



## Augmentative and Alternative Communication for Children with Communication Disorders Using Speech Recognition Machine Learning

*Parameswari .M<sup>1</sup>, Nivetha .S<sup>2</sup>, Preethiba .E<sup>3</sup>, Priyadharshini .R<sup>4</sup>, Narmatha .R<sup>5</sup>*

Professor<sup>[1]</sup>, UG Student<sup>[2,3,4,5]</sup>

[parameswarirajkumar@gmail.com](mailto:parameswarirajkumar@gmail.com)<sup>[1]</sup>, [narmatha5696733@gmail.com](mailto:narmatha5696733@gmail.com)<sup>[5]</sup>

### ABSTRACT :

Alternative and Augmentative Communication (AAC) is an area of clinical practice for people with severe expressive communication disorders. This paper introduces a solution for simplifying speech recognition comparison. integration of JavaScript and AI speech Recognition module. system employs advanced algorithms to conduct real-time comparison of voices.

Keywords— ML, Javascript, Communication disorder, Speech recognition.

### Introduction :

According to the National Institute on Deafness and Other Communication Disorders, approximately one in twelve children in the United States has a disorder related to voice, speech, language, or swallowing [1]. These communication disorders can stem from various conditions such as autism spectrum disorder (ASD), cerebral palsy, intellectual disability, hearing loss, traumatic brain injury, and others. For children affected by these disorders, effectively expressing thoughts, needs, and feelings through speech and language can be profoundly challenging.

Augmentative and alternative communication (AAC) encompasses methods and tools designed to supplement or replace speech and writing for those with impairments in producing or comprehending spoken or written language [2]. AAC strategies involve using gestures, sign language, picture communication boards/books, speech-generating devices, and other systems [3]. While beneficial, traditional AAC approaches have notable limitations. Physical communication aids can be cumbersome, lack flexibility for naturalistic communication, and require significant time and effort from professionals to create individualized resources.

Recent breakthroughs in automatic speech recognition (ASR) driven by neural network advances have opened up possibilities for transformative AAC solutions. By leveraging these technologies, we can develop systems capable of flexibly and accurately transcribing the diverse vocalizations, approximations of speech, and multimodal communication attempts of children with disorders. This transcribed input can interface with natural language processing (NLP) components to interpret the intended meaning, account for context, and enable naturalistic communication and language modeling tailored to each child's abilities and needs.

The proposed AAC system aims to significantly enhance communication abilities and independence for children across a wide range of speech and language disorders. It has the potential to reduce barriers, increase engagement, and create an empowering communication experience that dynamically adapts to each child's unique needs and communication modalities.

## 2. Background and Related Work :

### A. Communication Disorders in Children

Communication disorders encompass a range of challenges related to speech, language, and communication abilities. Some of the most common disorders impacting children include:

**Autism Spectrum Disorder (ASD):** Individuals with ASD often experience difficulties with social communication, including challenges with verbal and non-verbal communication, initiating and sustaining conversations, and understanding contextual cues [4].

**Cerebral Palsy:** This neurological disorder can affect muscle control and coordination, leading to difficulties with speech production, articulation, and intelligibility [5].

**Intellectual Disabilities:** Children with intellectual disabilities may have impairments in language development, comprehension, and expression, as well as challenges with social communication and pragmatic language use [6].

**Hearing Loss:** Children with hearing loss can experience delays or difficulties in speech and language development, as well as challenges in communicating effectively with others [7].

**Traumatic Brain Injury:** Depending on the area of the brain affected, traumatic brain injuries can result in various communication disorders, including aphasia, dysarthria, and cognitive-communication deficits [8].

These conditions can significantly impact a child's ability to communicate effectively using natural speech, necessitating the use of AAC strategies and tools.

### ***B. Traditional AAC Methods***

AAC methods have traditionally involved the use of low-tech and high-tech solutions to support communication for individuals with disabilities or impairments. Some common AAC approaches include:

**Picture Communication Boards/Books:** These low-tech AAC tools use pictures, symbols, or photographs to represent words, phrases, or concepts, allowing individuals to communicate by pointing or gesturing towards the desired symbols.

**Speech-Generating Devices (SGDs):** SGDs are electronic devices that can generate synthetic or digitized speech output. Users can select words, phrases, or symbols on the device, which then "speaks" the corresponding message.

**Sign Language:** Sign languages, such as American Sign Language (ASL), use hand shapes, movements, and facial expressions to convey meaning without relying on spoken language.

**Gestures and Body Language:** For individuals with limited speech abilities, gestures, body language, and facial expressions can be used to convey basic needs, emotions, or ideas.

While these traditional AAC methods have provided valuable support, they also have several limitations, including:

- Limited flexibility and naturalistic communication
- Reliance on pre-programmed vocabularies or symbol sets
- Difficulty adapting to individual communication needs and styles
- Potential stigma or self-consciousness associated with using physical communication aids
- Significant time and effort required from professionals to create individualized resources

These limitations have motivated researchers and developers to explore more advanced and personalized AAC solutions leveraging technological innovations.

### ***C. Related Work in AAC Technology***

Researchers have explored various technological solutions to enhance AAC capabilities and address the limitations of traditional methods. Some relevant works include:

- [9] describes a mobile AAC app that utilizes speech recognition along with word and phrase prediction to facilitate communication. However, it requires clear speech and does not adapt to unusual vocalizations or unconventional speech patterns.

- [10] presents an AAC system incorporating computer vision to detect gestures and sign language, but it lacks capabilities for interpreting vocalizations and does not use machine learning for personalization.

- [11] discusses the possibility of using brain-computer interfaces (BCIs) for AAC by detecting neural signals associated with intended speech. While promising, this approach requires specialized medical-grade equipment and training data.

- [12] proposes an AAC system that combines speech recognition, natural language processing, and adaptive language modeling to improve communication for individuals with dysarthric speech. However, the focus is on adults rather than children with diverse communication disorders.

Our work aims to leverage recent advancements in machine learning for speech recognition and natural language processing to create an AAC solution specifically tailored to the unique communication needs and modalities of children with a wide range of disorders. By incorporating personalization and continual adaptation capabilities, the proposed system can provide a more flexible, naturalistic, and empowering communication experience for these children.

---

## **3. LITERATURE REVIEW**

- Implementation of web application based on augmentative and alternative communication method for people with hearing and speech impairment by Khasanboy kodirov, Khasanboy kodirov, Young Sil Lee

This paper describes an AAC software developed that helps both disabled children and adults to communicate easily with other people especially in office work, school or social gatherings.

- Augmentative and alternative communication system using information priority and retrieval by Yoonseok Heo, Sangwoo Kang

Our proposed method can generate the required sentence using minimum keystroke input and reduced input errors. Here, the system takes the lead in suggesting symbol candidates to the user. In our experiments, we use a decision tree.

- A mega-review of literature reviews, systematic reviews, and meta-analyses on interventions using aided augmentative and alternative communication (AAC) interventions for children with intellectual and developmental disabilities from 2000 to mid-2020 was conducted by Becky Crowe, Wendy Machalicek, Qei Wei

Augmentative and Alternative Communication for Children with Intellectual and Developmental Disability: A Mega-Review of the Literature

➤ Augmentative And Alternative Communication For Children With Communication Disorders by Sai Aishwarya Ramania, Amudhu Sankara  
This device was administered on forty children with various communication disorders and whose language age ranges from two to four years. It was used to assess the accessibility and easy usage of the device.

## ***B. PROBLEM DESCRIPTION:***

In today's digital age, the reliance on electronic documents and transactions has surged, necessitating efficient methods for verifying voice comparison. However, the process of comparing voices manually can be cumbersome and prone to errors, leading to potential security vulnerabilities and delays in authentication. Existing solutions often require specialized software or complex procedures, limiting accessibility for users and hindering widespread adoption. Moreover, the lack of user-friendly interfaces and real-time feedback mechanisms further complicates the verification process. Therefore, there is a pressing need for a simplified and intuitive solution that leverages JavaScript's versatility to streamline voice recognition comparison. By addressing these challenges, the e-voice matcher seeks to enhance document security, improve authentication processes, and provide users with a seamless experience for verifying voices in digital environments.

---

## **4. Proposed System Architecture**

The proposed AAC system consists of the following key components:

### ***A. Speech Acquisition***

This module is responsible for capturing the vocalizations, speech attempts, and other audio input from the child user via a microphone or similar audio input device. The speech acquisition component should be designed to handle diverse acoustic environments and accommodate potential mobility or physical limitations of the child user.

### ***B. Speech Recognition Engine (ASR)***

The core of the system is a speech recognition engine powered by deep neural network models optimized for transcribing the unique speech patterns and communication modalities of children with disorders.

Unlike conventional ASR systems trained on clear, typical speech from adults, our models will be specifically trained on a diverse dataset of audio recordings and corresponding transcripts from children with various communication disorders, such as ASD, cerebral palsy, and intellectual disabilities. This specialized training data will include vocalizations, approximations of speech sounds, atypical prosody, and other communication modalities commonly exhibited by this population.

The speech recognition engine will employ state-of-the-art neural network architectures for acoustic modeling and language modeling, such as:

- Sequence-to-sequence models with attention mechanisms [13]
- Hybrid CTC/attention models [14]
- Transformer-based encoder-decoder architectures [15]
- Personalized language models adapted to each child's communication style
- Additionally, the ASR models will have mechanisms to continually learn and adapt to the unique speech patterns, utterances, and sounds produced by each individual child user over time through techniques such as transfer learning, personalized language modeling, and online adaptation strategies.

### ***C. Natural Language Understanding (NLU)***

The text output from the speech recognition engine will be processed by natural language understanding components to interpret the intended meaning, take into account contextual information, and facilitate accurate communication.

### ***D. Methodology:***

The methodology used in the Voice Matcher for simplifying handwritten voice comparison involves several key steps:

1. **Data Collection:** Handwritten voice samples are collected from individuals for both reference voices and input voices to be compared.
2. **Feature Extraction:** Various features of the voices are extracted, such as stroke direction, curvature, pen pressure, and timing information. These features help in quantifying the unique characteristics of each voice.

3. **Normalization:** The extracted features are normalized to ensure consistency and comparability across different voices. Normalization helps in removing variations in writing styles and sizes.
4. **Matching Algorithm:** A matching algorithm is applied to compare the features of the input voice with those of the reference voices. Common matching algorithms include Euclidean distance, dynamic time warping, or neural network-based approaches.
5. **Thresholding:** A threshold is set to determine the similarity between the input voice and the reference voices. If the similarity score exceeds the threshold, the voices are considered a match.
6. **Verification and Validation:** The matched voices are further verified and validated to ensure accuracy and reliability. This may involve additional checks or confirmation steps to confirm the authenticity of the matched voices.
7. **Result Interpretation:** The results of the voice comparison are interpreted and presented to the users. This may include a percentage indicating the similarity between the voices and a verification status (e.g., verified or not verified).

## 5. Conclusion:

In conclusion, the integration of a speech recognition system using JavaScript and AI speech recognition modules for children's speech analysis represents a significant advancement in pediatric healthcare. By leveraging cutting-edge technology, this system enables accurate and efficient assessment of speech patterns and characteristics in children, allowing healthcare professionals to identify speech impairments or developmental delays with precision. Subsequently, this information empowers doctors and speech-language pathologists to tailor therapy programs and interventions according to each child's specific needs, thereby optimizing the effectiveness of treatment. Ultimately, the implementation of this system holds the potential to significantly improve outcomes for children with speech-related challenges, ensuring they receive timely and targeted interventions to support their speech development and overall well-being.

## OUTPUT

S. NO	AGE	DISORDER	SCORE
1	3	CEREBRAL PALSY	10
2	3	DOWN'S SYNDROME	20
3	6	STUTTERING	30
4	6	DYSARTHRIA	10
5	4	OROFACIAL MYOFUNCTIONAL DISORDER	40
6	4	AUTISM SPECTRUM DISORDER	30
7	4	CEREBRAL PALSY	40
8	4	DOWN'S SYNDROME	50
9	6	STUTTERING	30
10	4	DYSARTHRIA	10
11	6	OROFACIAL MYOFUNCTIONAL DISORDER	20
12	6	AUTISM SPECTRUM DISORDER	10
13	6	CEREBRAL PALSY	30
14	6	DOWN'S SYNDROME	20
15	3	STUTTERING	20
16	6	DYSARTHRIA	40
17	6	OROFACIAL MYOFUNCTIONAL DISORDER	10
18	5	AUTISM SPECTRUM DISORDER	10
19	4	CEREBRAL PALSY	30
20	3	DOWN'S SYNDROME	40
21	5	STUTTERING	20
22	6	DYSARTHRIA	40

23	3	OROFACIAL MYOFUNCTIONAL DISORDER	10
24	4	AUTISM SPECTRUM DISORDER	30
25	3	STUTTERING	30

OUTPUT TABLE

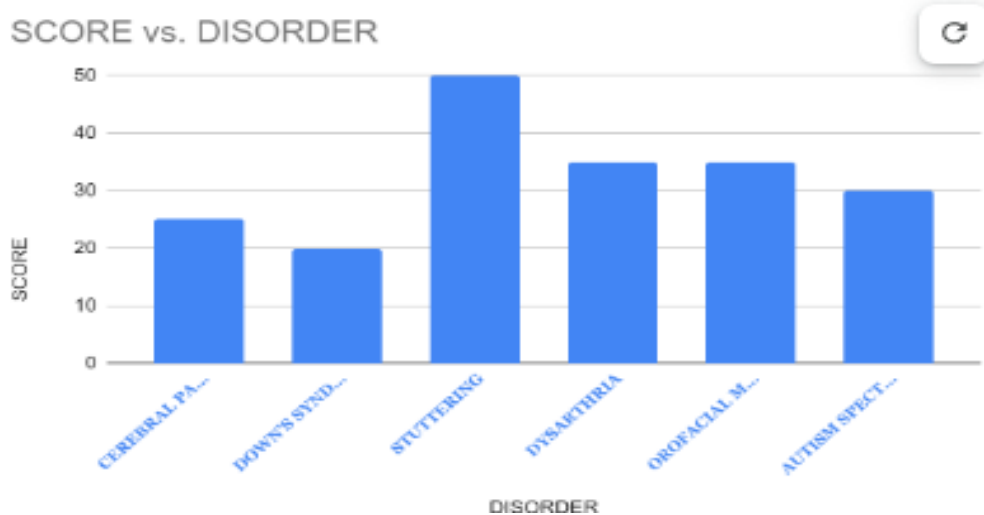


Fig. Overall Analysis

## 6. FUTURE ENHANCEMENTS

Many points can be considered in future work to conduct research, such as increasing the number of reference voice images in the case of offline systems or voice features in the case of online systems for each user to improve a machine learning system's decision. The second point is to create a multilingual voice dataset with a large number of users because as we notice in this paper, some voice dataset have a limited number of users (usually ranging from 10 to 100 users).

Future enhancements for the Voice Matcher system for simplifying handwritten voice comparison using the Doodle Net algorithm with JavaScript could include:

1. **Improved User Interface:** Enhance the user interface to make it more intuitive and user-friendly, allowing for easier voice input and result interpretation.
2. **Real-Time Feedback:** Implement real-time feedback mechanisms to provide users with immediate insights into the quality and confidence level of the voice matches as they input voices.
3. **Integration with Biometric Authentication:** Explore integration with biometric authentication technologies to further enhance security and authentication capabilities, ensuring that only authorized voices are accepted.
4. **Mobile Application Development:** Develop a dedicated mobile application version of the Voice Matcher system, allowing users to conveniently access and utilize the voice comparison functionality on their smartphones or tablets.
5. **Multi-Language Support:** Extend support for voices in various languages and writing styles, enabling users from diverse linguistic backgrounds to utilize the system effectively.

6. **Advanced Reporting and Analysis:** Introduce advanced reporting and analysis features to provide users with detailed insights and trends regarding voice matches, helping organizations identify patterns and anomalies more effectively.
7. **Machine Learning Model Optimization:** Continuously optimize the machine learning models used in the Doodle Net algorithm to handle large-scale voice datasets more efficiently and accurately.

---

**REFERENCES :**

---

- [1] De Wachter, M., et al., Template-based continuous speech recognition. *IEEE Transactions on Audio, Speech, and Language Processing*, 2007. 15(4): p. 1377-1390.
- [2] Resch, B., Automatic Speech Recognition with HTK. Signal Processing and Speech Communication Laboratory. Inffeldgase. Austria. Disponible en Internet: <http://www.igi.tugraz.at/lehre/CI>, 2003.
- [3] Kekre, H., A.A. Athawale, and G. Sharma. Speech recognition using vector quantization. in *Proceedings of the International Conference & Workshop on Emerging Trends in Technology*. 2011. ACM.
- [4] Moore, R.K., Twenty things we still don't know about speech, in *Proc. CRIM/FORWISS Workshop on Progress and Prospects of speech Research an Technology*. 1994.
- [5] Tebelskis, J., Speech recognition using neural networks. 1995, Siemens AG.
- [6] Pan, Y., P. Shen, and L. Shen, Speech emotion recognition using support vector machine. *International Journal of Smart Home*, 2012. 6(2): p. 101-108.
- [7] Anusuya, M. and S.K. Katti, Speech recognition by machine, a review. arXiv preprint arXiv:1001.2267, 2010.
- [8] Alhawiti, K.M., Advances in Artificial Intelligence Using Speech Recognition. *World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 2015. 9(6): p. 1397-1400.
- [9] Furui, S., Theory and Applications. 1 ed. *Speech Technology*, ed. K.J. Fang Chen. US: Springer. XXVII, 331.
- [10] Furui, S., 50 years of progress in speech and speaker recognition. *SPECOM*