The American Journal of Engineering and Technology
ISSN 2689-0984 | Open Access

TYPE Original Research
PAGE NO. 5-10
DOI 10.37547/tajet/Volume07Issue01-02

Check for updates

# Voice interface as a new challenge: how to test applications for users with screen readers

Nikita Klimov

Senior QA / QC Engineer in ADP inc. Miami, United States

**OPEN ACCESS**

SUBMITTED 20 October 2024
ACCEPTED 13 December 2024
PUBLISHED 04 January 2025
VOLUME Vol.07 Issue01 2025

CITATION

Nikita Klimov. (2025). Voice interface as a new challenge: how to test applications for users with screen readers. The American Journal of Engineering and Technology, 7(01), 5–10. https://doi.org/10.37547/tajet/Volume07Issue01-02



**Abstract:** The article identifies and describes testing methods that take into account the specifics of perception and orientation in the interface of such users. Interaction with voice interfaces creates a number of difficulties for users who use screen readers, as it requires the development of special approaches to create an accessible and user-friendly interface. Throughout the course of this work, the functional and cognitive aspects of voice interaction were investigated, compatibility with screen readers was evaluated, and synchronization of audio streams was studied.

Scientific articles were used as methodological sources, and for the practical part of the work, data contained in open access on the Internet was used, which made it possible to widely consider the chosen topic and form one's own position on the topic under consideration.

The results of the analysis showed that improved audio stream management and command settings reduce the burden on information perception and make it easier to interact with the interface. The conclusion of the work emphasizes the necessity of conducting tests in real-world conditions, considering the characteristics of the target audience.

The information presented in this article will be valuable to developers, designers, and testers who are interested in creating inclusive applications for visually impaired users.

**Keywords:** Voice interface, screen reader, accessibility, testing, cognitive load, synchronization of audio streams, inclusive design.

**Introduction:** Modern voice interface systems are evolving, expanding the range of capabilities available to users, particularly those with visual impairments. Voice

commands have become a key method of interacting with devices, providing essential solutions for daily use. For individuals using screen readers—software that converts text into audio format—it is necessary to consider specific aspects of information perception and processing. The audio mode of interaction differs significantly from visual modes, making voice interfaces an essential part of accessible technology.

Voice interfaces are particularly important for users with visual impairments, as traditional visual interactions pose substantial challenges. Screen readers serve as the primary tool, converting text content into audio output.

Research indicates that as the number of voice-controlled applications increases, so does the need to develop inclusive interfaces for users with disabilities. Developers strive to create accessible designs that consider how screen reader users perceive interfaces and analyze the barriers they encounter. Complexities such as managing parallel audio streams and improving command recognition accuarcy require careful consideration. To minimize cognitive difficulties and enhance ease of perception, specific testing methods and interface development approaches are essential.

The primary aim of this work is to investigate and establish a methodology for testing voice interfaces designed for users who rely on screen reader technology.

## METHODS

The preparation of this work involved systematic analysis and a comparative study of scientific literature on the development of voice interfaces designed for screen reader users. The research included studies addressing perception, cognitive load, accessibility of voice interfaces, and their interaction with screen readers. Functional analysis was employed to assess the effectiveness of interfaces under conditions of linear audio content perception, while cognitive analysis was used to evaluate the ease of use. Additionally, the method of observation was utilized, involving the modeling of real-life scenarios of voice interface usage. This comprehensive approach helped identify specific features and primary challenges in testing interfaces for screen reader users, establishing a practical foundation for refining testing methodologies.

The study by Vanukuru R. [1] demonstrated that spatial audio interfaces with simultaneous speech playback functions increase efficiency for users with visual impairments. Such technologies are used in screen readers to facilitate information search and display, making these interfaces more accessible for users with vision problems.

The article by Abdolrahmani A. et al. [2] emphasized the experience of blind users in mastering voice interfaces. Understanding their skills allows for consideration of specific needs during the development of voice user interfaces (VUI), enhancing accessibility for various user groups, including individuals with other disabilities.

Phutane M. et al. [3] investigated the perception of voice assistants by screen reader users. Experiments conducted in the study showed that personalized roles, such as "friend" or "expert," improve the perception of the interface, making it more comprehensible.

Yoshimura K. [4] proposed the KaraokeVUI system, where karaoke-style visual cues simplify navigation through voice commands, positively affecting command execution accuracy and improving interface understanding.

Iniguez-Carrillo A. L. et al. [5] conducted a survey-based analysis evaluating the usability of voice interfaces. The study highlighted optimal tools for assessing user satisfaction and efficiency, underscoring the need for developing VUI testing methodologies that enhance the quality of user interaction.

Guglielmi E. et al. [6] proposed an adapted method for testing voice interfaces based on approaches previously used for chatbots. The methodology presented by the authors generates numerous phrase variants, which helps identify interface errors and increases testing accuracy.

Microsoft's study outlines a comprehensive process for integrating accessibility into its technologies. The information, available on their official website [7], describes how the organization addresses barriers and creates an inclusive environment for users with disabilities, providing a basis for further research into specific technologies and approaches in accessible product design.

The next example considered was Apple's implementation of the VoiceOver feature. The data presented on their official website [8] provides detailed insights into how this screen access tool demonstrates the use of intuitive technologies to create user-friendly and functional interfaces, essential for developing visually accessible products.

Google's list of Android accessibility features, as presented on its official website [9], illustrates a wide range of tools designed to enhance accessibility. This source is significant for the practical study of how the operating system-level tools facilitate interaction for users with disabilities. It also highlights key aspects of designing accessible interfaces.

Each of these studies contributes a unique perspective on the development of user-friendly voice interfaces for

screen reader users, making their implementation in modern VUIs practical and effective.

## RESULTS AND DISCUSSION

The challenge of testing voice interfaces for screen reader users is one of the key issues in developing digital products aimed at people with special needs. The specific nature of information perception through sequential audio output, as in the case of screen readers, imposes additional requirements on the design and testing of voice interfaces [1]. Table 1 presents the challenges in the interaction between screen readers and voice interfaces.

Table 1. Challenges in the interaction between screen readers and voice interfaces [1].

| Challenge | Description |
|---|---|
| Linear perception of audio content | The use of screen readers limits the ability to perceive information in parallel, making it impossible to view the entire content structure at once. Users are forced to process data linearly, which alters navigation patterns and creates requirements for concise and clear voice interface content. Unlike visual perception, where interface elements are presented in a unified space, the voice interface must adapt to sequential output, avoiding excessive repetition and unnecessary details. This linearity imposes additional constraints on cognitive processing. Users must retain information in working memory and track the sequence of commands to form a complete understanding of the interface. Poorly structured information delivery can lead to cognitive overload, complicating navigation and causing rapid fatigue. |
| Conflict of audio streams | One of the significant barriers to creating an accessible voice interface is the need to manage concurrent audio streams. Simultaneous data output from both the screen reader and the voice interface can disorient the user, causing audio signal overlap. Reducing cognitive load is possible through proper segmentation and prioritization of audio streams. Audio stream management protocols may include timing markers for controlling pauses and a priority system for sound management. |
| Reliability of speech recognition | Adaptive speech recognition is a key element of an accessible voice interface, especially when interacting with screen reader users. Inaccurate speech recognition can lead to user frustration and completely disrupt navigation, as access to the interface relies on the system's ability to correctly interpret and process voice commands. The interface must demonstrate high resilience to variations in speech signals, such as accent, intonation, and speech rate. Testing the system in real-world conditions, including background noise levels and voice variability, enhances its adaptability and accuracy. |

Figure 1 illustrates the methods for testing voice interfaces tailored for screen reader users.

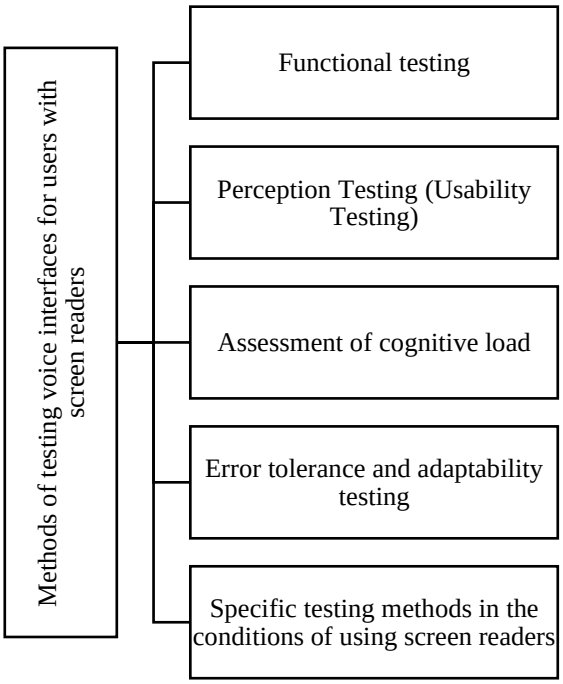


Fig.1. Methods of testing voice interfaces for users with screen readers [1].

As shown in Figure 1, functional testing of voice interfaces requires an assessment of the correct execution of all commands available to the user. It is necessary to ensure that the interface responds to commands sequentially, avoiding confusing or verbose replies. Information blocks should be divided and structured so that the screen reader does not duplicate unnecessary content, which is especially crucial when navigating large datasets.

An effective method of functional testing is conducting audits with experienced screen reader users who can suggest improvements based on real user experience. This approach helps identify weaknesses in the functional sequence of the interface and detect command logic that may cause difficulties.

Usability testing evaluates the ease and intuitiveness of using the interface from the perspective of screen reader users. A key task here is to assess the clarity and simplicity of navigation and analyze the need for repeated interactions. To measure usability, experts recommend using cognitive analysis methods, which help track user memory load, time spent on tasks, and error rates.

Contextual testing, involving multiple checks across various usage scenarios, helps accurately assess usability. It is essential to model test conditions, taking into account the real environment where the interface may be used, as factors like background noise and surroundings can significantly affect user perception.

Cognitive load analysis plays a central role in testing voice interfaces. Methods such as the NASA Task Load Index (NASA-TLX) can be adapted to evaluate task complexity and its impact on the cognitive perception of screen reader users. Cognitive testing helps optimize the amount of information delivered and enhances the interaction structure, preventing user overload [3].

To optimize the cognitive load, voice commands requiring active user participation should be minimized, and informational blocks should be divided into several messages. It is essential to provide options for repeat requests and feedback, allowing users to replay or clarify information without losing context.

The voice interface system should be resilient to errors caused by incorrect command recognition and offer the possibility for re-entering information without disrupting the interaction. The system needs to be assessed for its response to random sound signals to determine whether such signals interfere with accurate speech recognition. Mechanisms for auto-correction and prompts should be configured to help users quickly return to their current task [2].

The system's adaptability should also be evaluated: the voice interface must be capable of recognizing a wide range of voice patterns, including different timbres, accents, and variations in speech speed. Adapting the system to the user's unique voice characteristics can reduce recognition errors and enhance the accessibility and accuracy of the interface.

To ensure maximum accessibility, testing should be conducted using several popular screen readers, such as JAWS, NVDA, and VoiceOver. Differences in software platforms can affect user perception and navigation, making compatibility testing crucial to avoid unexpected conflicts. It is important to consider the

specific features of different operating systems and hardware configurations, as the interaction between the screen reader and the voice interface may vary.

The emotional tone of the voice interface also influences user perception. Aspects such as intonation, voice timbre, and speech fluidity impact the emotional response of the user and the comprehension of the information. Testing with the target audience can help identify optimal parameters for voice output, thereby enhanc the overall user experience. For instance, determining user preferences for the pace of information delivery can optimize the interface performance for individuals with different processing speeds [5].

The following relevant methods are illustrated in Figure 2.

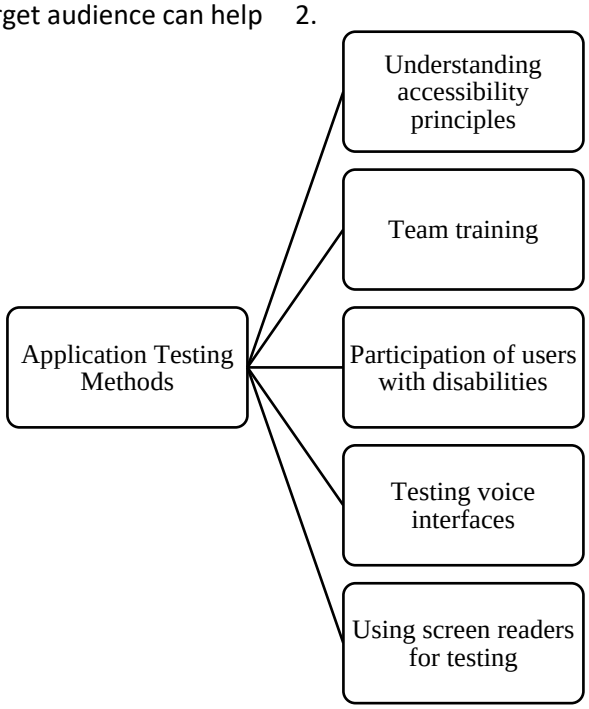


Fig. 2. Application testing methods [5-6].

The data presented in Figure 2 will be examined in greater detail to enhance understanding of its features.

1. Understanding Accessibility Principles: WCAG 2.1 (Web Content Accessibility Guidelines) provides standards for creating accessible applications.

2. Using Screen Readers for Testing:

- Testing should involve commonly used screen readers such as JAWS, NVDA for Windows, VoiceOver for macOS and iOS, and TalkBack for Android.
- Device-based testing helps identify specific issues that may not appear in emulators.

3. Testing Voice Interfaces:

Developing and testing scenarios where users interact with the application via voice commands allows for an assessment of usability and the accuracy of speech recognition.

It is essential to evaluate how effectively the application provides voice feedback, particularly in complex situations or when errors occur.

4. Involving Users with Disabilities:

Engaging visually impaired users in testing helps uncoverissues that might be overlooked by developers and testers.

5. Team Training:

Educating team members on accessibility principles and the use of screen readers contributes to the development of more inclusive products [6].

Adhering to these recommendations ensures application quality for users who rely on screen readers and voice interfaces, increasing user satisfaction and broadening the product's audience. As examples, consider how Microsoft, Apple, and Google achieve this goal.

1. Microsoft: The company integrates accessibility principles across all its products. Microsoft developed its own screen reader, Narrator, integrated into the operating system to provide accessibility for visually impaired users. Additionally, the company offers tools and guidelines for developers, including Accessibility Insights for testing accessibility [7].

2. Apple: The VoiceOver feature, built into iOS and macOS, allows visually impaired users to interact with devices. Apple provides developers with guidelines and tools such as the Accessibility Inspector [8].

3. Google: Google incorporates accessibility features into its products and offers tools for developers to

create accessible applications. The TalkBack screen reader for Android enables visually impaired users to interact with devices [9].

Thus, the challenges of testing voice interfaces require a multi-level approach.

## CONCLUSION

The research highlights the importance of adapted testing methods for voice interfaces aimed at enhancing accessibility for users relying on screen readers. An analysis of functional and cognitive testing approaches reveals that the linear perception of audio information and the limited capabilities for parallel data processing necessitate careful attention to the design and development of interfaces.

The main challenges include managing audio streams, ensuring speech recognition accuracy, and minimizing cognitive load — tasks addressed through specialized testing methods, such as audio stream synchronization and adaptive navigation protocols. Testing has demonstrated the effectiveness of these methods in creating inclusive interfaces that improve the user experience for individuals with visual impairments.

## REFERENCES

Vanukuru R. Accessible spatial audio interfaces: A pilot study into screen readers with concurrent speech //Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. – 2020. – pp. 1-6.

Abdolrahmani A. et al. Blind people are power users: An argument for centering blind users in the design of voice interfaces //UMBC Student Collection. – 2020.

Phutane M. et al. Speaking with My Screen Reader: Using Audio Fictions to Explore Conversational Access to Interfaces //Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility. – 2023. – pp. 1-18.

Yoshimura K. KaraokeVUI: Utilizing Karaoke Subtitles for Voice User Interfaces to Navigate Users What They Would Say //CHI Conference on Human Factors in Computing Systems Extended Abstracts. – 2022. – pp. 1-4.

Iniguez-Carrillo A. L. et al. Usability questionnaires to evaluate voice user interfaces //IEEE Latin America Transactions. – 2021. – Vol. 19. – No. 9. – pp. 1468-1477.

Guglielmi E. et al. Sorry, I don't Understand: Improving Voice User Interface Testing //Proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering. - 2022. – pp. 1-12.

Bridging the disability divide: Microsoft's ongoing accessibility and inclusion journey. [Electronic resource] Access mode: https://news.microsoft.com/en-au/features/bridging-the-disability-divide-microsofts-ongoing-accessibility-and-inclusion-journey / (accessed 07.11.2024).

VoiceOver. [Electronic resource] Access mode:https://developer.apple.com/documentation/accessibility/voiceover?language=objc (accessed 07.11.2024).

What special features are available on Android devices. [Electronic resource] Access mode: https://support.google.com/accessibility/answer/6006564 ?hl=ru-ru (accessed 07.11.2024).