

# A survey on Artificial Intelligent based solutions using Augmentative and Alternative Communication for Speech Disabled

Betsybha Evangeline (✉ [betwork24@gmail.com](mailto:betwork24@gmail.com))

Anna University

Anitha -

Thiagarajar College of Engineering

---

## Research Article

**Keywords:** Communication Disorders, Artificial Intelligence, Augmentative and Alternative Communication, Speech Generating Devices, Speech Recognition, Human-Computer Interaction

**Posted Date:** November 4th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-2225081/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

This paper aims to analyse how innovative Artificial Intelligence (AI) systems for non-standard speech recognition may revolutionize Augmentative Alternative Communication (AAC) technology for people with severe speech impairments. The AI-based system is personalized for each person's unique speech production and offers a real step forward in improving the efficiency of AAC. With impressive enhancements in recognizing non-standard natural language supported by Machine Learning and Deep Learning algorithms, AI is offering a turning point for personalized and customized Augmentative Alternative Communication (AAC). However, there is a need of understanding the contextual needs of the user which enhances the experience while using AAC. To address this, a systematic review has been done to identify existing applications and the technologies behind them. Also, challenges are explored that can lead to the future directions of the research work. This paper presents the results of the systematic review, after filtering 62 journal articles that are more relevant to the context from a pool of 1088 papers, in the perspective of the existing AI technologies in speech generation, speech reconstruction and speech generation. AI based AAC applications are comprehended in term of input/output, cost, infrastructure and user background. The paper also highlights the future research directions in this domain by identifying the research gaps.

## 1. Introduction

Speech is the form of expression or the ability to express thoughts and feelings by articulating sounds. A person who finds it difficult to communicate verbally is said to have a speech disorder. A person with a speech disorder can experience lowered self-esteem and overall quality of life. Speech disorders can be categorised into two: speech impairment and speech impediments.

Speech impairments are disorders of speech sounds, fluency, or voice that interfere with communication. Speech impairment can be categorized as follows:

- Fluency disorder, which can be described as an unusual repetition of sounds or rhythm, Examples of fluency disorders include childhood apraxia of speech happens when children start to talk, Dysarthria is a speech disorder that happens when the muscles that are used to talk are too weak and can't form words, and Orofacial Myofunctional Disorders(OMD) is categorised by an abnormal pattern of facial muscle use.
- Voice disorders have an atypical tone of voice. It could be a cause of unusual pitch, quality, resonance, or volume. Examples of voice disorders include stuttering, which means the repetition of words or sounds like "uh" and "um" (disfluencies) involuntarily. And it can be caused by a chronic cough or paralysis of the vocal cords, among others.
- Articulation disorder is the distorting of certain sounds or the fully omitting of sounds. Examples of articulation disorders include speech disorders for children who struggle as they learn to talk and pronounce certain sounds.

Speech impediments occur when the person's mouth, jaw, tongue, and vocal tract can't work together to produce recognizable words. Examples of speech impediments include stuttering, which indicates a developmental delay, Articulation errors which happen when a person can't form speech sounds because they have trouble placing their tongue in the right position, Tongue-tie (ankyloglossia) is a physical condition that makes it hard for the person to move their tongue, Apraxia is a condition that happens when a person's brain can't coordinate the muscles that enable speech, Dysarthria is a disorder that occurs when children slur their words because of brain damage, Selective mutism is a complaint that ensues when a person becomes so anxious about being in certain places and situations that they can't speak, Cerebral Palsy (CP) and Parkinson's disease occur when a person has degenerative muscle and motor conditions that can lead to speech impediments.

The potential causes of these disorders include

- Stroke
- Head Trauma
- Brain Tumor
- Lyme Disease
- Facial Paralysis, Such As Bell's Palsy
- Tight or Loose Dentures
- Alcohol Consumption

These causes cause the disorders to occur in people of all ages, from toddlers to adults. This might be due to an accident or brain damage, or even due to a developmental disability caused by differences in the brain. Some treatments are given to overcome speech impairments and impediments, like speech therapy exercises that focus on building familiarity with certain words or sounds and physical exercises that focus on strengthening the muscles that produce speech sounds. In the early days before the technological era, these speech and physical therapies were given in the form of traditional Face to Face therapy that involved sign language and gestures since its conception. It involves a patient travelling to a Therapy clinic/Center to obtain an assessment or treatment by a Speech-Language Therapist (SLT). The SLT used Printable worksheets, age-old flashcards, and activities as an *"alternative communication"* method. This method involves trained professionals and practical issues like travelling from one place to another to get therapy and the time taken to be trained is maximised. This has led to technological advancements called *"Augmentative and Alternative Communication"* (AAC) which have offered promising ways to enhance the communication and treatment of speech and language impairment by using technological advancement of digitized devices. These digitized devices that are available these days are more advanced and user-friendly. Many Artificial Intelligence-based Mobile/iPad Applications which employ AAC methods have become popular over the last decade. Such applications have made AAC technology accessible to millions of patients all over the world. Thanks to technological advancements and computational needs at a lower cost, which can produce speech by a digitized device for communication.

The rest of this paper is organized as follows. Section 2 describes AAC with the tools and strategies that this AAC uses and its limitations when people use low technology aids for communication. Section 3 emphasises the systematic review on the topic "Augmentative and Alternative Communication for the Speech Disabled." Section 4 explores the AAC applications that have been in existence years back and at present, along with the details of the people who use which application for which type of medical condition; the methodology that the particular application emphasizes; and the cost-effectiveness of the applications; along with the age group who can use this application. Section 5 illustrates the literature review on the basis of the algorithms used, the communicating partners, the cost-effectiveness of the solutions, and the hardware software requirements with the perspective of how well the high technology helps the speech disabled recognize their speech and communicate, whereas Section 6 elaborates on the future directions on what is to be done following the conclusions made from the literature review and Section 7 concludes the paper.

## 2. Augmentative And Alternative Communication

Augmentative and Alternative Communication is a means of communication besides talking. Those who have trouble with speaking skills use AAC. AAC stands for Augmentative and Alternative Communication. The word "augmentation" means *"to add something to someone's speech"*. Alternative means *"to be used in the place"* of speech. Therefore AAC means it is a communication method used to supplement or replace speech for those with impairments in the production or comprehension of spoken language. People of all ages who are born with speech disorders use AAC all their lives to communicate.

AAC includes all the tools and strategies a person can use to communicate when they are not able to speak. There are different types of communication media for the speech-disabled using AAC. It can be grouped into unaided and aided AAC.

- Unaided AAC is without the help of physical aid or tools. Examples of unaided AAC are facial expressions, body language, gestures, and sign language.
- Aided AAC devices are categorized as No-technology types of equipment, Low-technology instruments, and High-technology devices. The no-technology types of equipment include things like gestures and facial expressions, writing, drawing, and spelling words by pointing to letters; and the low-technology instruments include pointing to photos, pictures, or written words using picture cards and flashcards; and high-technology devices include a wired or wireless device with an application on an iPad or tablet to communicate with a "speech" as an output to communicate. It is called a "speech-generating device," technically termed an AAC device.

In this technological era, the use of AAC is proliferating, which helps or equips the person who is speech disabled to communicate and socialize with everyone around them who is known and unknown to them. AAC uses tools or materials like symbol boards, choice cards, communication books, Pragmatic Organisation Dynamic Display (PODD) books, keyboards, alphabet charts, and Speech-generating devices or communication devices. A text-based AAC system works with the keyboard by giving the text

as input to speech as output, in short Text-To-Speech (TTS). Symbol-based AAC is used by many people who use pictures or symbols to communicate, these symbols are in the form of pictograms, where when these pictograms are selected by a touch response as the input and speech are generated as the output. Finally, multimodal communicators have multiple ways to communicate their messages by using the Linguistic/Alphabetic mode which is written and spoken words, Visual modes like images that are moving or still, Aural mode which includes sound, music, and the Gestural mode which involves movement, expression, and body language. In this developing technological world, the AAC is in the form of mobile applications and web-based applications which require the use of a high-tech tool (e.g., Speech Generating Device - SGD or AAC app on an iPad/Android device) for communication.

The benefits of using high-technology AAC include improving the speech of AAC users for their language development, augmenting communication and literacy, and increasing the quality of work in general life. According to research by Millar et al. (2006), 89% of the participants who are AAC users improved their speech skills using AAC tools. Picture exchanges have also been proven to improve the language development of children with speech and language disabilities. It is shown that language and hearing research suggests that since AAC focuses on encouraging and providing communication, it reduces the psychological stress of individuals regarding the necessity of speaking and eases the speech development process. It is supported by the researchers Millar DC et al., (2016), Richard et al., (2016), Agius, M et al., (2019) that high-tech AAC also has a positive outcome in terms of self-expression in social situations, language development, and literacy. Children with speech disorders need help with developmental skills such as vocabulary, the length of sentences, syntax, and pragmatic skills. [44] Over 2 million persons who present with significant expressive language impairment use AAC.[45] Therefore high-tech AAC is a good source of communication.

The following section describes the systematic review done in order to show how the articles have been selected and reviewed on the topic "Augmentative and Alternative Communication Apps for Speech Disabled".

### **3. Systematic Review**

The aim was to identify the performance of AI in AAC for Speech Disabled and to provide a comprehensive overview of the state-of-the-art technology in terms of mobile technology. This was based on a systematic search conducted between June and August 2022 of published articles available on PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>), Springer (<https://link.springer.com/>), ASHA (<https://www.asha.org/>), Semantic Scholar (<https://www.semanticscholar.org/>), Elsevier (<https://www.elsevier.com/en-in>), ResearchGate (<https://www.researchgate.net/>), IEEE (<http://www.IEEE.org/>) using the phrase of the title. The searches were conducted twice for each repository.

The collection of reviews has been done from six different repositories, they are Pubmed, Springer, ASHA (American Speech-Language-Hearing Association), Semantic Scholar, Elsevier, ResearchGate and IEEE

from the time period between 1990–2022. The review articles and the papers published in English have been taken. The Instances of Search Keywords included in title: "Speech Disorder" + "Speech Disabled" + "AAC" + "Augmentative and Alternative Communication" + "Mobile Applications" + "Artificial Intelligence" + "Speech Recognition" + "Speech Reconstruction" + "Speech Generation".

The following Table 1 explains the selection criteria for the inclusion and exclusion of research articles.

Table 1  
The Inclusion and Exclusion criteria set for selecting the article of the required study.

<b>Specification</b>	<b>Inclusion Criteria</b>	<b>Exclusion Criteria</b>
Type of Research Articles	IC1. Research articles in peer-reviewed journals, conferences, and scientific research	EC1. Book reviews, thesis dissertations, opinion articles, lecture notes, and papers.
The Objective or Aim of the Research Article	IC2. Research articles that aim to enhance the communication skills of people with speech disorders using augmentative and alternative communication.	EC2. Research articles that focused on the development of other skills and screening processes, in particular for diseased patients, are not included.
Language	IC3. Articles or scientific research papers utilizing English as the mean of writing	EC3. Research articles in any other language other than English are excluded.
Database	IC4. Pubmed, Scopus, Springer, ASHA, Elsevier, Research Gate, Semantic Scholar, IEEE, and ScienceDirect	EC4. Research articles foreign to the included database and journals as well as the database is not available to the researchers'
Publication Time Frame	IC5. Research articles and papers focusing on communication for the speech disabled and AAC-based applications published from January 2017 to articles about the present	EC5. Research articles and papers focusing on communication in people with speech disorders and AAC-based applications published before January 2017 are not included.

The following is a comparison of articles in various repositories for the specified topic "Augmentative and Alternative Communication for Speech Disabled + Speech Recognition + Speech Reconstruction + Speech Generation"+ "Mobile Applications".

Table 2  
 “Augmentative and Alternative Communication for Speech Disorders”

S.No	Repository	Year	Total Articles	Published for the last 10 years	Articles Taken
1	Pubmed	1990–2022	487	242	14
2	Springer	2005–2022	9	7	7
3	ASHA	2002–2022	424	25	11
4	Semantic Scholar	1990–2022	71	35	5
5	Elsevier	1990–2022	15	11	11
6	Research Gate	1990–2022	54	25	9
7	IEEE	1999–2022	28	14	5
TOTAL			1088	359	62

The results of Table 2 have been illustrated in the following Fig. 2 shows the details of the number of articles present in the topic searched, the number of publications for the past 10 or 5 years, and how many articles have been taken for the literature review. The topics searched for are “Augmentative and Alternative Communication for Speech Disabled.” Out of 1088 research articles, 62 articles have been taken for study by focusing on the Aided AAC communication for the speech disabled that has been available for the past decade of their existence for the purpose of communication for the speech disabled. The Fig. 3 represents the graphical representation of the 62 articles selected for the study which is categorised in a 5 year ranges. These articles have been selected based on the inclusion criteria given in Table 1. This systematic review acknowledges the trend in AAC interventions as well as integrated features of AAC-based mobile applications for people with speech disabilities to enhance their language, learning, and emotional skills, which are linked to communication and social competence development. To achieve the research goal, 62 research articles were categorized and investigated. The findings of this systematic review can be beneficial to further exploring the potential of AAC interventions or mobile applications to augment communication ability in people with speech disorders. Moreover, it has been identified that Artificial Intelligence (AI) processes in mobile applications or systems have substantial potential to facilitate adaptability, contextual flexibility, data analysis, storage, and dissemination of information for advanced research.

The articles selected from the systematic review have been taken for further literature study. The following section describes some of the applications of AAC that are available for the speech disabled.

## 4. Mobile And Web Applications Of Aac

There are mobile and web applications for adults and children with developmental and speech disabilities. The user has the liberty to express their emotional states, needs, and opinions when using an AAC application. Table 3, presented below, explains the name of the applications that are being used for the required medical condition along with the methodology that is used. The application's pricing, then their ease of use of what tools or materials are used, followed by the age category of the people who use these applications, then the platform that is created, along with the popularity of end-user usage of the applications in terms of rating stars out of 5 and the purpose behind the applications that are being discussed.

In the following discussed mobile and web-based applications in Table 3, there is a wide usage of maximum usage of picture cards, flashcards and Text-to-speech methods. Additionally, the processes that are used, such as voice detection, reconstruction, and generation, are explained. The cost of the applications varies from free of cost to a maximum of \$300. The age group of people who uses these applications are from infants to adults. Few of these applications are built on iOS and Android platforms. The popularity of the applications is rated from 0 to 5, where a few applications like Autisay, DISSERO, VITHEA, and MOSOCO are not in use due to the limitations of the design work, the applications are not supported by the most common mobile device's browsers, whereas few applications rated from 1 star to 5 stars. These applications help disabled people to facilitate speech and language development, social communication, mental and social skills, communicate with non-verbal people, translate intangible pronunciation into understandable speech, and improve their speaking skills. It is also noted that the following applications are widely present in English and a few applications are also present in regional language communication.

## 5. Review Findings

Augmentative and Alternative Communication describes ways to support, brace, or amplify the communication of people who are not verbal communicators. Professionals with expertise in AAC, like Occupational Therapists, Speech and Language Pathologists, and Rehabilitation Technologists, use these unique AAC projects. For an AAC device speech recognition, speech reconstruction and speech generation play a vital role in helping speech-disabled persons to communicate. The results of the review findings have been categorized under various parameters which includes input and output that is sub-categorized as speech recognition, speech reconstruction and speech generation that speaks about the different types of input and output components that can be and that are used in the existing solutions, followed by the user background which states about the different age group of people who uses the AAC devices that varies from partially speech-disabled to fully speech-disabled, and infrastructure that describes about the hardware and the software components that are used to build the various web-based or mobile-based solutions along with its cost effectiveness of the solutions.

### 5.1 Input and Output:

An input/output device is any hardware that allows a human operator or other systems to interface with a computer. Input/output devices, as the name implies, are capable of delivering data (output) to and receiving data from a computer (input). The input devices can be a touch screen, camera, keyboard, mic, etc., whereas the output device can be a monitor screen, speaker, a printed paper, etc. The following AAC devices that are mentioned below are the mobile-based and the web-based devices. For the mobile-based device the screen of a smartphone or tablet computer is an input and output device. The input on the screen is through touch, and it outputs through graphics on the screen. Whereas for the web-based device the input device can be a camera, where it can capture the data and process it to the processor of a system or the keyboard where the input can be given as a typed text. And the output of the web-based can be the monitor screen of a system of the printed paper copy from the system. Here are some of the different AAC devices where these kinds of input and output components are being used.

Caron et al. (2017), suggested a few AAC apps like GoTalk Now, AutisMate, and EasyVSD that are used by adults with speech-disabled have pictograms-based touch responses and the mobiles keyboard request as the input mode and the visual scene display as their output method. It is noted that there is no regional-based pictograms are used.

Holyfield et al. (2020), propose that AAC technology with dynamic text for the single-word recognition of adults with intellectual and developmental disabilities like down syndrome, ASD and cerebral palsy; uses the dynamic display of text in conjunction with voice output as the output. The generalization of text communication is being used.

Vlachou, J. A., & Drigas. (2017), recommends Mobile Technology for Students & Adults with Autistic Spectrum Disorders (ASD) with the iPad touch response and the speech generating device for various purpose. Like as communication, academics, and entertainment.

Nakkawita et al. (2021), introduces the Proloquo2Go and Speak for Yourself mobile applications for persons with aphasia. These applications use typing and the pictogram-based touch response as input modes and speech as the output.

Janice et al. (2019), insist on visual scene displays (VSD) or grid-based AAC apps with the transition to literacy (T2L) supports (i.e., dynamic text paired with speech output upon selection of a hotspot) for individuals who require AAC who are preliterate.

The following Fig. 4 depicts the existing mobile applications and their input methods. Where we can see the collection of pictograms of how it is being used. The use of regional and conceptualized pictograms is absent in these applications.

## **5.2 Algorithms:**

An algorithm is a predetermined process for resolving a clearly stated computing issue. All facets of computer science, including artificial intelligence depend on the creation and analysis of algorithms. A finite set of instructions followed in a certain order to complete a task is known as an algorithm. In the

following session algorithms are categorised under different process like speech recognition, speech reconstruction, and speech generation. Some of the mentioned algorithms are mainly called as Machine Learning algorithms like Hidden Markov Method (HMM), Natural Language Processing (NLP), Case-Based Reasoning (CBR), Augmented Reality (AR), Surface ElectroMyoGraphy (sEMG), Random Forest (RF), Support Vector Machine (SVM), Dynamic Time Wrapping (DTW), Mel Frequency Cepstrum Coefficients (MFCCs), Principal Component Analysis (PCA).

(a). Speech Recognition:

E. Essa et al. (2008), I. Mohamad et al. (2016), and M. Rahman et al. (2018) have proposed a speech recognition system based on a Support Vector Machine (SVM) with the assistance of Dynamic Time Wrapping DTW for speakers of the isolated Bangla language, isolated Arabic words, and for voice-matching. The data has been collected from 40 speakers. The Mel Frequency Cepstrum Coefficients (MFCCs) were used as the static features of the speech signal. And they used the DTW algorithm after determining the feature vectors for the feature matching. They proposed a model that tested 12 speakers, and 86.08% was the recognition rate that the system achieved. Principal Component Analysis (PCA) is also been used for the extraction of images and MFCC for the extraction of voice inputs. The system achieved an accuracy of 87.14% with the DTW method and 92.85% with the Euclidean distance method. The system could achieve very high rates of speech recognition for isolated Arabic words of up to 96% using MFCC.

Y. Mittal et al. (2015) worked (on a voice-controlled smart home based on an IoT connection that could be installed easily and at a low cost. The authors suggested that using wireless technology can further be used for better performance. Withal, their solution suffers from authentication lockage and distance problems; the closer the user is to the sound device, the more suitable the system is. Additionally, the developed mobile application system can solve the distance problem, which requires that users be within a certain distance of the microphone.

Geoffrey et al. (2012), and Mark S et al. (2013), have used a Large Vocabulary Continuous Speech feature using Hidden Markov models. Here, HMM is used to deal with the temporal variability of speech and Gaussian Mixture Models (GMMs) to check how well each state of each Hidden Markov Model (HMM) fits a frame or a short window of frames of coefficients that represent the acoustic input. This gave an accuracy of 88%. The device can be configured to enable the user to create either simple or complex messages using a combination of a relatively small set of input "words". They have used the MFCC and HMM toolkit with the Baum-Welch algorithm and have got a recognition accuracy of 67%.

The challenges of existing applications that are used by the speech-disabled exhibit that there is no existence of speech as input for communication. This made the way for my research future directions. Table 4 compares the algorithms used by the researcher by giving the recognition rate, and determining which algorithm fits well for speech recognition. It is said that MFCC works well by giving the highest accuracy.

(b). Speech Reconstruction:

Several papers highlight the issues of conceiving VIVOCA (Voice Input Voice Output Communication Aid) for users with severe speech disorders and dysarthria [25–27]. In [28], Hawley et al. present a VIVOCA solution as a form of AAC device for people with severe speech disabilities: the researchers have applied the statistical Automatic Speech Recognition (ASR) techniques, based on HMMs, to dysarthric speakers to produce speaker-dependent recognition models.

W. Farzana et al.(2020), Uchoa et al. (2021), and Neamtu et al. (2019), highlight the use of Machine Learning (ML) and AI in communication in AAC, that an application named LIVOX is a machine learning-based android mobile application to recommend pictograms based on the location and time of the user device. The notable features of LIVOX are an artificial intelligence-based recommendation system by analysing past user data (used item, utilization time, touch time, GPS data, X and Y coordinates on touch screen), along with the random forest classification algorithm and another feature is Natural Language Processing (NLP) for speech recognition. Moreover, a feedback framework which applies to the Case-Based Reasoning (CBR) machine.

Gay et al. (2011) and Costa et al. (2017), explore their research with 1 in 68 children having Autism Spectrum Disease (ASD). Their intention is to develop a task recommendation. The system will use a Case-based Reasoning machine learning technique in order to supplement the child's regular therapy. Besides the tasks' recommendation, this application will allow closer monitoring by parents and better coordination with the therapists, contributing to improving the results of the child's development.

Liu et al. (2017), insist on Augmented Reality (AR), which may be especially effective as it may deliver visual and auditory cues while the user is simultaneously engaged in natural or structured social interactions. Additionally, wearable technologies contain sensors that can be used to record and quantitatively monitor a user's interaction. Their report highlights the need for further research into the use of augmented reality technologies as a therapeutic tool for people with ASD.

Elsahar et al. (2019), review give a range of signal sensing and acquisition methods used in conjunction with the existing high-tech AAC platforms for individuals with a speech disability, which includes imaging methods, touch-enabled systems, mechanical and electro-mechanical access, breath-activated methods, and Brain–Computer Interfaces (BCI). The listed AAC sensing modalities are differentiated in terms of ease of access, affordability, difficulty, and typical conversational speeds. A revelation of the associated AAC signal processing, encoding, and retrieval highlights the roles of Machine Learning (ML) and Deep Learning (DL) in the development of intelligent AAC solutions.

Geoffrey S et al. (2017), for training phoneme-based recognition models, a unique set of phrases were used for each of the 39 commonly used phonemes in English, and the remaining phrases were used for testing the word recognition of the models based on phoneme identification from running speech. This study provides a compelling proof-of-concept for sEMG-based laryngeal speech recognition, with the strong potential to further improve recognition performance.

(c). Speech Generation:

Clients who understand the language(s) spoken in their communities would arguably benefit from AAC systems that are aligned with these languages. Considerations may include the form of the AAC symbols. For literate clients making use of orthographic symbols the link between the written and spoken forms of the language is typically clear. Multilingual clients would merely need access to the relevant orthographic symbols for the different languages in which they would want to communicate. If access to synthetic speech generation is also desired, speech generation engines in the relevant spoken languages would need to be available. The development of such speech engines is a technically challenging task that has been accomplished for the global majority languages but is still lagging for many minority languages Calteaux et al. (2013); Louw, (2008); Titmus et al. (2016). A language agnostic speaker embedding mechanism for cross-lingual personalized speech generation. The proposed network adopts an encoder-decoder architecture to disentangle the language information from the latent representation via multi-task learning Zhou, X et al. (2021).

Table 4 compares the algorithms that are used to recognize and reconstruct speech to build an AAC device. the process flow of speech recognition and speech reconstruction along with the algorithms used under the respective process by reflecting the recognition accuracy rate and the purpose of usage of AAC is correlated.

Table 4  
Comparison of Algorithms used to recognize and reconstruct speech, to build AAC device

Process Flow	Algorithm Used	Purpose/ Medical Condition	Recognition Rate
Speech Recognition	SVM, DTW, MFCC	Isolated Bangla speech recognition	86.08%
	MFCC	speech recognition system for isolated Arabic words	96%
	PCA, MFCC, DTW	Voice matching of Hijaiyah Letters	87.14%
	HMM	Voice-controlled multi-functional smart home application	88%
	MFCC HMM	Portable, voice output communication aid controllable by automatic speech recognition	67%
	Baum-Welch algorithm.		
Speech Reconstruction	HMM	Dysarthric	Not Found
	RF, NLP, CBR	Non-verbal people	Not Found
	CBR	Autism	Not Found
	AR	Autism	Not Found
	NLP	Non-verbal people	Not Found
	sEMG	Learning Disability	Not Found

Hidden Markov Method (HMM), Natural Language Processing (NLP), Case-Based Reasoning (CBR), Augmented Reality (AR), Surface ElectroMyoGraphy ( sEMG), Random Forest (RF), Support Vector Machine (SVM), Dynamic Time Wrapping (DTW), Mel Frequency Cepstrum Coefficients (MFCCs), Principal Component Analysis (PCA).

### 5.3 User Background:

The users of these AAC applications vary from infants to adults and from partially speech-disabled to fully speech-disabled people.

An et al. (2017), ten 3–6-year-old children included in their study who had been diagnosed with ASD may or may not be able to communicate using speech or language. Based on criteria in the Diagnostic and Statistical Manual of Mental Disorders, had not been diagnosed with or suspected of having epilepsy, and had never previously used any other speech-generating devices or Picture Exchange Communication System (PECS).

C Wilson et al. (2018), undertook a field study over two school terms in an autism-specific primary school with 12 minimally verbal children aged 5 to 8 with their teachers and speech therapists. In Human-Computer Interaction (HCI), Ability-based Design addresses the importance of interactive technologies which adapt to users' abilities, skills, and contexts, as opposed to catering solely to their perceived inabilities.

Tariady et al. (2017), the subjects of their research were 12 autistic children in an inclusion school in the Pekalongan region. The experiment method was used with the single-subject research approach. The research concluded that the average ability level in communication before and after the treatment was 47%, while during the treatment the average level was 65%. Whereas there is an improvement after the intervention stage of the PECS method Multimedia Augmented Reality-Based with an average of 76%.

Shelley et al. (2017), motive for this paper is to explore how Speech-Language Pathologists (SLPs) who are AAC specialists approach the assessment process for 2 case studies, 1 child with cerebral palsy and 1 with autism spectrum disorder. The results of the current study provide an outline of an assessment protocol for children with complex communication needs.

Sarah Creer, et al. (2016), this paper provides an epidemiological approach to determine the prevalence of people who could benefit from AAC in the UK. As a result, a total of 97.8% of the people who could benefit from AAC have nine medical conditions: dementia, Parkinson's disease, autism, learning disability, stroke, cerebral palsy, head injury, multiple sclerosis, and motor neurone disease. The total expectation is that 536 people per 100,000 of the UK population (approximately 0.5%) could benefit from AAC.

The study of AAC on implementation shows that it has been checked with a limited count of speech-disabled people varying from infants to adults.

## **5.4 Infrastructure:**

The infrastructure of an AAC device refers to the digital technologies that provide the basic functionality for a device's operations. Examples of infrastructure include hardware and software components along with the cost-effectiveness required for the AAC device.

Castro et al. (2018), iCommunicate, which is a Sign Language Translator Application System using Kinect Technology. The C#, Kinect Studio Visual Gesture Builder, and Kinect for Windows V2 are the software and hardware used in this system. Language-Integrated Query (LINQ) makes the strongly-typed query to be a first-class language construct. As an object-oriented language, C# supports the concepts of encapsulation, inheritance, and polymorphism.

Ramon et al. (2017), A device with KY-038 Microphone Sound Sensor Module, WTV020M01 MP3 voice module SD, and Gizduino ATmega 168 with the LCD is implemented, whereas the program is implemented in a microcontroller as software to make the AAC device work.

N.H. Voon et al. (2015), The AutiSay application is developed for the hand-held smart device Apple iPad technology and requires iOS 6.1 or later. This device is equipped with a built-in camera, microphone, and Wi-Fi capabilities, which the application needs for customization and accessing the app store.

Haneen et al. (2015), the application will be downloaded on the teacher's/parent smartphones to use the setting page and also on the children's smartphones to use the pages provided to them by the application. The GPS service provided by the phones will be used to determine the children's location.

M.Manickavelu and Anil, (2013), KAVI-PTS are an AAC application built using the Android OS. It can be installed on any Android tablet or smartphone. It requires 2.3.3 or above Android OS with 512 MB RAM Memory and internet connectivity is preferred to download the application from the server.

While AutiSay: A Mobile Communication Tool for Autistic Individuals is focused on the autistic child, there is significant potential for the application to be scaled for adult autistic individuals N.H. Voon et al. (2015). The application is relatively inexpensive as it is distributed on the Apple App Store for a relatively low price in comparison to similar applications.

Voigt et al. (2014), Delayed Auditory Feedback (DAF) is a mobile application that is developed as an inexpensive app that contains a lot of insights we have gained in our research on stuttering during the last 10 years as the hardware or device expense is minimised.

Muharib et al. (2018), GoTalk Now is a relatively inexpensive application as it is an open-source application. Using iPad-based SGD may be feasible in educational settings. In addition to educational settings, family members may be taught to use an iPad as an SGD to support their children's communication at home and in the community.

Trishna et al. (2017), speak about Avaz, an iOS-based application that helps autistic kids by providing picture-based communication and natural voices in multiple languages. Roger Voice is a free-of-cost cross-platform application for people with difficulty hearing. Using Roger Voice, people with hearing problems can comprehend phone calls by reading. Vaakya is a free-of-cost Android-based Augmentative and Alternative Communication application which is helpful for people with speech disabilities, especially for autism, cerebral palsy, etc. The application contains a combination of images and phrases, through which users can communicate effectively in the Hindi language.

The above Fig. 5 represents the cost of the existing mobile application of AAC followed by the future directions of research and conclusion.

## 6. Future Directions

The many solutions that are now on the market are reviewed throughout the study, including their potential use of input/output components and the algorithms that were used to create them. Next, the study describes how the solutions are used by users of all ages, before moving on to a discussion of the hardware and software components that were used to create the devices/solutions and their cost of

production. But still there are some limitations that are need to be addressed from the research findings, they are, (1). the review findings show that the AAC applications available display the scarcity of the use of, the combination of speech recognition and speech reconstruction on a single device that is used to communicate, for the partially and fully speech-disabled. (2). The input and the output modes of the existing applications generalize that the pictograms or the text buttons used as a touch response is lacking from regional context and lack conceptualizing. (3). The cost-effectiveness has to be maintained. (4). Lack of local regional language speech-generating applications. (5). There is only a limited source of vocabulary, often with a heavy reliance on the fringe (nouns), Positions of words change in picture exchange systems. These limitations are to be overthrown by future research work. An AI-based AAC solution is to be developed for the partially and fully speech disabled with the combination of speech recognition and speech reconstruction as a single solution that helps the speech-disabled people to communicate by generating a speech, and also to help them in speech therapy for the partially speech disabled. As the existing solutions do not support speech recognition it is to be implemented with the regional and contextual communicating data.

## 7. Conclusion

In this review article, a brief introduction about speech disorders and Augmentative and Alternative Communication is given with the comparison of the AAC mobile and web-based applications used for various speech disorders with the systematic review of articles in AAC. In the review findings, it is found that there are various speech disorders affected by different situations. The results indicate how AAC is effective and useful for the person who is affected by speech. With the usage of the latest techniques like PECS (Picture Exchange Card System), TTS (Text-To-Speech), Picture Cards, Flash Cards, Image Animation, and Augmentation, it is made easy for people who are speech disabled to communicate.

Finally out of these findings, an AI-based solution is to be created to recognise the user's speech, rebuild the words, and then generate the speech, taking into account the restrictions discussed in Section VI of the future directions. Further research suggests a method that also reads speech as input and recognises it, followed by the contextualised prediction and reconstruction to provide the proper speech output for communication to overcome the restrictions, using the contextualised and region-based collection of pictograms.

## References

1. Miller, S. E. Investigating the Advantages and Disadvantages of Augmentative and Alternative Communication (AAC) Use in Individuals with Autism Spectrum Disorder (ASD).
2. Wendt, O., Hsu, N., Simon, K., Dienhart, A., & Cain, L. (2019). Effects of an iPad-based speech-generating device infused into instruction with the picture exchange communication system for adolescents and young adults with severe autism spectrum disorder. *Behavior modification*, 43(6), 898-932.

3. Voon, N. H., Bazilah, S. N., Maidin, A., Jumaat, H., & Ahmad, M. Z. (2015). Autisay: A mobile communication tool for autistic individuals. *In Computational Intelligence in Information Systems* (pp. 349-359). Springer, Cham.
4. Hani, H., & Abu-Wandi, R. (2015, November). DISSERO Mobile Application for AUTISTIC Children's. *In Proceedings of the international conference on intelligent information processing, security and advanced communication* (pp. 1-6).
5. Muharib, R., Correa, V. I., Wood, C. L., & Haughney, K. L. (2019). Effects of functional communication training using GoTalk Now™ iPad® application on challenging behavior of children with autism spectrum disorder. *Journal of Special Education Technology*, 34(2), 71-79.
6. Neamtu, R., Camara, A., Pereira, C., & Ferreira, R. (2019, September). Using artificial intelligence for augmentative alternative communication for children with disabilities. *In IFIP Conference on Human-Computer Interaction* (pp. 234-243). Springer, Cham.
7. Pompili, A., Abad, A., Trancoso, I., Fonseca, J., & Martins, I. P. (2020, March). Evaluation and extensions of an automatic speech therapy platform. *In the International Conference on Computational Processing of the Portuguese Language* (pp. 43-52). Springer, Cham.
8. Escobedo, L., Nguyen, D. H., Boyd, L., Hirano, S., Rangel, A., Garcia-Rosas, D., ... & Hayes, G. (2012, May). MOSOCO: a mobile assistive tool to support children with autism practicing social skills in real-life situations. *In Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 2589-2598).
9. Krcek, T. E. (2015). Effectiveness Of Proloquo2go Tm In Enhancing Communication In Children With Autism During Aba Therapy.
10. Castro, M. D. B. iCommunicate: A Sign Language Translator Application System using Kinect Technology.
11. Viner, M., Singh, A., & Shaughnessy, M. F. (2020). Assistive technology to help students with disabilities. *Special Education Design and Development Tools for School Rehabilitation Professionals*, 240-267.
12. Wang, E. H., Zhou, L., Chen, S. H. K., Hill, K., & Parmanto, B. (2018). Development and evaluation of a mobile AAC: a virtual therapist and speech assistant for people with communication disabilities. *Disability and Rehabilitation: Assistive Technology*, 13(8), 731-739.
13. Salwerowicz, D. A. (2019). Design Proposal for a Software Tool for Speech Therapy. *Modern Application Structure of Visual Speech Therapy App for Children* (Master's thesis, UiT Norges arktiske universitet).
14. Satterfield, B. (2013). Studies in AAC and Autism: The Impact of LAMP as a Therapy Intervention.
15. Voigt, T., Hewage, K., & Alm, P. (2014, April). Smartphone support for persons who stutter. *In IPSN-14 Proceedings of the 13th International Symposium on Information Processing in Sensor Networks* (pp. 293-294). IEEE.
16. Binti Rosli, S. (2019, December). The Use of Alternative and Augmentative Communication (AAC) In Classroom Settings: A Case Study. *In the 3rd International Conference on Special Education (ICSE*

- 2019) (pp. 311-314). Atlantis Press.
17. Anand, A., Nair, R. R., Kodamanchili, S., Panda, R., Bhardwaj, K. K., & Gowthaman, T. B. (2022). Communication with Patients on Mechanical Ventilation: A Review of Existing Technologies. *Indian Journal of Critical Care Medicine: Peer-reviewed, Official Publication of Indian Society of Critical Care Medicine*, 26(6), 756.
  18. J. Shor, D. Emanuel, O. Lang, O. Tuval, M. Brenner, J. Cattiau, F. Vieira, M. McNally, T. Charbonneau, M. Nollstadt, et al., Personalizing ASR for dysarthric and accented speech with limited data, 2019, *arXiv preprint arXiv:1907.13511*.
  19. R.N. Bloor, K. Barrett, C. Geldard, The clinical application of microcomputers in the treatment of patients with severe speech dysfunction, in: IEE Colloquium on High-Tech Help for the Handicapped, 1990, pp. 9/1–9/2.
  20. U. Sandler, Y. Sonnenblick, A system for recognition and translation of the speech of handicapped individuals, in: *MELECON'98. 9th Mediterranean Electrotechnical Conference. Proceedings* (Cat. No. 98CH36056), Vol. 1, IEEE, 1998, pp. 16–19.
  21. R. Palmer, P. Enderby, M. Hawley, A voice input voice output communication aid: What do users and therapists require? *J. Assist. Technol.* 4 (2) (2010) 4–14.
  22. M.S. Hawley, S.P. Cunningham, P.D. Green, P. Enderby, R. Palmer, S. Sehgal, P. O'Neill, A voice-input voice-output communication aid for people with severe speech impairment, *IEEE Trans. Neural Syst. Rehabil. Eng.* 21 (1) (2012) 23–31.
  23. Farzana, W., Sarker, F., Hossain, Q. D., Chau, T., & Mamun, K. A. (2020, July). An evaluation of augmentative and alternative communication research for ASD children in developing countries: Benefits and barriers. In *International Conference on Human-Computer Interaction* (pp. 51-62). Springer, Cham.
  24. Costa, M., Costa, A., Julián, V., & Novais, P. (2017, June). A task recommendation system for children and youth with autism spectrum disorder. In *International Symposium on Ambient Intelligence* (pp. 87-94). Springer, Cham.
  25. Liu R, Salisbury JP, Vahabzadeh A and Sahin NT (2017) Feasibility of an Autism-Focused Augmented Reality Smartglasses System for Social Communication and Behavioral Coaching. *Front. Pediatr.* 5:145. doi: 10.3389/fped.2017.00145
  26. Neamtu, R., Camara, A., Pereira, C., & Ferreira, R. (2019, September). Using artificial intelligence for augmentative alternative communication for children with disabilities. In *IFIP Conference on Human-Computer Interaction* (pp. 234-243). Springer, Cham.
  27. 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.
  28. Meltzner, G. S., Heaton, J. T., Deng, Y., De Luca, G., Roy, S. H., & Kline, J. C. (2017). Silent speech recognition as an alternative communication device for persons with laryngectomy. *IEEE/ACM transactions on audio, speech, and language processing*, 25(12), 2386-2398.

29. An, S., Feng, X., Dai, Y., Bo, H., Wang, X., Li, M., ... & Wei, L. (2017). Development and evaluation of a speech-generating AAC mobile app for minimally verbal children with autism spectrum disorder in Mainland China. *Molecular autism*, 8(1), 1-12.
30. Wilson, C., Brereton, M., Ploderer, B., & Sitbon, L. (2018, June). MyWord: Enhancing engagement, interaction and self-expression with minimally-verbal children on the autism spectrum through a personal audio-visual dictionary. *In Proceedings of the 17th ACM Conference on Interaction Design and Children* (pp. 106-118).
31. Kurniawan, I. (2018). The improvement of autism spectrum disorders on children communication ability with PECS method Multimedia Augmented Reality-Based. In *Journal of Physics: Conference Series* (Vol. 947, No. 1, p. 012009). IOP Publishing.
32. Lund, S. K., Quach, W., Weissling, K., McKelvey, M., & Dietz, A. (2017). Assessment with children who need augmentative and alternative communication (AAC): *Clinical decisions of AAC specialists. Language, speech, and hearing services in schools*, 48(1), 56-68.
33. Creer, S., Enderby, P., Judge, S., & John, A. (2016). Prevalence of people who could benefit from augmentative and alternative communication (AAC) in the UK: *determining the need. International Journal of Language & Communication Disorders*, 51(6), 639-653.
34. Dressler, R., Bland, L., & Baumgartner, M. (2016). The benefits of alternative and augmentative communication: A quality of life issue. *Internet Journal of Allied Health Sciences and Practice*, 14(4), 6.
35. Soto, G., & Yu, B. (2014). Considerations for the provision of services to bilingual children who use augmentative and alternative communication. *Augmentative and alternative communication*, 30(1), 83-92.
36. Handberg, Charlotte; Voss, Anna Katarina (2017). Implementing Augmentative and Alternative Communication in Critical Care Settings: Perspectives of Healthcare Professionals. *Journal of Clinical Nursing*, ().
37. Sigafos, J., O'Reilly, M. F., Lancioni, G. E., & Sutherland, D. (2014). Augmentative and alternative communication for individuals with autism spectrum disorder and intellectual disability. *Current Developmental Disorders Reports*, 1(2), 51-57.
38. Sennott, S. C., Akagi, L., Lee, M., & Rhodes, A. (2019). AAC and artificial intelligence (AI). *Topics in language disorders*, 39(4), 389.
39. Beukelman, D., & Mirenda, P. (2013). *Augmentative and alternative communication: Supporting children and adults with complex communication needs*, Paul Brookes. Baltimore, MD.
40. Barman, T., & Deb, N. (2017, September). Development of "Kotha" for people with speech impairments. *In 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)* (pp. 2652-2655). IEEE.
41. Garcia, R. G., Ibarra, J. B. G., Paglinawan, C. C., Paglinawan, A. C., Valiente, L., Sejera, M. M., ... & Villegas, M. C. (2017, December). Wearable augmentative and alternative communication device for paralysis victims using Brute Force Algorithm for pattern recognition. *In 2017 IEEE 9th International*

- Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)* (pp. 1-6). IEEE.
42. Rahman, M.; Dipta, D.; Hasan, M. Dynamic Time Warping Assisted SVM Classifier for Bangla Speech Recognition. *In Proceedings of the 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*, Rajshahi, Bangladesh, 8–9 February 2018; pp. 1–6.
43. Kong, X.; Choi, J.Y.; Shattuck-Hufnagel, S. Testing automatic speech recognition systems in comparison with human perception results using distinctive feature measures. *In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA, USA, 5–9 March 2017; pp. 5810–5814.
44. Mohamad, I.; Zainal, I.; Guntur, R. Implementation of Dynamic Time Warping algorithm on an Android-based application to write and pronounce Hijaiyah letters. *In Proceedings of the 2016 4th International Conference on Cyber and IT Service Management, Bandung, Indonesia, 26–27 April 2016*.
45. Ruben, G.; Muñoz, J.; Salazar, J.; Duque, N. Voice Recognition System to Support Learning Platforms Oriented to People with Visual Disabilities. *In Proceedings of the International Conference on Universal Access in Human-Computer Interaction, Toronto, ON, Canada, 17–22 July 2016; Springer: Cham, Switzerland, 2016; pp. 65–72*.
46. Mittal, Y.; Sharma, S.; Toshniwal, P.; Singhal, D.; Gupta, R.; Mittal, V. A voice-controlled multi-functional smart home automation system. *In Proceedings of the Annual IEEE India Conference (INDICON)*, New Delhi, India, 17–20 December 2015.
47. Essa, E.; Tolba, A.; Elmougy, S. Combined classifier based Arabic speech recognition. *In Proceedings of the INFOS 2008, Cairo, Egypt, 27–28 March 2008; pp. 27–29*.
48. Koskela, T.; Mattila, K. Evolution towards smart home environments: An empirical evaluation of three user interfaces. *Pers. Ubiquitous Comput.* 2004, 8, 234–240. [CrossRef]
49. Ismail, A., Abdlerazek, S., & El-Henawy, I. M. (2020). Development of smart healthcare system based on speech recognition using support vector machine and dynamic time warping. *Sustainability*, 12(6), 2403.
50. Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A. R., Jaitly, N., ... & Kingsbury, B. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal processing magazine*, 29(6), 82-97.
51. Öztürk, N., & Ünözkan, U. (2010, October). Microprocessor based voice recognition system realization. *In 2010 4th International Conference on Application of Information and Communication Technologies* (pp. 1-3). IEEE.
52. Sharma, M. K. (2014). Speech recognition: A review. *International Journal of Advanced Networking and Applications* (IJANA), 62-71.
53. Caron, J., Light, J., Davidoff, B. E., & Drager, K. D. (2017). Comparison of the effects of mobile technology AAC apps on programming visual scene displays. *Augmentative and Alternative*

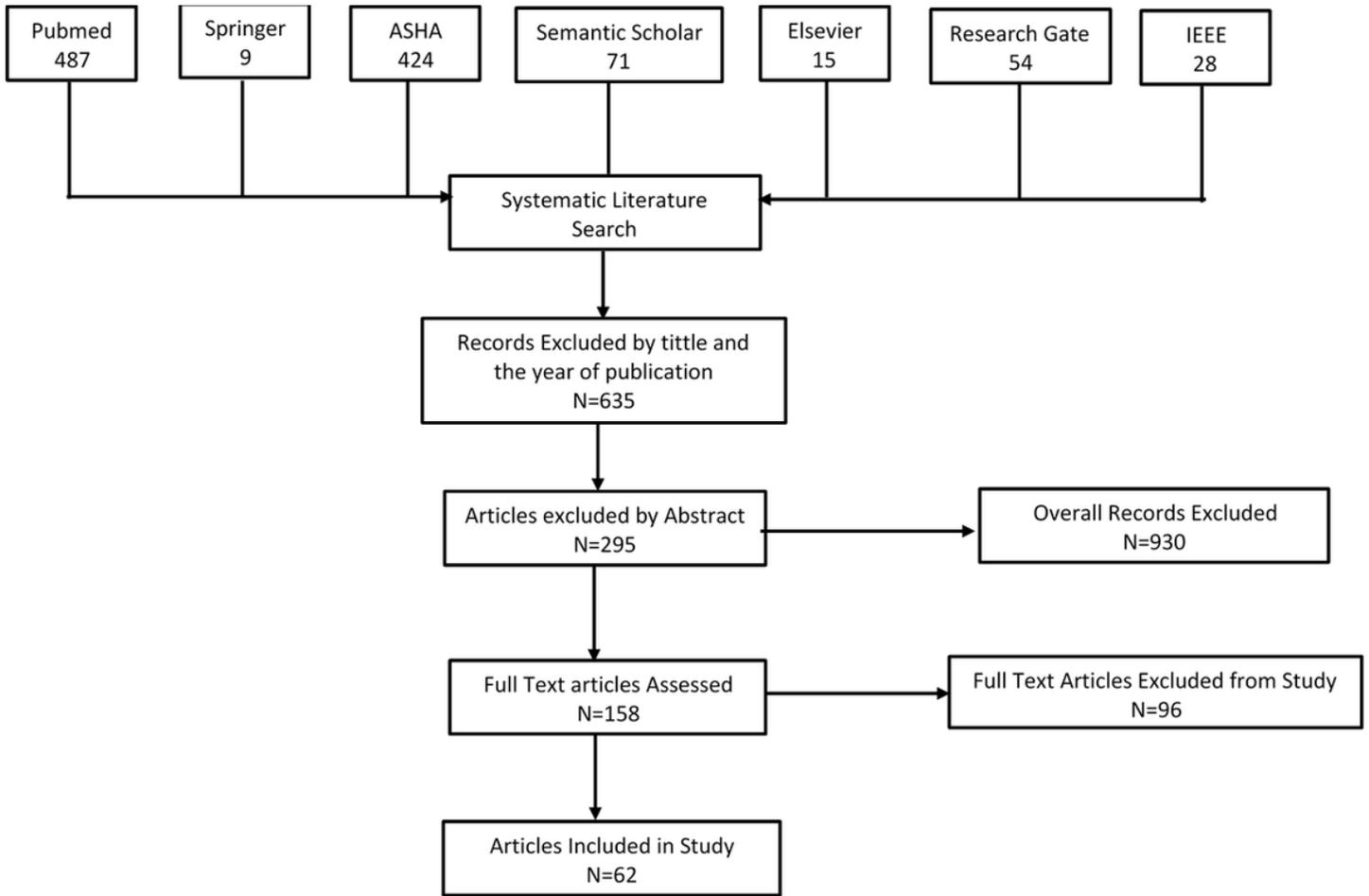
*Communication*, 33(4), 239-248.

54. Holyfield, C., Light, J., Mcnaughton, D., Caron, J., Drager, K., & Pope, L. (2020). Effect of AAC technology with dynamic text on the single-word recognition of adults with intellectual and developmental disabilities. *International Journal of Speech-Language Pathology*, 22(2), 129-140.
55. Vlachou, J. A., & Drigas, A. S. (2017). Mobile Technology for Students & Adults with Autistic Spectrum Disorders (ASD). *International Journal of Interactive Mobile Technologies*, 11(1).
56. Nakkawita, S. G., Duncan, E. S., & Hartzheim, D. U. (2021). AAC apps for aphasia: a pilot study on the role of intuition and learning. *Disability and Rehabilitation: Assistive Technology*, 1-11.
57. Chien, M. E., Jheng, C. M., Lin, N. M., Tang, H. H., Taelle, P., Tseng, W. S., & Chen, M. Y. (2015). iCAN: A tablet-based pedagogical system for improving communication skills of children with autism. *International Journal of Human-Computer Studies*, 73, 79-90.
58. Calteaux, K., De Wet, F., Moors, C., Van Niekerk, D., Mcalister, B., Davel, M., & Van Heerden, C. (2013). Lwazi II final report: Increasing the impact of speech technologies in South Africa Pretoria, South Africa.
59. Louw, J. A. (2008). Speech: A multilingual text-to-speech system. *Proceedings of the 19th annual symposium of the pattern recognition association of South Africa (PRASA)*, 165–168
60. Titmus, N., Schlünz, G. I., Louw, A., Moodley, A., Reid, T., & Calteaux, K. (2016). Lwazi III project final report: Operational deployment of indigenous text-to-speech systems. *Meraka Institute, Council for Scientific and Industrial Research, Pretoria, South Africa, Tech. Rep.*
61. Y. Zhou, X. Tian and H. Li, "Language Agnostic Speaker Embedding for Cross-Lingual Personalized Speech Generation," in *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 3427-3439, 2021.
62. Uchoa, J. P., Falcão, T. P., Nascimento, A. C., Miranda, P. B., & Mello, R. F. (2021, July). Fostering Autonomy through Augmentative and Alternative Communication. *In 2021 International Conference on Advanced Learning Technologies (ICALT)* (pp. 320-324). IEEE.
63. Gay, P., López, B., & Meléndez, J. (2011, December). Sequential learning for case-based pattern recognition in complex event domains. *In Proceedings of the 16th UK Workshop on Case-Based Reasoning* (pp. 46-55).

## Table 3

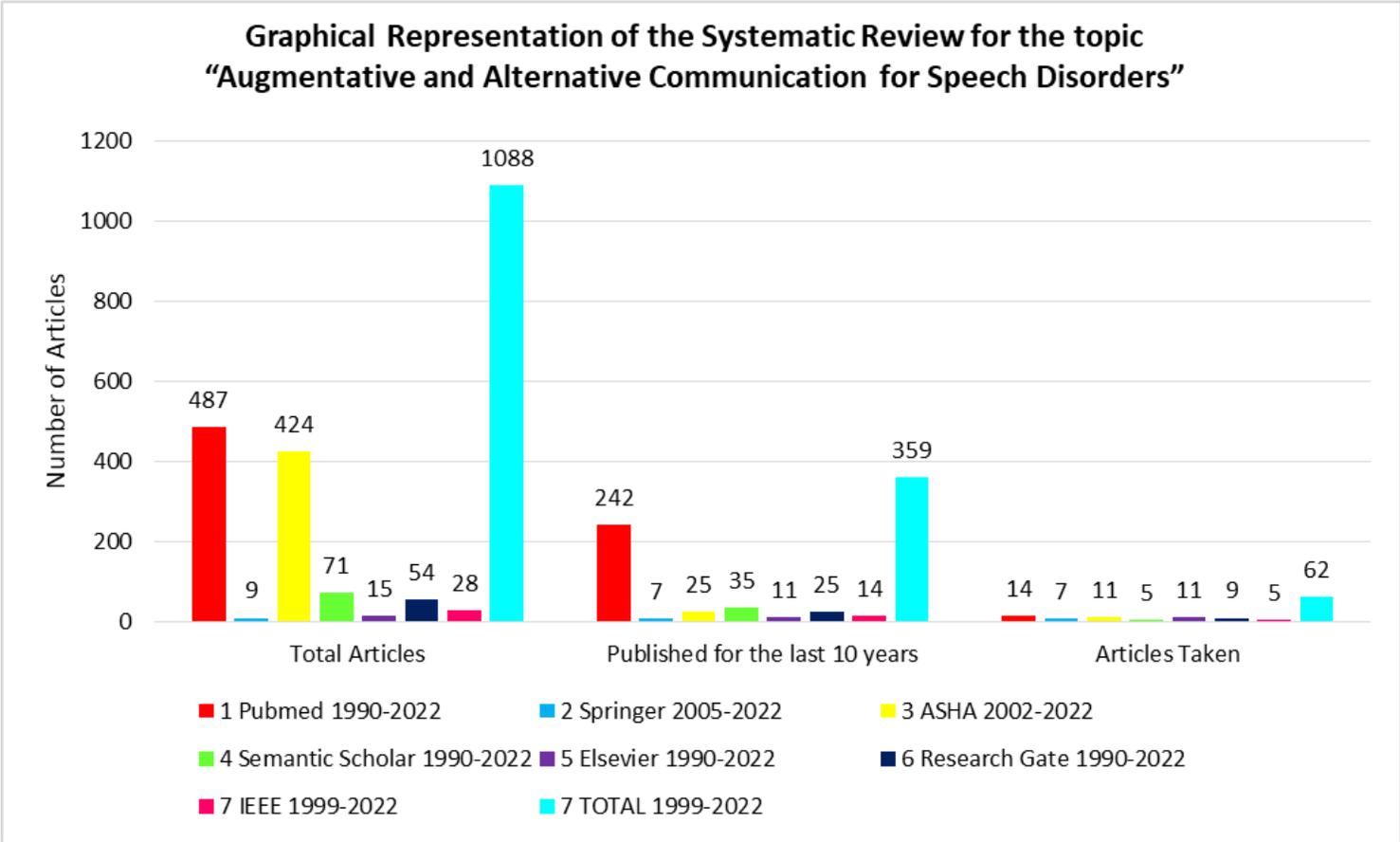
Table 3 is available in Supplementary Files section.

## Figures



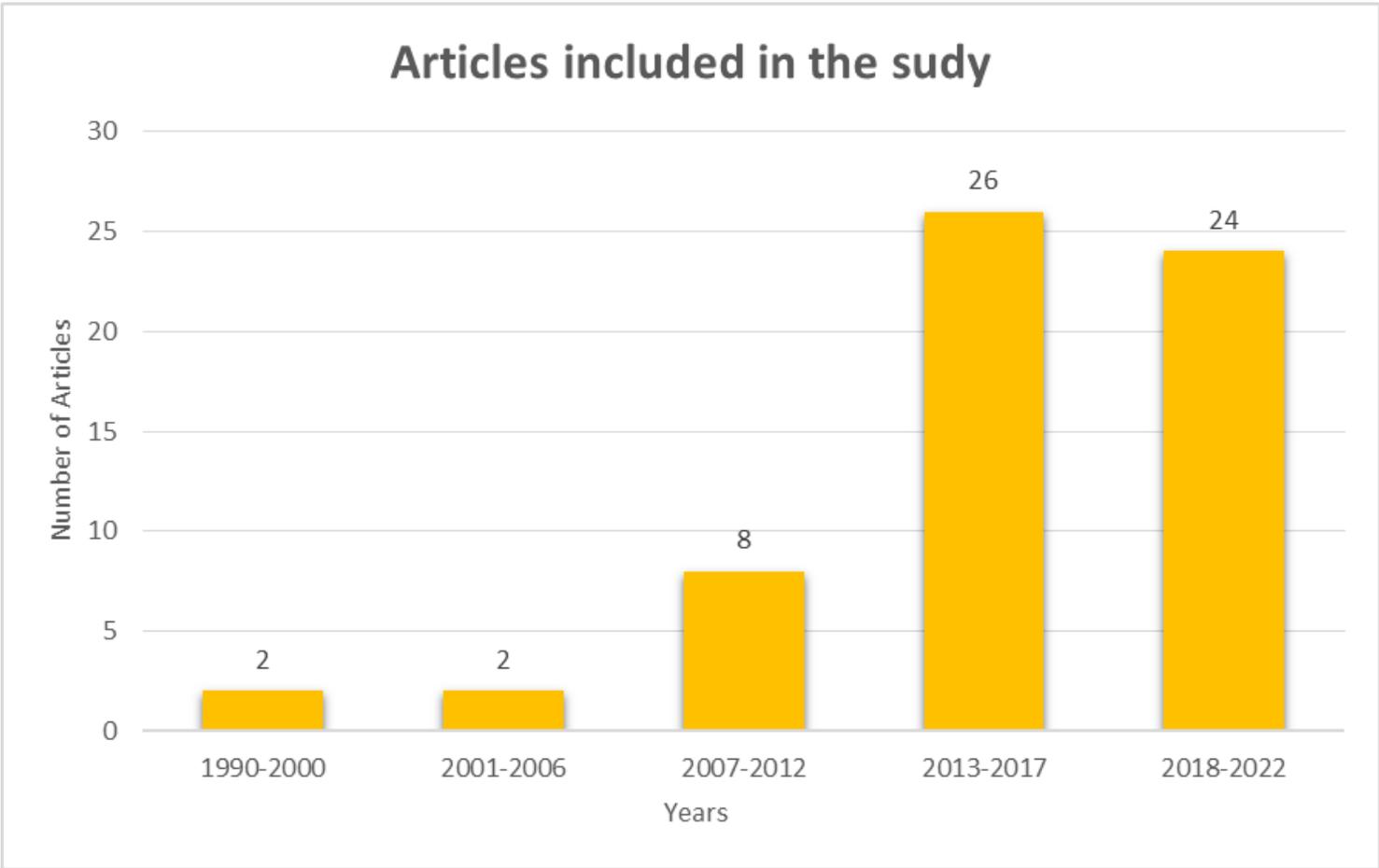
**Figure 1**

The flowchart of systematic search



**Figure 2**

Graphical Representation of the Systematic Review for the topic “Augmentative and Alternative Communication for Speech Disorders”

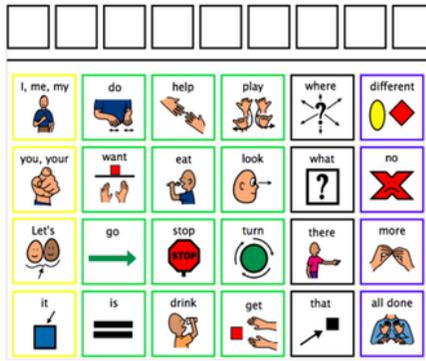


**Figure 3**

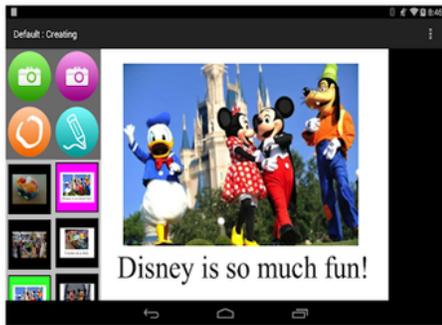
Graphical representation of the number of articles selected for the study, distributed in 5 year ranges.



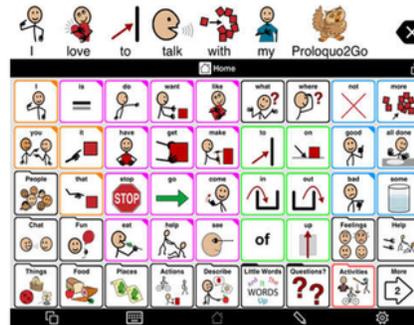
GoTalk Now



AutisMate



EasyVSD



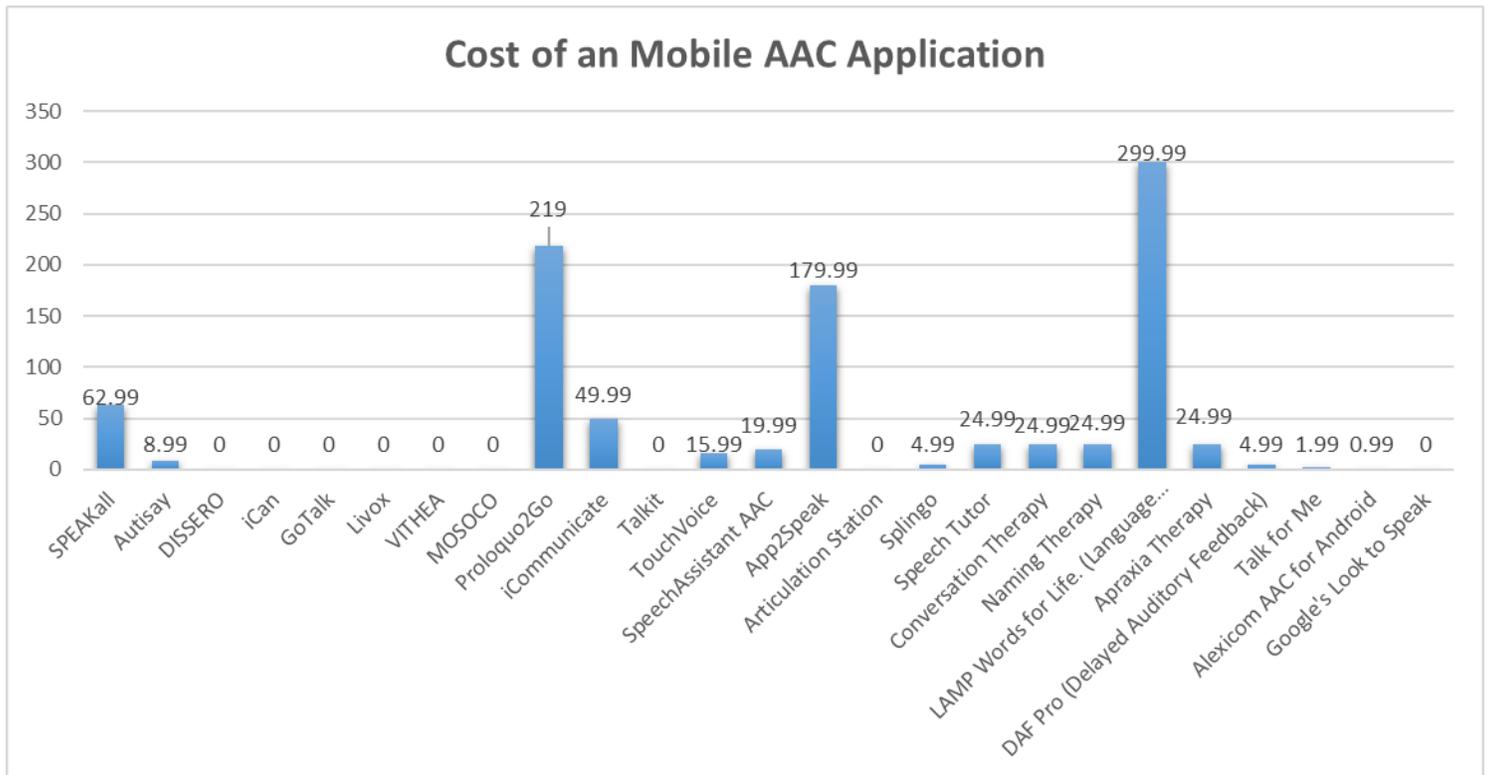
Proloquo2Go



Speak for Yourself

Figure 4

Existing mobile applications and their input methods.



**Figure 5**

The comparison of the cost of an AAC Mobile Application

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table3.docx](#)