or system abandonment (Cassidy et al., 2016).

1.5. Review of Related works

The efficacy of AI-driven tools is rooted in their capacity to accurately process diverse input modalities. The following section critically reviews selected empirical studies focusing on how STT and gesture recognition systems, frequently powered by machine learning (ML), have been applied to communication support for individuals with communication disabilities, including those on the autism spectrum.

Kambouri et al. (2023) conducted a foundational group-based intervention study in the UK, utilizing commercial Speech-to-Text (STT) software, Dragon, to improve writing skills in thirty children with Special Educational Needs and Disabilities (SEND) who experienced difficulties in spoken and written communication . The results demonstrated that the use of STT technology significantly **boosted the quantity and quality of handwritten text** and resulted in positive and statistically significant improvements in self-esteem. The post-test screen-written text was significantly superior to handwritten text. This study, while focusing on a broader SEND population, supports the feasibility of STT to circumvent motoric and linguistic challenges inherent in traditional writing, directly supporting individuals who struggle with verbal output .

Ullah et al. (2023) proposed a body-worn multi-sensor Internet of Things (IoT) platform specifically designed to recognize the complex sign language and gestures of speech-impaired children with Autism Spectrum Disorder (ASD) . Recognizing that only a subset of people understands traditional sign language, the platform placed multiple sensors on the body to capture complex movements . By comparing various classification techniques, including the Artificial Neural Network (ANN) and Random Forest, the researchers achieved greater than **96% recognition accuracy** . This study provides strong empirical evidence that fusion-based, body-worn sensing platforms can successfully recognize and translate complex, non-standardized gestures, offering a crucial communication bridge for speech-impaired autistic individuals .

In a study included within the systematic review conducted by Kambouri et al. (2023), reported on the positive effects of speech-to-text software, specifically the Dragon system, on the written expression of three students with various disabilities aged 7–16 . This work affirmed the high functional **accuracy** and benefit of commercial STT systems when used as accommodations for producing written output for individuals with communication disorders . This finding reinforces the notion that even standard STT technology, when properly implemented, can overcome expressive language barriers.

Omoyemi (2024) proposed a novel Machine Learning (ML) framework aimed at enhancing AAC systems by integrating multimodal data: predictive text, advanced speech recognition, and gesture recognition . The central hypothesis was that ML could provide **faster, more accurate, and contextually relevant communication assistance** than traditional AAC . This study notably addressed the challenge of processing speech patterns "outside standard norms" and advocated for leveraging computer vision to interpret non-verbal gestural communication, identifying multimodal sensing as key to addressing the restrictions and slow responsiveness of conventional AAC devices .

Cassidy et al. (2016) focused on the potential of **expressive visual text-to-speech (TTS)** as an assistive technology for adults with Autism Spectrum Conditions (ASC) . This research addressed the recognized difficulty autistic adults often have in recognizing and appropriately responding to the emotions of others . By exploring a system that generates visual speech alongside text, the study moves beyond mere transcription to address socio-communicative skills, positioning AI to facilitate emotion recognition, a critical component often impaired in ASD

Saraswat (2021) investigated the use of Natural Language Processing (NLP) techniques for sign gesture identification, with the goal of translating sign language video frames into appropriate English sentences . This work is significant because it highlights the fundamental technical process of converting complex visual, gesture-based input into linguistic output, a core requirement for supporting non-verbal autistic individuals who communicate through gesture . The methodology relies on computer vision techniques (like CNN and SVM) for classification, demonstrating the viability of creating systems that bridge the communication gap for individuals relying on gesture-based communication .

1.6. Emerging Insights

The AI trend in assistive communication predetermines a number of essential research and development directions, which could be aimed at making sure that technological advances would address the needs of the non-verbal autistic population in the most refined way.

1.7. Need for Multimodal Sensing

The multimodal form of language, which combines vocals, gestures and facial expressions and posture, has been the focus of contemporary communication studies (de Marchena et al., 2025). This understanding should be capitalized upon in the future of AI-based accessibility as it should not be based on the systems that focus on speech or gesture alone (Omoyemi, 2024). Multimodal sensing will help users with complex requirements (including those with both speech and physical disabilities) to communicate effectively through malleable input strategies such as mixed voice inputs and gestures (Kooli and Chakraoui, 2025). Because it is possible to merge data of multiple sensors on the body, nowadays the high accuracy is achieved and the complex gestures that a single sensor system cannot recognize are accurately interpreted (Ullah et al., 2023). This aims at a combined strategy capable of handling outliers in vocal inputs and nonverbal communication in real-time to precisely intend the user, hence, providing rich communication outputs in a contextual manner (Omoyemi, 2024).

1.8. Sensory-Friendly Interface Design

The familiar sensory processing includes in autistic people need interface design to shift its functionality to actively being sensory-friendly (Voultsiou and Moussiades, 2025). To eliminate the risk of sensory overload, highly customizable sensory settings must also be included in the development in the future (Strömberg et al., 2022). The prototype testing has already given user feedback on the parameters that will be essential to the control, such as speed and pitch of Text-to-Speech (TTS) voices, the need to adjust text color, and the size of buttons on the interface to address motor issues (Curtis et al., 2024). Moreover, since the dynamics of audio-visual integration is often atypical in autism (Stevenson et al., 2018), the interdependence between the visual feedback (e.g., expressive avatars in the TTS systems) and audio output will have to be dialled to the dot to create a comfortable and effective understanding (Cassidy et al., 2016).

1.9. Diverse, Autism-Specific Datasets

The intrinsic soundness of the machine learning models depends solely on the training data quality and representativeness (Omoyemi, 2024; Ullah et al., 2023). The primary challenge to creating a strong AI to enable access to non-verbal autistic individuals is the dramatic lack of appropriate, varied, and autism-specific datasets (Kooli and Chakraoui, 2025; Yusuf et al., 2025). There are intrinsic ethical and logistical issues in the collection of data among children and adults having neurodevelopmental disorders (Yusuf et al., 2025). The existing studies frequently work with small groups of people and restricted variations in gestures, which undermines the potential of the model to be generalized to be applicable in various levels of ASD, linguistic, and personal peculiarities (Nadeem et al., 2025; Ullah et al., 2023). Thus, one of the future directions is the creation of large, ethically sourced datasets, which may involve special sensors or computer vision algorithms, that would cover the entire spectrum of minimally verbal patterns of communication (Ullah et al., 2023; Nadeem et al., 2025).

1.10. Interpretable and Clinically Aligned Models

To implement AI-based AAC that will be trusted by the parents, therapists, and educators, the underlying models should be changed to the non-black box systems that can be interpreted and explained (Atlam et al., 2024). Clinical alignment requires the justification of an AI model, e.g. why a particular utterance has been challenging to transcribe or why a gesture was categorized in a particular way, to be provable and explainable (Choi et al., 2025). It aims at creating open AI models that would be applied not only in translating communications but also as a diagnostic or therapeutic aid, and deliver useful information on the communication pattern of the user that would meet the standard clinical measures and psychometric tools (Atlam et al., 2024).

2. Discussion

The introduction of sophisticated AI into AAC systems is a turning point towards increasing the quality of life and independence of non-verbal autistic people. Nevertheless, to exploit this possibility, it is required to take a sober evaluation of both the achieved strengths, as well as the technical and ethical constraints that remain in place.

2.1. Strengths and Limitations

The ability of AI to provide personalisation in real-time and increased responsiveness is the strongest feature in this respect (Kooli and Chakraoui, 2025; Omoyemi, 2024). Artificial intelligence provides diverse and custom-made solutions (Kooli and Chakraoui, 2025). The empirical evidence of this theory is that AI-enhanced STT can enhance the amount and quality of written text (Kambouri et al., 2023). Multi-sensor gesture fusion systems, on the non-verbal front, are shown to be remarkably accurate in decoding both difficult, functional movements and this offers a reliable source of expression where words are not used (Ullah et al., 2023). Moreover, the ASR use can offer instant and precise feedback, which bypasses the difficulties users tend to experience with monitoring speech production (Ballard et al., 2019; Kooli and Chakraoui, 2025).

In spite of these achievements, there are a number of constraints that hinder general effectiveness. The nature of the Automatic Speech Recognition (ASR) technology also means that it is ineffective with the inputs that imply a speech distortion, replacement, or fragmentation, which is a characteristic feature of some communication disorders, as the performance of the ASR technology is based on clear and consistent speech (Omoyemi, 2024; Radford et al., 2023; Sanguedolce et al., 2023). In addition to technical precision, infrastructural needs, such as compatibility of hardware and the need to have a well-developed technical support, are also problematic, restricting access in under-resource systems (Esquivel et al., 2024; Kooli and Chakraoui, 2025). Lastly, as mentioned earlier, the lack of high-quality training data is also a major obstacle to providing models that represent the Autistic community in terms of their linguistic and gestural diversity (Nadeem et al., 2025).

2.2. Ethical Issues: Privacy, Sensory Burden, and Bias

The AI ethical situation in the field of special education is complicated and requires urgent and stringent efforts on the part of researchers, policymakers, and industry participants (Kooli and Chakraoui, 2025).

2.2.1. Privacy and Data Security

The development of multimodal data, including voice, gesture kinetics, physiological responses, to train the models raises significant questions of privacy and data protection as the list of users is frequently a minor and a vulnerable adult (Yusuf et al., 2025). The possibility of advanced mechanisms to read and process an extremely customized information, like pointing patterns or emotional indications (Hamidi et al., 2018), requires the use of strong regulatory protection and complete disclosure of data collection and storage (Kooli and Chakraoui, 2025). Also, there exists an ethical conflict of dependence and deskilling; excessive dependence on the technology may cause the loss of the natural sign language or the non-digital communication abilities of the user (Saraswat, 2021). The solutions should be promoted as improvements and not substitutes to human interaction.

2.2.2. Sensory Burden and Cognitive Load

According to the theoretical framework, systems developed without considering neurodiversity may pose significant sensory load to the user (Stromberg et al., 2022). In case the results of the TTS in a system are too rapid or rough, or the interface is overstimulating, the technology can be discarded, even in terms of technical precision (Curtis et al., 2024). Moreover, preliminary studies on voice assistants indicate that their utilization and configuration may place a significant cognitive burden particularly on users with cognitive dissimilarities, and this could be because of the unavailability of flexible nonverbal control choices (Esquivel et al., 2024). The developers should invest in intensive testing on autistic participants to ensure that interfaces are optimally sensory-comfy and the least amount of mental work is required.

2.2.3. Algorithmic Bias

The most toxic ethical issue is, perhaps, the problem of algorithmic bias, which creates a threat of repeating and spreading systemic discrimination (Kooli and Chakraoui, 2025; Lyerly, 2023). Unless models are conditioned on non-verbal autistic datasets of speech or mainstream sign language, they will always classify or overlook the communication patterns peculiar to non-verbal autistic individuals (Ullah et al., 2023). Research indicates that AI systems may be biased towards persons with disabilities (Lyerly, 2023). This needs to be mitigated by inclusive standards of design that will involve the joint participation of the autistic self-advocates, community members and clinical experts during the design, development, and evaluation stages (Guo et al., 2020; Kooli & Chakraoui, 2025). AI should not be an exclusion source since it does not comprehend difference.

3. Conclusion

This paper has demonstrated that the implementation of AI-based Speech-to-Text and Gesture Recognition technologies is an incredible potential to make non-verbal autistic people more accessible, independent, and included. These technologies go further and transcend the constraints of traditional AAC by providing more dynamic, personalized, and responsive communication channels that can integrate these complex multimodal inputs (Omoyemi, 2024). Among the crucial findings are the current technical ability to have high precision when deciphering both vocal and nonverbal communication (Kambouri et al., 2023; Ullah et al., 2023).

However, it is through a rigorous effort of scaling the prototypes into scalable, ethically responsible clinical solutions that the process will see the reduction of the issues documented. The issue of strong, heterogeneous, and neurodiversity-specific data sets is directly connected to the main problem of accuracy (Nadeem et al., 2025). Simultaneously, in order to achieve its success, one will need to adhere rigidly to the principles of creating sensory-friendly, and ethical frameworks that will take into account the risk of algorithmic bias, threat to privacy, and cognitive load (Esquivel et al., 2024; Kooli and Chakraoui, 2025).

The future research priorities should be strategic and aim at:

- Longitudinal Efficacy Studies: The researchers will carry out a large-scale study of diverse socio-economic and cultural settings to determine the long-term effectiveness of AI on learning, well-being, and digital literacy (Kooli & Chakraoui, 2025; Nadeem et al., 2025).
- Standardized Policy and designer Frameworks: Jointly setting up cross-country policy systems and all-embracive architecture standards that require adjustable sensory interfaces and solid moral protection (Kooli & Chakraoui, 2025; Voultsiou and Moussiades, 2025).
- Multimodal Framework Development: Keeping exploring the extent to which sensor fusion and AI processing of vocalization and gestures and other non-verbal cues can be deeply integrated to create communication output that reflects the fluidity and complexity of natural discourse (Omoyemi, 2024).
- Making Autistic Experience Central: Making sure that research, development, and evaluation of the same tools always integrate the lived experience and agency of autistic adults and least verbally expressive people to ensure that the technology is indeed a tool of empowerment (Kooli and Chakraoui, 2025; Nadeem et al., 2025).

AI-powered assistive technologies can live up to their hype and become not only a catalyst for human wellbeing and social justice in a non-verbal community of autistic people, but also a valuable addition to the group of innovations, by prioritizing care, equity, and neurodiversity-affirming principles in their design (Kooli and Chakraoui, 2025).

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclsoed.

References

[1] Adams, J. (2017). *The effectiveness of using speech-to-text technology to support writing of students with learning disabilities*. Master's dissertation, Rowan University, US

[2] Alam, I., Hameed, A., & Ziar, R. A. (2024). Exploring Sign Language Detection on Smartphones: A Systematic Review of Machine and Deep Learning Approaches. *Advances in Human-Computer Interaction* .

[3] Almufareh, M. F., Kausar, S., Humayun, M., & Tehsin, S. (2024). A conceptual model for inclusive technology: advancing disability inclusion through artificial intelligence. *Journal of Disability Research*, *3*(1), 20230060

[4] Atlam, E.-S., Masud, M., Rokaya, M., Meshref, H., Gad, I., & Almars, A. M. (2024). Easdm: Explainable autism spectrum disorder model based on deep learning. *Journal of Disability Research*, *3*(1), 20240003

[5] Ballard, K. J., Etter, N. M., Shen, S., Monroe, P., & Tien Tan, C. (2019). Feasibility of automatic speech recognition for providing feedback during tablet-based treatment for apraxia of speech plus aphasia. *American Journal of Speech-Language Pathology, 28*(2S), 818–834

[6] Bondy, A., & Frost, L. (1994). *PECS: the Picture Exchange Communication System training manual*. Pyramid Educational Consultants

[7] Bruce, C., Edmundson, A., & Coleman, M. (2003). Writing with voice: An investigation of the use of a voice recognition system as a writing aid for a man with aphasia. *International Journal of Language & Communication Disorders, 38*(3), 263–274

[8] Cassidy, S. A., Stenger, B., Van Dongen, L., Yanagisawa, K., Anderson, R., Wan, V., Baron-Cohen, S., & Cipolla, R. (2016). Expressive visual text-to-speech as an assistive technology for individuals with autism spectrum conditions. *Computer Vision and Image Understanding, 148*, 193–200

[9] Cheng, X., Jin, T., Huang, R., Li, L., Lin, W., Wang, Z., Wang, Y., Liu, H., Yin, A., & Zhao, Z. (2023). Mixspeech: Cross-modality self-learning with audio-visual stream mixup for visual speech translation and recognition. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 15735–1574

[10] 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*

[11] Cook, A. M., & Polgar, J. M. (2015). *Assistive Technologies Principles and Practices* (4th ed.). Elsevier

[12] 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*

[13] 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)

[14] Elsahhar, 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

[15] 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*

[16] Garrett, T., Heller, W., Fowler, P., Alberto, A., & Fredrick, D. (2011). Using speech recognition software to increase writing fluency for individuals with physical disabilities. *Journal of Special Education Technology, 26*(1), 25–41

[17] Guo, A., Kamar, E., Vaughan, J. W., & Wallach, H., Morris, M. R. (2020). Toward fairness in AI for people with disabilities. *ACM SIGACCESS Accessible Computing*

[18] Haigh, S. M., Brosseau, P., Eack, S. M., Leitman, D. I., Salisbury, D. F., & Behrmann, M. (2022). Hyper-Sensitivity to Pitch and Poorer Prosody Processing in Adults With Autism: An ERP Study. *Frontiers in Psychiatry, 13*(844830)

[19] Hamidi, F., Poneres, K., Massey, A., & Hurst, A. (2018). Who should have access to my pointing data? Privacy trade-offs of adaptive assistive technologies. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*, 203–216

[20] Higgins, E. L., & Raskind, M. H. (1995). The state of research on computers and writing for students with learning disabilities. *Learning Disabilities: A Multidisciplinary Journal, 6*(2), 65–78

[21] Hodges, H., Fealko, C., & Soares, N. (2024). Autism spectrum disorder: Definition, epidemiology, causes, and clinical evaluation. *Translational Pediatrics, 13*(1), 25–45

[22] Iacono, T., Lyon, K., Johnson, H., & West, D. (2014). Experiences of adults with complex communication needs. *Augmentative and Alternative Communication, 30*(2), 163–175

[23] Jaliaawala, M. S., & Khan, R. A. (2020). Can autism be catered with artificial intelligence-assisted intervention technology? A comprehensive survey. *Artificial Intelligence Review, 53*(2), 1039–1069

[24] 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

[25] Kapp, S. K., Gillespie-Lynch, K., Sherman, L. E., & Hutman, T. (2013). Deficit, difference, or both? Autism and neurodiversity. *Developmental Psychology, 49*(1), 59–71

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

[27] Lee, Y. Y. (2016). *Using speech recognition software to improve independent (fluency) writing of students in a mainstream primary school*. Master's Dissertation, University College London Institute of Education

[28] Lyerly, E. (2023). Study finds AI language models show bias against individuals with disabilities. *Disability Compliance for Higher Education, 28*(12), 9

[29] MacArthur, C. A., & Cavalier, A. R. (2004). Dictation and speech recognition technology as test accommodations. *Exceptional Children, 71*(1), 43–58

[30] Manwaring, S. S., Swineford, L., O'Reilly, B., Meyer, S., Dager, S., Lee, D. H., & Tager-Flusberg, H. (2018). Characteristics of gesture use in children with autism spectrum disorder and developmental delay. *International Journal of Language & Communication Disorders, 53*(5), 966–979

[31] McCollum, D., Nation, S., & Gunn, S. (2014). The effects of a speech-to-text software application on written expression for students with various disabilities. *National Forum of Special Education Journal, 25*(1), 1–13

[32] Millar, D. C., McNaughton, D. B., & Light, J. C. (2005). A comparison of accuracy and rate of transcription by adults with learning disabilities using a continuous speech recognition system and a traditional computer keyboard. *Journal of Postsecondary Education and Disability, 18*(1), 12–22

[33] 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 Technology & Interaction, 9*(98)

[34] Noakes, M. (2018). Does speech-to-text assistive technology improve the written expression of students with traumatic brain injury? *Dissertation Abstracts International Section A: Humanities and Social Sciences, 79*(1-A(E))

[35] 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

[36] Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2023). Robust speech recognition via large-scale weak supervision. In *International Conference on Machine Learning*, 28492–28518

[37] Raskind, M. H., & Higgins, E. L. (1999). Speaking to read: the effects of speech recognition technology on the reading and spelling performance of children with learning disabilities. *Annals of Dyslexia, 49*, 251–281

[38] Robertson, C. E., & Baron-Cohen, S. (2017). Sensory perception in autism. *Nature Reviews Neuroscience, 18*(11), 671–684

[39] Sanguedolce, G., Naylor, P. A., & Geranmayeh, F. (2023). Uncovering the Potential for a Weakly Supervised End-to-End Model in Recognising Speech from Patient with Post-Stroke Aphasia. In *Proceedings of the 5th Clinical Natural Language Processing Workshop*, 182–190

[40] Saraswat, P. (2021). NLP-based Sign Gesture Identification for Disabled People. *International Journal of Innovative Research in Computer Science & Technology, 9*(6), 41–45

[41] Smith, E. G., & Bennetto, L. (2007). Audiovisual speech integration and lipreading in autism. *Journal of Child Psychology and Psychiatry, 48*(8), 813–821

[42] Stevenson, R. A., Segers, M., Ncube, B. L., Black, K. R., Bebko, J. M., Ferber, S., & Barense, M. D. (2018). The cascading influence of multi-sensory processing on speech perception in autism. *Autism, 22*(5), 609–624

[43] Strömberg, M., Liman, L., Bang, P., & Igelström, K. (2022). Experiences of Sensory Overload and Communication Barriers by Autistic Adults in Health Care Settings. *Autism in Adulthood, 4*(1), 66–75

[44] Svensson, I., Nordström, T., Lindeblad, E., Gustafson, S., Björn, M., Sand, C., Almgren/Bäck, G., & Nilsson, S. (2021). Effects of assistive technology for students with reading and writing disabilities. *Disability and Rehabilitation: Assistive Technology, 16*(2), 196–208

[45] Szechy, K. A., Turk, P. D., & O'Donnell, L. A. (2024). Autism and Employment Challenges: The Double Empathy Problem and Perceptions of an Autistic Employee in the Workplace. *Autism in Adulthood, 6*(2), 205–217

[46] 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

[47] Ullah, F., Ullah, A., Siddiqui, U. A., Paracha, S., Iqbal, A., Khan, R., & Kwak, K. S. (2021). Wearable-Sensors-Based Platform for Gesture Recognition of Autism Spectrum Disorder Children Using Machine Learning Algorithms. *Sensors,*

[48] Valencia, S., Cave, R., Kallarackal, K., Seaver, K., Terry, M., & Kane, S. K. (2023). "The less I type, the better": How AI Language Models can Enhance or Impede Communication for AAC Users. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*

[49] 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

[50] Wobbrock, J. O. (2017). SIGCHI Social Impact Award Talk – Ability-Based Design: Elevating Ability over Disability in Accessible Computing. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 5–7

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