

Futuristic Voice System For Visually Impaired Persons And Today's Scenario on Voice Email System For Communication

Samrat Dutta
22BAI70872

Dept. of Computer
Science & Engineering
Chandigarh University

Freestylewinds777@gmail.com
ORCID

ID:0009-0007-2691-
3691

Bhavya Jain
22BDA70046

Dept. of Computer
Science & Engineering
Chandigarh University

Bhavyajain2109@gmail.com
ORCID

ID:0000-
0003-3401-477X

Abstract— Accessing and managing email presents major difficulties for blind people. Despite recent advancements made by screen readers and other assistive technologies, reading and replying to emails can still be a laborious and time-consuming process. Voice email systems, which enable blind users to receive and communicate with their email using natural language voice commands, present a promising solution to this issue. In this review, we look at the most recent developments in voice email tools for visually impaired people. Artificial neural networks are used in deep learning, a subset of machine learning, to process and evaluate data. It can be applied in a number of ways to improve the blind's ability to use voice instructions. People who are blind or physically impaired can benefit greatly from improved accessibility thanks to AI and machine learning. Using natural language processing (NLP) methods, it is possible to decipher and comprehend spoken language. Devices can now understand voice commands and react correctly as a result. For people with restricted vision, vision loss has a significant negative effect on spatial memory and object identification.

Utilizing other sensory modalities, such as contact or sound, to make up for this is difficult.

The internet has recently evolved into one of the most essential tools for daily life. The internet is extensively used by all people to obtain knowledge, information, and for communication. However, blind individuals encounter challenges when attempting to utilize these internet-based resources as well as when utilizing any online service. The development of computer-based accessible tools has greatly expanded the range of opportunities available to people who are blind or visually challenged. Blind people have benefited greatly from the accessibility of voice-based search engines and audio-based virtual environments like screen readers. We outline the voice-based mail system architecture that a blind individual can use to quickly and easily receive emails.

Email is one of the most important means of business contact, but people who are blind or visually challenged cannot use features like email attachments. The interchange of attachments is one of the most crucial elements of business communication and transactions. The suggested research study created a framework that uses a text-to-speech and speech-to-text model to make it easy and effective for people who are blind or visually disabled to use email and send attachments. The suggested system can operate completely without the aid of a computer, mouse, or other external devices. The suggested model's security was made easier by using face detection. Ordinary users and people with limited literacy skills can use the suggested method.

With the Saarbruecken Voice Database, a variety of speech characteristics and models, including Deep Neural Networks

(DNN), are used for categorization jobs between pathological and sound speech. (SVD). However, the best recorded accuracy numbers for audio are 80.71% for phrases and 82.8% for the vowel /aiu/samples in SVD instead of a pick of a few pathologies when the assessment contains the large number of pathologies in the database.

Keywords—Artificial Intelligence, Machine learning, Natural Language Processing, Deep Neural Networks

I. INTRODUCTION

With advancements in digital healthcare The following topics are covered in this scholarly article: As they offer a simple and effective method of managing emails, voice email systems have grown to be an indispensable aid for people

with visual impairments. Due to the rising accessibility of speech recognition and text-to-speech conversion technology, voice email systems have grown in popularity in recent years. By using voice commands instead of the conventional email interfaces that require visual input, this technology enables people with vision impairments to access their emails. The creation and use of voice email systems for the blind has been the subject of a considerable amount of research. Studies have uncovered a number of crucial elements that affect how usable voice email systems are, such as the user interface's layout, the efficacy of speech detection, and the precision of text-to-speech conversion. Accessibility: The effectiveness of voice email tools for the blind is significantly impacted by accessibility. People are using voice email systems more frequently as they look for methods to communicate with others without having to type out lengthy messages. Without a keyboard or mouse, these systems enable users to transmit and receive email messages using only their voice. Voice email systems come in a variety of varieties, each with its own special features and powers.

One important component of artificial intelligence is voice translation (VC). It is the study of vocal imitation, or how to change one's speech to sound like another without altering the linguistic substance. Voice synthesis is a broad technological area that transforms text to speech or modifies the characteristics of speech, such as voice identification, expression, and accents.

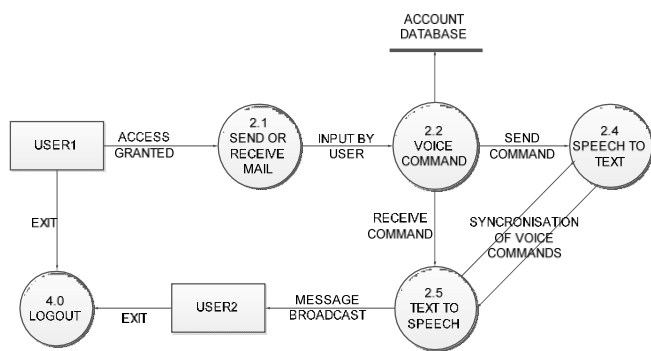


Figure 1. AI in acquiring and analyzing data of a patient in personalizing the treatment.

Figure-Shows the basic login page to send voice emails.

A standalone program that can be downloaded and installed on a computer or mobile device is one form of voice email system. These programs usually convert the user's voice into text using speech recognition technology, which is then used to create and send emails. Natural language processing features are also included in some of these programs, enabling users to dictate notes in a more conversational manner. A web-based application that is accessed through a browser is another kind of voice email system. Similar to standalone apps, these services use speech recognition technology to convert the user's voice to text. They are more adaptable and practical for users who need to access their email from various places, though, because they are web-based and accessible from any device with an internet connection. Particularly for users who have trouble typing or using a computer, they are quicker and simpler to use. Additionally, they make it possible for users to communicate with others while on the go without having to pause and type out a message. Blind people's voice systems are created to assist physically impaired people in navigating their environment through their sense of hearing. These systems come in a variety of formats, from dedicated devices made just for the blind to smartphone applications.

Since the introduction of computer-based speech synthesis in the 1950s, efforts to manipulate speech characteristics effectively have been ongoing. The control of the voice manipulation factors was significantly aided by the 1970s' rapid creation of digital signal processing. Personalized speech synthesis communication aids for the speech-impaired speaker de-identification voice mimicry and disguise, and voice dubbing for movies are just a few of the real-world applications that have benefited from technological advancements from statistical modeling to deep learning. The original motivation for voice conversion may have simply been novelty and curiosity.

Smartphone apps that use speech recognition technology to recognize items and read aloud text are a common voice system for the blind. These apps can be used to read books and other written materials, as well as to recognize anything from common objects to street signs.

Another type of voice system for the blind is a navigation system that uses voice prompts to guide users through unfamiliar environments. These systems typically use GPS technology to provide turn-by-turn directions

A. Vocally Powered Systems

1) Interactive Voice Response: Blind individuals can benefit from interactive vocal response (IVR) technology because it enables voice-only computer interaction. Although IVR systems are frequently used in customer support applications, they can be used for a variety of other tasks as well, like making payments or gaining access to information.

IVR systems can offer blind people an easy-to-use method of interacting with technology without the need for visible interfaces. Many IVR systems are created with accessibility in mind and might have features like text-to-speech and voice recognition to make it simpler for blind people to use the system. A blind individual can use an IVR system by simply dialing a phone number and speaking out loud to communicate with the system. The user may be prompted by the system to provide input by speaking precise commands or by expressing their requirements naturally in language. The system can then reply with pertinent data or prompts to assist the user in finishing their job.

IVR technology can be a useful tool for the blind, giving them a handy and accessible way to communicate with computers through voice commands.

A form of automated phone system known as interactive voice response (IVR) technology enables callers to engage with computerized menus or prompts using voice commands or touch-tone keypad responses.

IVR systems give a way to access information or services without requiring visual input or output, which makes them particularly helpful for blind people. Blind users can browse the interface and make selections using only their voice, eliminating the need for a visual display or the use of a mouse or keyboard.

Many businesses and groups have included features like speech recognition and text-to-speech capabilities in IVR systems to make them more available to the blind. With the aid of these features, the system is able to hear, comprehend, and carry out spoken instructions as well as repeat back information or prompts in a synthesized voice.

IVR systems can also be developed to include particular accessibility features for blind users, such as voice prompts that outline the options or the use of standardized menu layouts to make it easier for users to recall how to use the system. If users have trouble using the system, some IVR systems even give them the choice to speak with a live customer support agent.

IVR technology can be a useful instrument for blind people to access information and services, and its accessibility features can help to remove obstacles and enhance their overall user experience.



Figure 2 .shows the use of ivr technology
Source- [What is Interactive Voice Response \(IVR System\)? \(ivrtechgroup.com\)](http://www.ivrtechgroup.com)

2)Integrating Ai with Voice System for blind:AI can offer major advantages in terms of usability and accessibility in voice email systems for the blind. The following are some methods that AI could improve the functionality of voice email systems for blind people.

Voice messages can be converted into text using AI-powered speech-to-text technology, making it simpler for blind users to access and view their messages. Natural Language Processing (NLP).

It enables the voice email system to comprehend the message's context and offer pertinent ideas or responses. The system may prompt the user to dictate their answer if they request to reply to a message, for instance. AI can gradually acquire a user's preferences, such as the people they communicate with most frequently, and then recommend relevant messages or contacts. Vocal recognition technology makes it simpler for blind users to use vocal commands to navigate the voice email system rather than relying on conventional keyboard input.

By adding text-to-speech features for incoming messages, AI can help the voice email system become more accessible. This will allow blind users to hear their messages even if they don't have a screen reader. Sentiment analysis: Computer programs can interpret a speaker's tone and facial expressions to ascertain their mood and modify their answer accordingly.Contextual understanding: AI systems are able to respond to voice commands in a more personalized and pertinent way by using contextual data such as a user's location, the time of day, and recent activities.Algorithms for machine learning are used to identify spoken language and convert it into text. These algorithms can distinguish between various accents, dialects, and languages because they were taught on a lot of speech data.

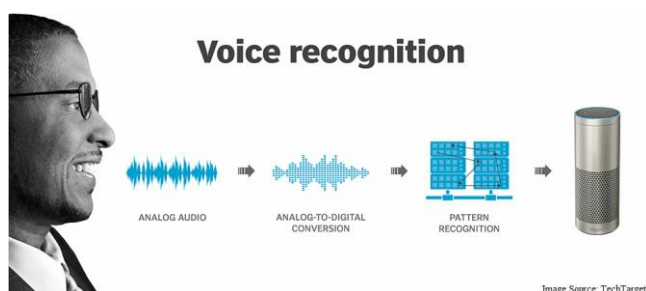


Figure 4 integrating the ai and machine learning

3)Today's scenario on voice system: There are an estimated 253 million blind or vision-impaired people in the world, and many of them depend on assistive technologies like screen readers, voice recognition software, and voice assistants to help them carry out tasks, including those that are work-related. As these technologies advance and become more accessible, they are being adopted more frequently, and it is possible that more blind people will use voice systems in the future.

Certainly! Similar to how email allows users to send text messages, voice email systems, also referred to as voicemail, are a type of communication technology that allow users to record and send audio messages to other users.

Use of voice email and voice commands is growing at a very high speed. Voice email systems, which enable users to leave messages for others when they are unable to accept a call or when they prefer to speak by voice rather than text, are frequently used for personal and business communications.

Service to Customers Customers can send messages with questions or concerns using voice email systems, which are used by many businesses as part of their customer service operations.

These messages can then be returned by a customer service representative at a later time. Voice email systems are essential for keeping in touch between team members who may be in different time zones or unavailable to accept calls during regular business hours as more people work remotely.

For people with visual impairments or other disabilities that make it difficult to view or type text messages, voice email services are especially helpful.

In comparison to traditional phone calls, voice email systems can be more effective because users can send messages without having to hold a conversation for a long time and can listen to the messages whenever it is convenient for them.



Figure 3 using voice commands for daily works.
Source- <https://www.improovo.com/assistentivocali/>

II.Deploying a voice System

It takes careful planning and consideration of the special requirements and difficulties faced by blind users when implementing a voice system. Organizations can make sure that blind users can use the system to communicate and complete duties by designing a system that is accessible, compatible, and user-friendly.

However, many nations have put in place a variety of assistive technologies to aid people who are blind or visually challenged, and voice systems are one popular kind of assistive technology. Many developed nations have laws and regulations in place to guarantee that businesses and organizations give people with disabilities, including those who are blind or visually impaired, equitable access to their goods and services. Additionally, there are a large number of non-profit groups and projects that promote

the use of assistive technologies by people with disabilities in various nations around the world. These groups strive to provide resources, training, and assistance to make it easier for people with disabilities to access and use assistive technologies, such as voice systems.

It is evident that the use of voice systems and other assistive technologies for the blind is expanding globally as more nations and organizations realize how important it is to grant everyone, regardless of ability, equal access to their goods and services. Due to their reliance on auditory feedback for communication, navigation, and information gathering, blind people frequently use voice systems. Voice systems give blind people access to an interface that they can use and navigate by speaking instructions, enabling them to carry out tasks that would otherwise be challenging or impossible.

A. Application of ai and deep learning with voice commands

For the blind, voice aides like Siri, Alexa, and Google Assistant can be of great assistance. These systems can comprehend commands given in natural language and respond correctly with the aid of artificial intelligence and machine learning. Voice assistants allow blind people to make calls, send messages, send reminders, and even manage their smart home devices.

AI and machine learning can also be used to translate written text into spoken phrases and the other way around. This is known as text-to-speech and speech-to-text. For blind people who might find it difficult to read written text, this can be particularly helpful. Books, papers, and other written materials can be read aloud using text-to-speech technology, and spoken words can be converted into written text using speech-to-text technology.

Navigation systems: Artificial intelligence (AI) and machine learning can be used to develop navigation systems that can lead blind people through unfamiliar settings. These systems could, for instance, use computer vision to identify impediments and offer audio guidance on how to avoid them.

Object recognition: Artificial intelligence (AI) and machine learning can also be used to assist blind people in recognizing objects and people in their surroundings. For instance, a smartphone camera could use computer vision to instantly recognize persons and objects while also giving the user audio feedback.

B. Voice assistants systems and senior citizens

Exploring how older adults view, engage with, and incorporate sIPAs into their daily routines is becoming more and more popular. Kim [23] interviewed older people (74+ years old) who had never used a virtual helper in task-based semi-structured interviews. The authors observe that while senior people in the research politely responded to their encounters with VAs, VAs have limited follow-up natural conversational skills.

The author suggests incorporating typical human conversational tenors into VA's conversational powers so that users can quickly comprehend and carry on natural-sounding discussions. The need to make VAs more personable is reiterated by Shalini et al. To achieve the goal of creating an intuitive health interface for older people, they contrasted a number of speech assistants, including Amazon Echo and Google Assistant. They claim that older people respond well to automatic health warnings from these systems, but these signals need to be more personable.

As older people become more experienced with VAs, their

opinions of them shift. According to the authors, participants' behavior shifted over time from simply savoring speech-based interactions to actively planning their responses to unwinding and admiring the freedom and ease of the interaction mode. Many participants found the device's basic talks to be enjoyable and helpful for fostering companionship. There are some suggestions for improving the conversational skills of VAs to support older people's use of the device to foster friendship.

C. Difficulties To Use

According to numerous studies, senior people find it difficult to master new technologies. The learnability of VA has also been a topic of discussion, along with voice non-recognition. The experts suggest alterations to conversational features and VA designs as a whole to improve the technology's learnability. Older people were able to grasp voice system. low-income older people were concerned about getting reliable information and maintaining the privacy of their personal information.

These demonstrate that older people are uninterested in using sIPAs due to a dearth of understanding, as well as false information and misinformation about the conquest by artificial intelligence. The authors advise seeking objective instruction of the subject matter to potential users in order to ensure that they have a thorough grasp of sIPAs and can make an educated decision about using a VA or a vocally powered system.

Alexa is something senior people would like to use less of. After conducting qualitative study with older people for a year, they discovered an intriguing trend that was related to and related to worries on usable instances. The results demonstrate that older adults view Alexa as a toy companion that can understand basic commands like playing music but fails to function properly in more complex use cases for assisted living and wellbeing.



Figure- shows the voice activated home device.

III. Benefits Of The Methods Discussed Above

A. Interactive interface using modern technologies

Benefits of the methods discussed above The majority of studies demonstrate how the speech-to-text and text-to-speech conversion processes make it easier and more interactive for people who are blind or visually impaired. People with disabilities feel like regular users thanks to this method. Voice-based technology is also helpful for those who are ignorant or disabled. Automatic voice recognition is one of the main benefits. We can observe a decrease in the mental effort required by blind people to recall and type characters on a keyboard. A user-friendly system is the voice-based messaging system. Limitations/Disadvantages of the Techniques Examined Above It is evident that mouse clicks are used for many duties in almost all of the papers.

Voice systems can help blind people complete tasks on their own, without requiring aid from sighted people.

Voice systems frequently complete duties more quickly than conventional techniques, like typing or reading, enabling blind people to work more productively.

Convenience to use is the most important in the deployment of voice commands.

Voice systems are a convenient way for blind people to engage with technology because they can be accessed from a variety of devices, including smartphones, computers, and smart speakers.

Using voice systems hands-free while conducting activities like driving or cooking can increase safety.

Voice systems can be tailored to the requirements of the user, including linguistic preferences and voice detection settings, which can enhance the user experience.

Voice systems are a useful aid for blind people because they can handle a variety of duties, such as reading texts, using the internet, and operating smart house devices.

IV. VOICE CONVERSION CHALLENGES In

Limitations/Disadvantages of the Technique Examined

It is evident that mouse movements are used for many duties in almost all of the documents. It becomes more challenging for those who are blind. Additionally, the Indian peninsula does not profit from this because there are numerous languages spoken there that voice recognition software cannot understand. The favored vernacular is primarily English.

Limited precision: Voice systems depend on imperfect voice detection software. Certain words or sentences may be challenging for the system to identify, especially if the person has an accent or speaking impediment.

Limited lexicon: The vocabulary that has been put into voice programs is its only limitation. If the user needs to use a term or sentence that is not listed in the system's database, this could be an issue.

Limited context: Voice algorithms may occasionally be unable to discern the meaning of a word or sentence. Confusion and communication mistakes may result from this.

Limited input: Voice systems frequently give the user only a small amount of feedback, making it challenging for the user to determine whether the system has properly grasped their request.

Limited movement: Since voice systems generally require the user to be near the device, this can restrict the user's freedom and mobility.

Limited privacy: Blind users of voice systems may have privacy concerns because they may not always be aware of who else is listening to their orders and replies.

Limited compatibility: Voice systems may not work with all hardware and software, which can reduce their utility for blind users who depend on a range of technological solutions.

A.Lack of enthusiasm, skepticism, and continued use

As previously mentioned, many vocal conversion methodologies are data-driven, so speech data are necessary for model training as well as conversion assessment. It is necessary to have a shared database that clearly defines training and evaluation data in order to compare such data-driven methods to one another. But a shared library like that didn't exist until 2016. Without shared datasets, academics are forced to recreate other people's systems using their own databases before attempting any novel concepts. It is not always the case that a

system that has been modified and re-implemented will perform as anticipated in the initial work.

The adoption of new technologies and their ongoing use are influenced by a variety of personal variables, including self-efficacy, personal inventiveness, technology anxiety, and technology familiarity, according to recent study. According to authors in, the major obstacles to using new technologies are skepticism and a sluggish acceptance rate. Anytime consumers are resistant to a technology, the adoption and utilization situation are delayed. According to the diffusion of innovation curve, roughly 50% of the consumers fall into this category of late users. This set of users typically has doubts about any new products.

Voice aides have developed quickly in recent years, becoming almost universally present now that they are built into all current smartphone operating systems. Voice helper services have developed along with technological developments, and they are now an essential component of a smart home setup. Thus, with the introduction of new categories of smart products, the world of digital technology is changing quickly. Therefore, it is crucial to design goods that users use consistently for a long time in addition to having a high adoption rate.

G.The components of voice system

All existing voice-based email systems provide their own user-developed email services and do not incorporate the use of Google's Gmail, so taking this into consideration, the intention is to develop the application by linking it with the Gmail Client, thereby giving users an additional advantage.

Here, this was developed using internet tools. Directly obtaining and converting speech to writing is what a speech to text device does. The various components of speech recognition systems include feature extraction, acoustic models generated from training data, dictionaries, language models, and voice recognition algorithms. 2. Convert Text to Speech It transforms text to voice output using speech synthesis methods. It helps the blind focus on printed material and is now widely used to transmit financial data, according to e- International Journal of Computer Applications.

Speech-to-Text Conversion Program A speech-to-text translator can identify your voice, assess the noises you make by filtering what you say, and then transcribe your speech into a text file that can be viewed by a computer. It is possible to store the identified text in a file. The tools used to create this are Net and C#.Net. speech-to-text system acquires and transforms speech straight into writing. The various components of speech recognition systems include feature extraction, acoustic models generated from training data, dictionaries, language models, and voice recognition algorithms. Convert Text to Speech It transforms text to voice output using speech synthesis methods.

The blind use it to focus on written material, and it is now widely used to transmit banking information, e-mail communications, and other information over the telephone for everyone. When providing instructions, text-to-speech is also used on gadgets like handheld GPS units to read out city names.

An email application that uses voice The user can use vocal commands to transmit emails, listen to what they have written, and receive emails. The SMTP and POP3 protocols are used by

Email for delivering and getting emails, respectively.

The SMTP (Simple Mail Transfer system) server rapidly transmits the email messages, making it a dependable system for sending emails. In order to receive communications, POP3 (Post Office Protocol) is used. Emails are stored on the POP3 service and made available for viewing upon request. The same is done in our program so that emails are downloaded at the user's request.

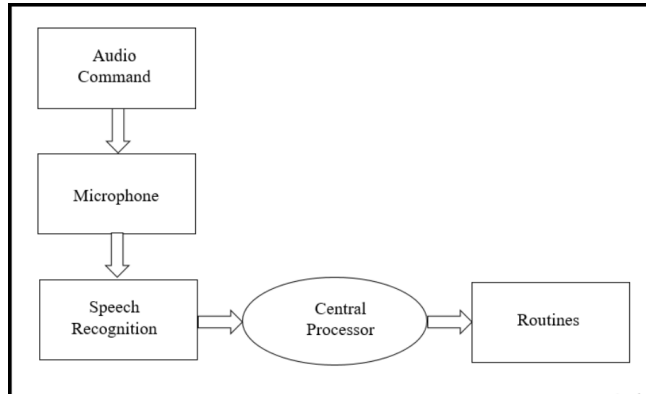


Figure –shows the basic approach to execute voice commands

C.Voice assistant for blind usability

The classification is founded on advancements made during the speech assistant era. The first group covers the years 2000 to 2006, which was also known as the year of the Y2K bug in telecoms and the year of social media and camera phones. With the debut of innovations like Honda's Advanced Step in Innovative Mobility (ASIMO**) humanoid robot, talking bots began to gain attention during this time. The second group includes years 2007 to 2014.

The second group includes years 2007 to 2014. Through the integration of speech aides into cellphones and laptops, technical developments during this time period increased user exposure to them. For instance, Microsoft unveiled Cortana in 2014, while Apple debuted SIRI in 2011. The final group spans the years 2015 to 2021. The widespread use of speech assistants peaked during this period, setting a record.

It is abundantly obvious that, over the past six years, research on voice systems has greatly increased. (2014–2021). The development of voice-activated smart speakers and phones is to blame for this. The COVID-19 outbreak, which has renewed interest in touchless interaction technologies like voice, is another factor in vocally powered systems growing popularity

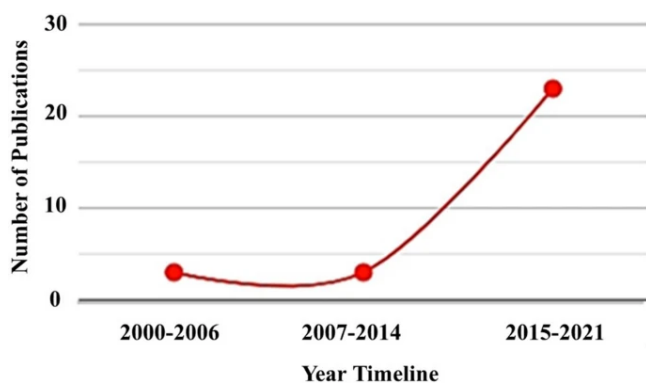


Figure –shows the advancement of voice systems over the time.

V. Voice systems with Various Embodiment Types

Smart speakers are the most prevalent embodiment of virtual assistants (VAs) used in our chosen articles. The present appeal of commercial smart devices like Alexa, HomePod, etc. is to blame for this. According to a 2019 survey, there are smart speakers in 35% of US homes, with that number expected to rise to 75% by 2025. The use of humanoids is also widespread because anthropomorphism is a key component of usability metrics for speech assistants.

Car interfaces are also verbal aides that serve as a conduit between the driver and the vehicle. The VA car interface enables users to access vehicle data and complete tasks without diverting their attention away from the road. The speech assistant software that is integrated into cellphones or laptops is the fourth form of software interface. The studies we have gathered have either used the commercialized form of the software interface, like Alexa and Siri, or others have developed new voice interfaces that are easily usable by users due to the adoption of smartphones and computer assistants using programming codes and skills. Both, however, take the shape of various program entities.

Smart Speakers – Smart speakers are stand-alone gadgets that can be used for a variety of tasks, including playing music, answering questions, and operating smart home gadgets. Apple HomePod, Google Home, and Amazon Echo are a few examples of smart devices.

cellphones - A lot of contemporary cellphones come equipped with speech aides like Siri, Google Assistant, and Alexa, which let users carry out a variety of tasks by speaking instructions.

Wearables - Wearable gadgets like fitness trackers and smartwatches can include voice assistants, enabling users to complete a variety of tasks without having to remove their phone from their pocket or bag.

Home automation systems - mouth assistants can be used to manage home automation systems, giving users the ability to regulate everything with just their mouth, from lighting to climate to security systems.

In-Car Systems - A lot of modern vehicles have voice-activated systems that let the driver operate the radio, make phone calls, and even travel without taking their hands off the wheel. Voice assistants can also be built into robots, enabling more intuitive exchanges between people and computers. These robots include the Echo Look from Amazon and the Pepper from SoftBank Robotics.

These are just a few instances of the many different platforms and kinds of equipment that can use speech aides or voice-controlled devices. Future developments should bring about even more creative manifestations.

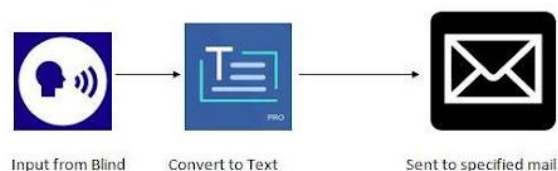


Figure Voice-based messaging system design for visually

VI.DISCUSSION AND CONCLUSIONS

This paper outlines a voice-based email system for people who are blind or visually impaired. It was created as a tool to make it easier and more effective for them to receive emails. It offers a voice-based mailing service so that people who are blind can receive and send mail on their own, without assistance. Basic knowledge of keyboard commands is necessary. The system has done away with all of these ideas and solved every issue the blind have. It makes use of a speech recognition programme to give blind users of devices an effective voice input technique. People who are disabled or illiterate can also benefit from it.

In the future, it is intended to make the system entirely voice-based and without a computer. This app might be enhanced in the future and used for a variety of services besides email, like messaging, taking notes, and voice-operated control of other programmes. Furthermore, with the aid of high-tech instruments, even the tiniest number of keyboard shortcuts may be eliminated. The capabilities of voice could be enhanced in the future to include image attachments and other features found in standard E-Mail, such as indentation and font choice.

Therefore, it is simple for those who are visually impaired to obtain the services. The method as it is now only functions on desktop computers. There is potential to include this feature as an application in mobile phones since their use is currently on the rise. To make the system safer, security measures that will be used during the login period can also be revised. This strategy might help overcome a number of challenges that blind people previously faced when attempting to receive communications.

There are no longer any screen readers, which might have made it easier for users to memorise duties. Increased feeling of community among those who are blind in this small environment is the main objective of creating the kind of system described in the research. Each action has a unique result, so the decision tree follows a specific route. This makes the system much more compatible. With the help of our technology, people with bodily disabilities will be able to interact with the outside world.

Regardless of age, anyone can use this messaging system with simplicity. It has speech to content and speech to content with discourse reader capabilities, making planned structure usable for people with exterior disabilities as well. Now, people who are blind or visually impaired can send and receive messages easily with just vocal commands and very little computer or mouse use. It has assisted in removing the difficulties that impaired people face and transforming them into more sociable individuals. It has done away with the idea of using terminal shortcuts in addition to screen readers, which would have reduced the brain load.

This paper gives a thorough introduction to speech conversion technology, outlining its principles and applications. We explore the potential and constraints of the underlying technologies and how they relate to statistical methods to deep learning. We also research speech conversion assessment methods. We also list a number of tools and difficulties with voice conversion that scholars and engineers can use to start their own voice conversion studies.

REFERENCES

- [1] S. Kim, "Exploring how older adults use a smart speaker-based voice assistant in their first interactions: Qualitative study", *JMIR mHealth uHealth*, vol. 9, no. 1, Jan. 2021.
- [2] S. Shalini, T. Levins, E. L. Robinson, K. Lane, G. Park and M. Skubic, "Development and comparison of customized voice-assistant systems for independent living older adults", *Proc. Int. Conf. Hum.-Comput. Interact.*, pp. 464-479, Jul. 2019.
- [3] S. Kim and A. Choudhury, "Exploring older adults' perception and use of smart speaker-based voice assistants: A longitudinal study", *Comput. Hum. Behav.*, vol. 124, Nov. 2021.
- [4] J. Kowalski, A. Jaskulska, K. Skorupska, K. Abramczuk, C. Biele, W. Kopeć, et al., "Older adults and voice interaction: A pilot study with Google home", *Proc. Extended Abstr. CHI Conf. Hum. Factors Comput. Syst.*, pp. 1-6, May 2019.
- [5] J. Kowalski, A. Jaskulska, K. Skorupska, K. Abramczuk, C. Biele, W. Kopeć, et al., "Older adults and voice interaction: A pilot study with Google home", *Proc. Extended Abstr. CHI Conf. Hum. Factors Comput. Syst.*, pp. 1-6, May 2019.
- [6] M. Hanley and S. Azenkot, "Understanding the use of voice assistants by older adults", *arXiv:2111.01210*, 2021.
- [7] M. Trajkova and A. Martin-Hammond, "'Alexa is a toy': Exploring older adults' reasons for using limiting and abandoning echo", *Proc. CHI Conf. Hum. Factors Comput. Syst.*, pp. 1-13, 2020.
- [8] N. K. Prakash, T. Monlam, R. Singh, R. Aravindhan, P. V. Reddy and R. Jain, "Voice based E-Mail with Attachment for Blind," 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2022, pp. 01-07, doi: 10.1109/ICOEI53556.2022.9776840.
- [9] S. Kim, "Exploring how older adults use a smart speaker-based voice assistant in their first interactions: Qualitative study", *JMIR mHealth uHealth*, vol. 9, no. 1, Jan. 2021.
- [10] P. Nallam, S. Bhandari, J. Sanders and A. Martin-Hammond, "A question of access: Exploring the perceived benefits and barriers of intelligent voice assistants for improving access to consumer health resources among low-income older adults", *Gerontol. Geriatric Med.*, vol. 6, Dec. 2020.
- [11] M. Hanley and S. Azenkot, "Understanding the use of voice assistants by older adults", *arXiv:2111.01210*, 2021.
- [12] L. M. Bowen, "Resisting age bias in digital literacy research", *College Composition Commun.*, vol. 62, no. 4, pp. 586-607, 2011.
- [13] Parkhi Bhardwaj, Gunjan Sethi "Voice Based E-mail System for Visually Impaired: A Review" published in International Research Journal of Engineering and Technology (IRJET) on December 12th, 2020.
- [14] Ingle, P., Kanade, H. and Lanke, A., 2016. Voice based e-mail System for Blinds. International Journal of Research Studies in Computer Science and Engineering (IJRSCSE), 3(01), pp.25-30. Jain, V., AK, K., Shenoy, R.N. and Ahmed, M., 2021. Voice Based Email for the Visually Impaired. Divesh Jethani, Simmi Saluja, Saurabh Gautam Published 26 March 2018 Computer Science International Journal of Advance Research, Ideas and Innovations in Technology
- [15] Biruntha, S., Priya, M.G., Kiruthika, R., Indupriya, N. and Ashwini, R., 2021. Voice Based Email for Blind People Using Speech Recognition through Artificial Intelligence. International Journal of All Research Education and Scientific Methods (IJARESM), 9(04).
- [16] Parkhi Bhardwaj, Gunjan Sethi Voice Based E-mail System for Visually Impaired: A Review published in International Research Journal of Engineering and Technology (IRJET) on December 12th, 2020.
- [17] Pathan, N., Bhoyar, N., Lakra, U. and Lilhare, D., 2019. V-Mail (Voice Based E-Mail Application). International Research Journal of Engineering and Technology (IRJET), 6(03).
- [18] Tharani K K, Shalini R, Jeyanthi I, Dr. Deepalakshmi R (2017), Voice Based Mail Attachment For Visually Challenged People, in International Journal of Scientific and Engineering Research, Vol.8, Issue.5, pp. 126-130.
- [19] D Kiran kumar, User Interface for Visually Impaired People, IOSR Journal of Electronics and Communication Engineering (IOSR-JECE) (Jan.-Feb. 2017), PP 65-71
- [20] Shoba, G., Anusha, G., Jeevitha, V. and Shanmathi, R., 2014. An interactive email for visually impaired. International Journal of Advanced Research in Computer and Communication Engineering (IJARCE), (5089-5092)
- [21] Bhowmick, A. and Hazarika, S.M., 2017. An insight into assistive technology for the visually impaired and blind people: state-of-the-art and future trends. Journal on Multimodal User Interfaces, 11(2), pp.149- 172.
- [22] Ummuhanysifa, U. and Nizar Banu, P.K., 2013. Voice based search engine and web page reader. International Journal of Computational Engineering Research (IJCER), pp.1-5.
- [23] Tiwari, P.A., Zodawan, P., Nimkar, H.P., Rotke, T., Wanjari, P.G. and Samarth, U., 2020, February. A Review on Voice based E-Mail System for Blind. In 2020 International Conference on Inventive Computation Technologies (ICICT) (pp. 435-438). IEEE.
- [24] Hari Priya, S.L., Karthigasree, S. and Revathi, K., 2015. VoiceBased E- Mail (V-Mail) for blind. International Journal.
- [25] Kulkarni, O., Alhat, A., Tejanekar, N. and Patil, M., 2019. Voice based e- mail system for blind people. Open access international journal of science and engineering, 4(01)
- [26] K. Gupta, R. Hajika, Y. S. Pai, A. Duenser, M. Lochner, and M. Billinghurst, "Measuring human trust in a virtual assistant using physiological sensing in virtual reality," in Proc. IEEE Conf. Virtual Reality 3D User Interfaces (VR), Mar. 2020, pp. 756-765.
- [27] M. Trajkova and A. Martin-Hammond, "'Alexa is a toy': Exploring older adults' reasons for using, limiting, and abandoning echo," in Proc. CHI Conf. Hum. Factors Comput. Syst., 2020, pp. 1-13.
- [28] J. Balsa et al., "Intelligent virtual assistant for promoting behaviour change in older people with T2D," in Proc. EPIA Conf. Artif. Intell. Cham, Switzerland: Springer, Sep. 2019, pp. 372-383.
- [29] L. M. Bowen, "Resisting age bias in digital literacy research," College Composition Commun., vol. 62, no. 4, pp. 586-607, 2011.
- [30] Y. Liao, J. Vitak, P. Kumar, M. Zimmer, and K. Kritikos, "Understanding the role of privacy and trust in intelligent personal assistant adoption," in Proc. Int. Conf. Inf. Cham, Switzerland: Springer, Mar. 2019, pp. 102-113.
- [31] J. Lau, B. Zimmerman, and F. Schaub, "Alexa, are you listening? Privacy perceptions, concerns and privacy-seeking behaviors with smart speakers," Proc. ACM Hum.-Comput. Interact., vol. 2, pp. 1-31, Nov. 2018.
- [32] H. Chung, M. Iorga, J. Voas, and S. Lee, "Alexa, can I trust you?" Computer, vol. 50, no. 9, pp. 100-104, 2017.
- [33] P.-S. Chiu, J.-W. Chang, M.-C. Lee, C.-H. Chen, and D.-S. Lee, "Enabling intelligent environment by the design of emotionally aware virtual assistant: A case of smart campus," IEEE Access, vol. 8, pp. 62032-62041, 2020.
- [34] K. Weitz, D. Schiller, R. Schlagowski, T. Huber, and E. André, "'Do you trust me?': Increasing user-trust by integrating virtual agents in explainable AI interaction design," in Proc. 19th ACM Int. Conf. Intell. Virtual Agents, Jul. 2019, pp. 7-9.

- [35] P. M. Podsakoff, S. B. MacKenzie, J. Y. Lee, and N. P. Podsakoff, "Common method biases in behavioral research: A critical review of the literature and recommended remedies," *J. Appl. Psychol.*, vol. 88, no. 5, pp. 879–903, 2003.
- [36] N. K. Malhotra, S. S. Kim, and A. Patil, "Common method variance in IS research: A comparison of alternative approaches and a reanalysis of past research," *Manage. Sci.*, vol. 52, no. 12, pp. 1865–1883, Dec. 2006. [68] N. Kock and G. Lynn, "Lateral collinearity and misleading results in variance-based SEM: An illustration and recommendations," *J. Assoc. Inf. Syst.*, vol. 13, no. 7, pp. 546–580, Jul. 2012.
- [37] J. F. Hair, C. M. Ringle, and M. Sarstedt, "PLS-SEM: Indeed a silver bullet," *J. Marketing Theory Pract.*, vol. 19, no. 2, pp. 139–152, 2011.
- [38] J.-M. Becker, K. Klein, and M. Wetzels, "Hierarchical latent variable models in PLS-SEM: Guidelines for using reflective-formative type models," *Long Range Planning*, vol. 45, nos. 5–6, pp. 359–394, 2012.
- [39] J. F. Hair, M. Sarstedt, and C. M. Ringle, "Rethinking some of the rethinking of partial least squares," *Eur. J. Marketing*, vol. 53, no. 4, pp. 566–584, Apr. 2019.
- [40] D. Lacobucci, "Everything you always wanted to know about SEM (structural equations modeling) but were afraid to ask," *J. Consum. Psychol.*, vol. 19, no. 4, pp. 673–680, 2009.
- [41] R. P. Bagozzi and Y. Yi, "On the evaluation of structural equation models," *J. Acad. Marketing Sci.*, vol. 16, no. 1, pp. 74–94, 1988.
- [42] J. Hulland, "Use of partial least squares (PLS) in strategic management research: A review of four recent studies," *Strategic Manage. J.*, vol. 20, no. 2, pp. 195–204, 1999.
- [43] C. Fornell and D. F. Larcker, "Evaluating structural equation models with unobservable variables and measurement error," *J. Marketing Res.*, vol. 18, no. 1, pp. 39–50, 1981.
- [44] J. F. Hair, G. T. M. Hult, C. M. Ringle, and M. Sarstedt, *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*, 2nd ed. Thousand Oaks, CA, USA: Sage, 2017.
- [45] H. Pirkkalainen, M. Salo, and M. Makkonen, "IT engagement as a blessing and a curse? Examining its antecedents and outcomes in organizations," *Int. J. Inf. Manage.*, vol. 53, Aug. 2020, Art. no. 102130, doi: 10.1016/j.ijinfomgt.2020.102130.
- [46] J. Marbach, C. R. Lages, and D. Nunan, "Who are you and what do you value? Investigating the role of personality traits and customer-perceived value in online customer engagement," *J. Marketing Manage.*, vol. 32, nos. 5–6, pp. 502–525, Mar. 2016.
- [47] M. Dehghani, K. J. Kim, and R. M. Dangelico, "Will smartwatches last? Factors contributing to intention to keep using smart wearable technology," *Telematics Informat.*, vol. 35, no. 2, pp. 480–490, May 2018.
- [48] M. Herz and P. A. Rauschnabel, "Understanding the diffusion of virtual reality glasses: The role of media, fashion and technology," *Technol. Forecasting Social Change*, vol. 138, pp. 228–242, Jan. 2019.
- [49] P. Kowalczyk, "Consumer acceptance of smart speakers: A mixed methods approach," *J. Res. Interact. Marketing*, vol. 12, no. 4, pp. 418–431, Oct. 2018.
- [50] D. Pal, C. Arpnikanondt, S. Funilkul, and V. Varadarajan, "User experience with smart voice assistants: The accent perspective," in *Proc. 10th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, Kanpur, India, Jul. 2019, pp. 1–6, doi: 10.1109/ICCCNT45670.2019.8944754.
- [51] D. Ribas, M. A. Pastor, A. Miguel, D. Martínez, A. Ortega and E. Lleida, "Automatic Voice Disorder Detection Using Self-Supervised Representations," in *IEEE Access*, vol. 11, pp. 14915–14927, 2023, doi: 10.1109/ACCESS.2023.3243986.
- [52] B. Sisman, J. Yamagishi, S. King and H. Li, "An Overview of Voice Conversion and Its Challenges: From Statistical Modeling to Deep Learning," in *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 132–157, 2021, doi: 10.1109/TASLP.2020.3038524.
- [53] A. Kain and M. W. Macon, "Spectral voice conversion for text-to-speech synthesis," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, (ICASSP), 1998, pp. 285–288.
- [54] [3] M. Zhang, X. Wang, F. Fang, H. Li, and J. Yamagishi, "Joint training framework for text-to-speech and voice conversion using multi-source tacotron and wavenet," 2019, arXiv:1903.12389.
- [55] C. Veaux, J. Yamagishi, and S. King, "Towards personalised synthesised voices for individuals with vocal disabilities: Voice banking and reconstruction," Aug. 2013.
- [56] B. Srivastava, N. Vauquier, M. Sahidullah, A. Bellet, M. Tommasi, and E. Vincent, "Evaluating voice conversion-based privacy protection against informed attackers," Nov. 2019.
- [57] Z. Wu and H. Li, "Voice conversion versus speaker verification: An overview," *APSIPA Trans. Signal Inf. Process.*, vol. 3, p. e17, 2014.
- [58] C. Y. Huang, Y. Y. Lin, H. Y. Lee, and L. S. Lee, "Defending your voice: Adversarial attack on voice conversion," vol. abs/2005.08781, 2020.
- [59] [67] T. Toda et al., "The voice conversion challenge," in *Proc. Interspeech*, 2016, pp. 1632–1636.
- [60] J. Lorenzo-Trueba et al., "The voice conversion challenge 2018: Promoting development of parallel and nonparallel methods," in *Proc. Odyssey, Speaker Lang. Recognit. Workshop*, 2018, pp. 195–202. [Online]. Available: <http://dx.doi.org/10.21437/Odyssey.2018-28>
- [61] X. Hu, A. Song, Z. Wei and H. Zeng, "StereoPilot: A Wearable Target Location System for Blind and Visually Impaired Using Spatial Audio Rendering," in *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 1621–1630, 2022, doi: 10.1109/TNSRE.2022.3182661.