



“Empowering visually impaired students: Integrating text-to-speech and automated keyboard technologies”

Yogeeta, Anshul, Chaitanya Dhawan, Dr. Meena Chaudhary, Dr. Gunjan Chandwani
Department of Computer Science and Technology, Manav Rachna University, Faridabad, India

Abstract

This document presents an assistive system designed to empower students with visual impairments by integrating Text-to-Speech (TTS) technology and automated keyboard solutions. The system enables users to interact with digital materials by converting on-screen text into natural-sounding speech via advanced TTS engines, while also offering voice or Braille input for tasks such as note-taking and answering exam questions. This approach enhances accessibility, reduces reliance on human transcribers, and promotes autonomy during assessments. A comparative analysis of TTS and automated keyboard systems is included, evaluating aspects such as accuracy, usability, speed, and cost. The proposed system aims to bridge the digital divide for visually impaired students and support inclusive education through intelligent, voice-activated interaction.

Keywords: TTS (text-to-speech). automated keyboard · stt (speech-to-text) · voice-based recognition system · AI-generated voice system · ai-based educational support

Introduction

In recent years, technological advancements across the globe have grown at an astonishing rate, making them one of the most vital aspects of daily life in today's world. According to recent surveys, it has been revealed that approximately 200,000 children in India are either visually impaired or completely blind, with around 15,000 of them attending specialized schools for visually impaired individuals. Unfortunately, due to their physical disabilities and impairments, these children face numerous challenges when it comes to accessing education. These difficulties make it even harder for them to fully participate in academic activities and thrive in their studies. Due to this, it has large impact on countries economy also. Also, students, regardless of their physical disability, have a right to get education just like normal peoples, within their body potential and ability. However, with the rapidly increase in population and number of people with disability along with other challenges, need of the advancement in technology is also important.

This introduction to the difficulties faced by blind people often highlights the numerous challenges they encounter in daily life due to the lack of visual information. Blind individuals rely heavily on auditory and tactile cues, making traditional interfaces like screens and keyboards less accessible. This can hinder their ability to navigate digital environments, access information, and communicate effectively.

In our Research paper, questions and options are delivered with help of voice in real-time, making it attainable for visually impaired students. This system helps students with the automated keyboard and Text-To-Speech system. It could be beneficial process, providing better results and relieving invigilator from these time-consuming duties.

a. Text-to-Speech System

Text-to-Speech (TTS) tools and automated keyboards play a crucial role in bridging these gaps. TTS tools convert written text into spoken words, allowing blind users to listen

to content such as articles, books, and messages, enabling greater independence and access to information through auditory means. This enables them to gather information, stay connected, and perform various tasks independently, without relying on others to read aloud for them.

b. Automated Keyboard

Automated keyboards, on the other hand, can assist in inputting text through alternative methods, such as voice commands or tactile feedback, helping blind individuals interact more efficiently with computers, smartphones, and other devices.

Voice-activated systems, for example, enable users to dictate text, search for information, and control devices with their voice, while tactile keyboards can provide a physical feedback system that helps with accurate typing.

Together, these tools significantly improve the quality of life for blind individuals by providing greater independence, so they can participate in an increasingly digital world. TTS tools and automated keyboards are key components in reducing the digital divide and empowering blind users to engage with technology on their terms.

Problem Statement

Blind individuals face significant challenges in accessing digital content and interacting with technology due to the lack of visual information. Traditional interfaces like screens and keyboards are not designed with their needs in mind, making it difficult to read, navigate, or input text independently. This creates barriers in tasks such as reading documents, browsing the internet, and communicating. To address these challenges, Text-to-Speech (TTS) tools and automated keyboards are essential. TTS tools convert written text into spoken words, while automated keyboards offer alternative input methods like voice commands or tactile feedback. These technologies are crucial for improving accessibility, promoting independence, and ensuring full participation of blind individuals in the digital world.

Existing Methods

Existing methods to assist blind people in accessing digital content and interacting with technology primarily focus on Text-to-Speech (TTS) tools, voice recognition systems, and alternative input devices.

a. Stereo Camera

The authors introduced a navigation system using an intelligent assistant called Tyflos. This system uses two vision cameras to capture images of the surrounding 3D environment, either through user command or continuous video mode. It then converts these images into verbal descriptions for communication with the user.

This system seamlessly integrates a binocular camera, an inertial measurement unit (IMU), and earphones, all built into a bicycle helmet. The binocular camera captures the user's surroundings, while the IMU tracks motion and orientation, allowing the system to understand the user's position and movements in space.

When the system detects an object at a specific location, it processes this information and converts the object's position into a corresponding sound. This sound is then transmitted to the user through the earphones, providing auditory feedback about the object's placement in the environment.

b. IP Camera Network

Chaccour et al. proposed a navigation system that uses IP cameras installed on the ceilings of rooms. The cameras capture images of the environment, which are processed by a remote system using computer vision algorithms. A mobile application allows users to input their destination and receive directions based on the visual data. While this system offers an innovative way for indoor navigation, it faces a significant challenge due to the high installation costs of multiple cameras throughout the environment. This expense makes it a less feasible solution for widespread use, especially in spaces with limited budgets.

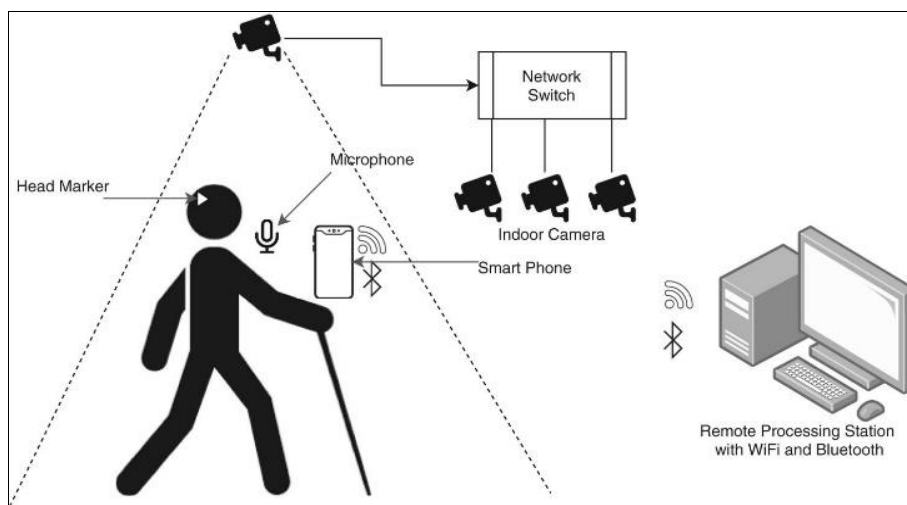


Fig-B

c. Braille Keyboards and Displays

Braille-based input methods, such as Braille displays and tactile keyboards, provide blind users with direct, tactile feedback when typing. Research often investigates ways to make these devices more portable, affordable, and responsive.

d. Gesture-Based Interaction

Some studies explore gesture recognition systems that use cameras or wearable to detect hand or body movements, allowing blind users to control devices with gestures. These systems can be paired with haptic feedback to guide users through interactions.

e. Smartphone and App Accessibility

Research into app development for blind users explores how mobile operating systems (iOS, Android) can offer better integration of accessibility features like TTS, Braille support, and voice navigation.

f. Voice Recognition Systems

Voice input systems, like those used in Google Assistant or Apple's Siri, allow blind individuals to interact with devices via speech. Research on this topic explores improvements in speech recognition accuracy, natural language

understanding, and adaptability to various accents and speech patterns.

Working Methodology of TTS (Text-to-speech)

Text-to-Speech (TTS) system converts written text into spoken words, enabling blind or visually impaired individuals to access textual content audibly. The core working methodology of a TTS system involves several stages

a. Text Preprocessing

This step involves tasks such as tokenization (breaking text into smaller units like words or sentences), normalization (converting abbreviations, numbers, and symbols into their full forms, such as "Dr." to "Doctor" or "12" to "twelve"), and text parsing (analyzing the structure of the sentence for correct pronunciation).

Example: the TTS system first tokenizes the text into smaller parts, such as:

"Once", "upon", "a", "time", ",", "in", "a", "faraway", "kingdom", etc.

It then normalizes abbreviations (though there are none here), but if there were something like "Dr." it would convert it to "Doctor".

It ensures the punctuation marks are considered, so the system knows to pause briefly after "time," and "kingdom"

b. Linguistic Analysis

This step involves deeper linguistic analysis to determine the appropriate prosody (intonation, stress, rhythm) and phonetics. The system identifies how to segment the text into phonemes (the smallest units of sound) and determines the correct prosody based on punctuation, sentence structure, and context.

Example: The system also identifies prosody, such as which words should be stressed or how the intonation should rise or fall. *For example*, it may emphasize “brave knight” to make the speech more expressive.

c. Speech Synthesis (Generation)

In this step, the system generates actual speech output. There are several approaches to speech synthesis, with the two main methods being:

▪ **Concatenative Synthesis**

This method involves piecing together pre-recorded human speech segments (such as syllables or words) stored in a large database. The system selects the appropriate segments and joins them together to form fluent speech. This method offers high-quality, natural-sounding speech but requires a large database of recordings.

▪ **Formant Synthesis**

This technique generates speech by modeling the acoustic properties of speech sounds (formants) using mathematical algorithms. While formant synthesis requires less storage and can generate speech in real-time, it tends to sound more robotic and less natural.

▪ **Neural Network-Based Synthesis**

(e.g. WaveNet): Modern TTS systems increasingly use deep learning models like WaveNet (developed by Google) or Tacotron (developed by Google DeepMind). These models learn to generate highly natural and expressive speech by training on vast datasets of human speech. Neural TTS models synthesize speech directly from text, producing more lifelike and fluid voice output.

Example: TTS engine would pull together pre-recorded chunks of speech from a database. It might select a chunk for “Once upon a time,” a chunk for “in a faraway,” and so on. The segments are then smoothly combined to sound natural.

d. Post-Processing

The final speech waveform undergoes post-processing to enhance the quality of the generated audio. This may include noise reduction, volume normalization, and smoothing to ensure the speech sounds clear and natural. The output speech is typically delivered through audio playback (such as speakers or headphones) or directly into assistive devices like screen readers or smartphones.

e. Output Delivery

The synthesized speech is then output as audio that can be heard by the user. The system may allow the user to control speech speed, pitch, volume, and even select different

voices (male or female, regional accents, etc.) for a personalized experience.

Example 3: Voice Command (Using TTS with Voice Recognition)

- **Input Command:** "Set the alarm for 7:00 AM."

1. Text Preprocessing

- The voice recognition system processes the spoken input, converts it to text (speech-to-text), and passes it to the TTS engine.
- The system might normalize the command to ensure it understands "7:00 AM" as a time, converting it into a suitable format for setting the alarm.

2. Linguistic Analysis

- The system parses the structure of the command to recognize that it's a request to set an alarm for a specific time, ensuring the correct intonation and phrasing for confirmation.

3. Speech Synthesis

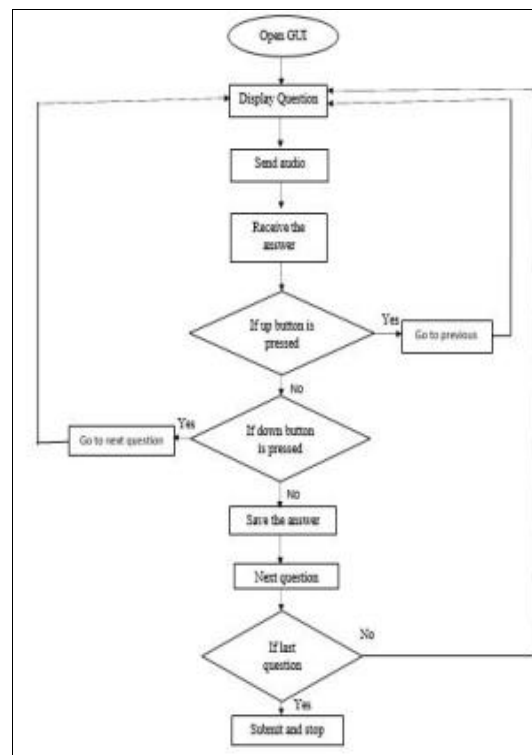
- The TTS system would respond to the user by confirming the alarm setting. For example, if the system uses concatenative synthesis, it would pull pre-recorded segments to say, “Your alarm is set for 7:00 AM.”
- If neural TTS is used, the response might be generated in real-time, sounding more natural and dynamic.

4. Post-Processing

- The synthesized speech is refined to ensure clarity, such as adjusting volume or tone if needed.

5. Output delivery

- **The user hears:** “Your alarm is set for 7:00 AM,” providing confirmation of their request.



The visual representation of input and Output

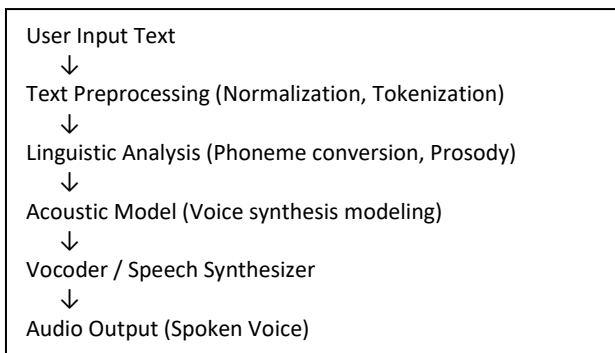
Comparison of TTS vs Automated -Keyboards

Table 1: (I) Comparison of TTS and Automated keyboard

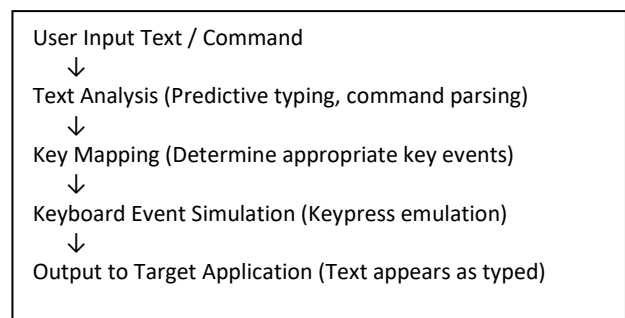
Criteria	Text-To-Speech	Automated keyboard
Independence	Allows blind students to independently read and comprehend exam questions.	Provides independence in answering questions via voice or Braille input.
Response Time	Allows students to process exam material at their own pace	Voice recognition enables faster response times for long answers
Usability	Easy to use for reading textual content, but struggles with non-textual information (e.g. equations, images).	Braille keyboards are tactile and accurate, but voice recognition may struggle with complex language.
Accuracy	High accuracy in reading text but can struggle with complex or specialized content (e.g., math equations)	Voice recognition accuracy can vary; Braille keyboards offer high accuracy for those proficient in Braille
Accessibility & Availability	TTS is widely available in digital platforms, but may require specialized adaptations for complex exam formats	Voice recognition is increasingly available, while Braille devices may be limited by cost and availability.
Cost	Generally low-cost in software implementation	Voice recognition systems may be affordable; Braille devices can be expensive.
Speed & Efficiency	Effective for reading text-based content but slower for handling complex queries	Voice recognition can speed up input, while Braille devices offer precision.

Data Flow Diagram

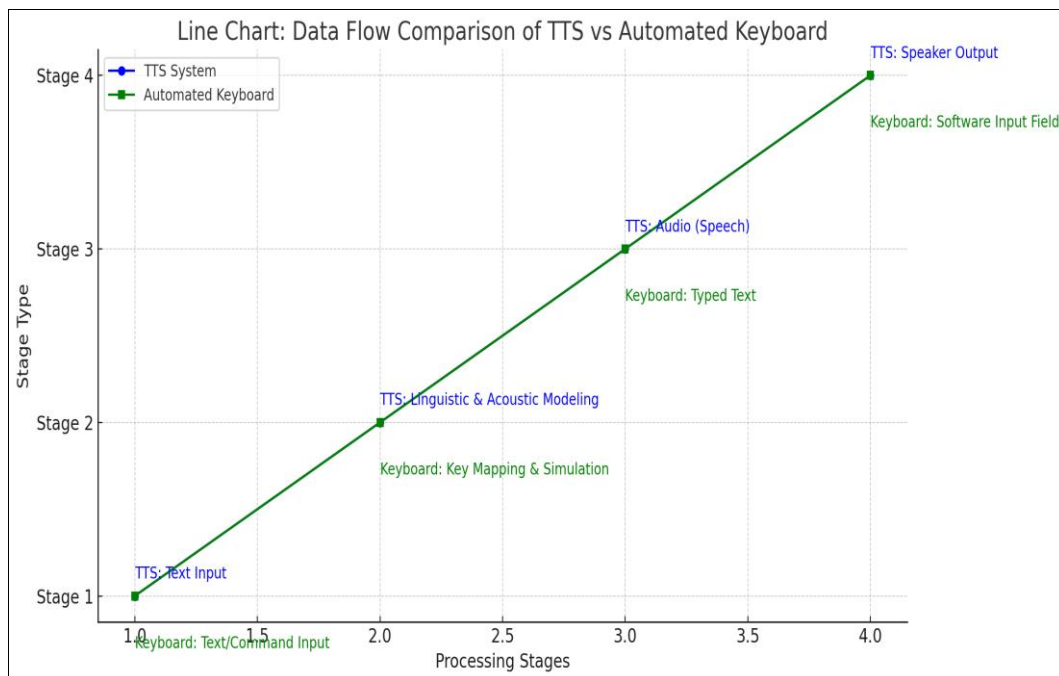
1. Text-to-Speech (TTS) System



2. Automated Keyboard System



Line-Chart- TTS vs Automated Keyboard



Result and Comparison

The integration of Text-to-Speech (TTS) systems and automated keyboards has transformed the way blind individuals participate in examinations, offering them much-needed independence and equal opportunities. These assistive technologies, when combined, address the core

challenges faced by blind students during exams—accessing written content and providing accurate, timely responses without relying on a scribe. The system enhances the examination process by offering real-time auditory feedback, where students can listen to the text at a comfortable speed, adjust the voice pitch and tone, and

control the pace of reading. This independence ensures that blind students are not disadvantaged in comparison to their sighted peers.

However, TTS systems have limitations. They may struggle with complex content, such as mathematical equations, diagrams, and other non-textual elements. Specialized systems are required to handle these intricacies, which can make some exam formats less accessible. Additionally, while modern TTS systems produce natural-sounding voices, they still require continual improvement to read complex or unfamiliar terms with perfect accuracy. Despite these challenges, TTS technology has significantly improved exam accessibility for blind students.

Automated keyboards, including voice recognition systems and Braille-based input devices, offer blind students the ability to input responses efficiently and accurately. Voice recognition allows students to dictate their answers, speeding up response times and ensuring that lengthy answers can be transcribed without delay. Braille keyboards, on the other hand, provide a tactile, accurate way for students who are familiar with Braille to type their responses, mimicking the physical act of writing.

Voice recognition systems, though convenient, may still face accuracy issues in noisy environments or with complex terminology, leading to occasional transcription errors. On the other hand, Braille keyboards provide precision and are well-suited for students who are proficient in Braille. However, these devices may have a steeper learning curve and may not be as widely available due to cost constraints. Despite these limitations, both methods have proven to be effective and supportive in enabling blind students to take exams independently.

Conclusion

The integration of assistive technologies, such as Text-to-Speech (TTS) systems and automated keyboards, is crucial in enabling blind individuals to participate fully in examinations without the need for a scribe. These tools empower blind students to independently access exam content, input responses, and navigate digital examination interfaces, thereby ensuring fairness, accessibility, and inclusivity.

Text-to-Speech (TTS) systems allow blind students to listen to the exam questions, instructions, and any additional text-based content, removing the dependency on a human scribe. Moreover, TTS systems can read out options in multiple-choice questions, long-form passages, or problem statements, ensuring that blind students can engage with the material in real-time without missing any vital information.

The use of automated keyboards, including voice-controlled input systems or tactile keyboards, allows blind users to input their answers efficiently. Automated keyboards can include voice recognition systems, where the student can dictate their answers, or alternative tactile devices, such as Braille keyboards, which give direct feedback and allow accurate typing. With advancements in speech recognition technology and neural networks, these systems can transcribe spoken words into text with high accuracy, making it possible for blind students to complete exams independently without relying on manual writing or assistance.

Together, these technologies offer an alternative and equitable method for blind students to participate in exams. They facilitate the complete independence of the user,

reducing the need for human intervention (like a scribe or reader), thus promoting a more inclusive educational environment. These tools not only ensure equal access to exam content but also foster greater autonomy and confidence in blind individuals, allowing them to demonstrate their knowledge and skills effectively, just as sighted students do.

In conclusion, the use of Text-to-Speech systems and automated keyboards can substantially enhance the examination experience for blind individuals, ensuring that they are not disadvantaged by their visual impairment. As these technologies continue to evolve and become more sophisticated, they will play a vital role in promoting inclusivity and equality in education, allowing blind students to engage in assessments independently, confidently, and without barriers.

Together, Automated Keyboards and TTS supplement one another: TTS facilitates inputting information by sound, and Automated Keyboards facilitate outputting and managing interfaces by simulated input.

Text-to-Speech (TTS) and Automated Keyboard systems are crucial in assisting blind people by increasing accessibility and independence in digital spaces. Text-to-Speech (TTS) enables blind users to read and comprehend written information by transforming text into speech. It offers an auditory interface for reading messages, navigating information, and browsing through applications without having to look at the screen.

Automated Keyboard systems assist users in entering and communicating with digital devices through key simulation. It supports effective communication, task performance, and app management—particularly when combined with screen readers and voice input.

References

1. Bourbakis NG, Kavraki D, "An intelligent assistant for navigation of visually impaired people," in Proceedings 2nd Annual IEEE International Symposium on Bioinformatics and Bioengineering, IEEE, 2001, 230–235. (https://ieeexplore.ieee.org/document/974434?utm_source=tfo&getft_integrator=tfo)
2. Chaccour K, Badr G, "Novel indoor navigation system for visually impaired blind people," in International Conference on Applied Research in Computer Science and Engineering (ICAR), IEEE, 2015, 1-5 (https://ieeexplore.ieee.org/document/7338143?utm_source=tfo&getft_integrator=tfo)
3. An Application for Blind using Text to Speech Conversion (willson joseph) (https://www.academia.edu/84079063/An_Application_for_Blind_using_Text_to_Speech_Conversion)
4. International Journal for Research in Applied Science Engineering Technology (IJRASET), 2022. (<https://doi.org/10.22214/ijraset.2022.45740>) \
5. Bourbakis NG, Kavraki D, "An intelligent assistant for navigation of visually impaired people," in Proceedings 2nd Annual IEEE International Symposium on Bioinformatics and Bioengineering (BIBE 2001), IEEE, 2001, 230–235. (https://ieeexplore.ieee.org/document/974434?utm_source=tfo&getft_integrator=tfo)
6. Jennifer MF, Holly T. Tu, is your web page accessible? a comparative study of methods for assessing web page accessibility for the blind. in Proceedings of the

- SIGCHI Conference on Human Factors in Computing Systems, 2005,
7. Prof Vidhyashree c, Supriya AM, Supriya H, Vedala Dinesh, Kavya R. International Research Journal of Engineering and Technology (IRJET), 2020, 07(07). (<https://www.ijraset.com/research-paper/ml-based-text-to-speech-converter-for-visually-impaired>)
 8. Sankaran1 AR. Barath Priyadarshan2 A. Mohamed Thaaric2, S. Nataraj authors of their research papers discussed in their thesis of IOSR Journal of Engineering (IOSR JEN) (<https://iosrjen.org/Papers/Conf.ICETETR-2019/Volume-4/EEE/11.%2046-51.pdf>)
 9. Bourbakis NG, Kavraki D, “An intelligent assistant for navigation of visually impaired people,” in Proceedings 2nd Annual IEEE International Symposium on Bioinformatics and Bioengineering, IEEE, 2001, 230–235. (https://ieeexplore.ieee.org/document/974434?utm_source=tfo&gettf_integrator=tfo)
 10. Chaccour K, Badr G, “Novel indoor navigation system for visually impaired and blind people,” in International Conference on Applied Research in Computer Science and Engineering (ICAR), IEEE, 2015, 1-5. (https://www.researchgate.net/publication/326224279_Text-to-Speech_Device_for_Visually_Impaired_People)
 11. Prof. Vidhyashree c, Supriya A M, Supriya H, Vedala Dinesh, Kavya R. International Research Journal of Engineering and Technology (IRJET), Volume: 07 Issue: 07 July 2020 (<https://www.ijraset.com/research-paper/ml-based-text-to-speech-converter-for-visually-impaired>)