Universal Design 2024: Shaping a Sustainable, Equitable and Resilient Future for All
K.S. Fuglerud et al. (Eds.)

doi:10.3233/SHTI241038

Mastering Gesture-Based Screen Readers on Mobile Devices - Exploring Teaching and Practice Strategies

Joschua Thomas SIMON-LIEDTKE[a,1] Kristin Skeide FUGLERUD[a], Eva Elida SKRÅMESTØ[b] , and Wolfgang LEISTER[a]
[a] *Department of Applied Research in Technology (DART), Norsk Regnesentral (NR), Norway*
[b] *Norges Blindeforbund, Norway*
ORCiD ID: Joschua Thomas Simon-Liedtke https://orcid.org/0000-0003-4809-1688, Kristin Skeide Fuglerud https://orcid.org/0000-0002-5648-0264, Wolfgang Leister https://orcid.org/0000-0002-8310-2031

Abstract. Gesture-based screen readers like VoiceOver or TalkBack provide visually impaired users with a means to interact with digital content. However, there is a significant lack of both strategies and resources for teaching the use of these screen readers, and standardized teaching guidelines are notably absent. Furthermore, there is no free, universally designed, and accessible app for practicing gestures in mobile screen readers. This study aims to identify best practice strategies for teaching and practicing the use of gesture-based screen readers among visually impaired users, based on observations from an IT course directed at visually impaired individuals. Moreover, we present common challenges related to usability, attitudes, emotions, technical aspects, and user guidance and education, as well as key traits and facilitators for learning gesture-based screen readers. Lastly, we assess the feasibility of an app to practice gestures and propose a framework for a gesture practice app to enhance user accuracy and patience.

Keywords. Visual Disability; Screen Reader; Gesture; Adult Education; Mobile Technology; Accessibility; Usability; Universal Design

1. Introduction

Screen readers, such as VoiceOver (Apple) and TalkBack (Android), are key tools for visually impaired users in mobile technology [1–3]. These tools enable interaction with digital content through audible narration of textual and graphical elements, employing various finger-swiping gestures like swiping, tapping, or rotating for functions such as reading, writing, navigating, or adjusting settings [1–5]. Although these gesture-based screen readers offer numerous benefits and advantages, and their gestures are presumably simple and natural, certain limitations persist [6–10]. Mastering these gestures requires practice and patience, with the goal of ingraining these movements into the user's muscle memory for increased comfort and consistency [6–9]. Despite the availability of various training resources from producers, institutions, and public authorities for learning and using gesture-based screen readers [1–5,11,12], these resources are often not widely

[1] Corresponding Author: Joschua Thomas Simon-Liedtke, joschua@nr.no.

known, difficult for beginners of screen readers to access, and lack proper educational formatting. Likewise, research provides only a few strategies to teach gesture-based screen readers, including strategies based on auditory and verbal feedback and gamification [6,13–15]. Thus, significant challenges remain, related to the absence of consistent and standardized guidelines and strategies for teaching gestures in gesture-based screen readers [6,16]. Also, a free, and universally designed practice app for visually impaired people has not been identified so far [6,14].

This study aims to identify the best strategies for teaching and practicing gesture-based screen readers to visually impaired users and assess if a gesture practice app can simplify learning and boost comfort and skill. We obtained insights from silently observing an IT course for the visually impaired.

This article outlines our data collection and analysis methods. We will then present our qualitative findings and discuss potential future research directions.

2. Methodology and implementation

We participated in a four-day in-person IT course, facilitated by the Norwegian Association of the Blind and Partially Sighted (NAPB) at their continuing education center located in Northern Norway. The course emphasized the use of digital media and screen readers tailored for individuals with visual impairments. Two main instructors and six assistants oversaw one-on-one and group training sessions. All of them were peer supporters, either having disabilities or being closely related to visually impaired individuals. Fifteen participants attended the ICT course, aged late twenties to seventies, with visual impairments ranging from total blindness to residual vision. Their experiences with ICT varied from beginner to expert levels.

The research team, consisting of three members, attended the course in-person as non-intrusive observers, monitoring individual sessions and conducting informal, unstructured interviews with instructors, assistants, and participants during breaks and free time. We scrutinized both attendee learning and instructor teaching methods during each session. The data from these observations and interviews were compiled and thematically analyzed, identifying recurring patterns among attendees. The patterns were cross-examined among the researchers, leading to the extraction of themes related to strategies and challenges in teaching screen reader gestures, and the potential of a gesture practice app. We focused primarily on iOS's VoiceOver due to its popularity among individuals with visual impairments, as also confirmed by the NABP. These themes were subsequently organized and summarized.

All participants were members of the NABP, a partner in the research project [17]. Participants were informed about the project’s scope, team, objectives, and data type, and were aware of their voluntary participation and right to withdraw consent. To ensure participant privacy, no personal data or audiovisual recordings of the participants were collected. Thus, all collected data was anonymous.

3. Teaching-strategies for gesture-based screen readers

A few weeks prior to the IT course, instructors conducted remote sessions with participants to understand their motivation, assess their skills, inquire about commonly used technology and assistive devices, and discuss expectations and desired outcomes. Depending on the participant's comfort and familiarity with technology, these sessions

were conducted via video conferencing or phone. Instructors individually engaged with each participant at the start of the four-day course, adjusting learning objectives based on their preferences and needs, and demonstrating basic functions in VoiceOver. Specific learning goals were set as informal agreements between instructors and participants.

Participants had two or three individual training sessions per day, each lasting either 45 or 90 minutes. Breaks were scheduled for activities, practice, or socializing, during which many participants shared their experiences, learning from and teaching each other. At the end of each day, common topics were discussed in joint sessions, including useful apps or devices, assistive technology, digital skills training funding schemes, and online security strategies. Each learning session focused on achieving previously set goals, usually practicing specific functions. To teach gestures, instructors demonstrated in the air, on a flat surface, traced on the participant's body, like the back of the hand or shoulder, or guided the participant's finger.

For beginners, instructors began with the fundamental concepts of VoiceOver: navigation, actions, and gestures. Navigation involves selecting functions and elements, actions pertain to executing tasks like messaging, and gestures are finger movements used for navigation and action execution. They differentiated VoiceOver from Siri, Apple's digital assistant, underlining that they operate differently and independently, requiring separate learning. Participants were then taught to de/activate VoiceOver and basic gestures like swiping and tapping, with practical examples. For more experienced users, advanced functions like the reverse "Z" gesture, and rotor function were covered. The VoiceOver rotor function, activated by a two-finger rotation, provides advanced shortcuts like direct navigation, and typo correction for experienced users. Throughout the learning sessions, participants had the opportunity to ask general questions or adjust their learning goals.

4. Common challenges when using and learning screen readers

The study identified various challenges encountered by participants and instructors.

4.1. Usability

Users faced usability challenges with VoiceOver gestures due to the similarity of multiple movements and their executions. Gestures such as double-tapping with two or three fingers or triple-tapping with three fingers were often mixed up, leading to unexpected phone behavior and incorrect actions. The confusion occurred because users frequently thought they were making one gesture, while the phone recognized it as a different one. Likewise, precisely performing gesture was challenging when initial finger placement was incorrect, inappropriate swiping speed, or swiping beyond the intended screen area. The complexity of performing actions in the user interface with VoiceOver, often required multiple steps and multiple and varied gestures across applications. This presented significant challenges, leading to users desiring more streamlined processes for performing similar actions.

The identical default voices of Siri and VoiceOver often led to confusion among users of both VoiceOver and Siri. Some users stated that they would prefer different voices to distinguish between these two functions. Moreover, many users were unaware of the auditory feedback provided by VoiceOver, such as sounds indicating skipped elements or reaching the end of a page or menu.

Frequent updates to apps and software, which alter functionality and user interfaces, were overwhelming and challenging for blind and visually impaired users who must adapt to new interfaces and relearn app navigation and functions.

4.2. Attitude and emotions

Participants often showed impatience or hastiness while interacting with VoiceOver, leading to skipping over crucial elements, accidental activations, and a loss of orientation due to swiping too quickly or not waiting for full readouts.

Privacy concerns arose from VoiceOver potentially exposing sensitive information when used without headphones. Similarly, users felt exposed and embarrassed by the audible reading of personal data and desired a headlines-only option for privacy in public settings. While some preferred headphones or earplugs to minimize disruption, others worried about losing them.

Participants exhibited varied emotional responses, including feelings of irritation, stress, anxiety, discomfort, and embarrassment. Frustration led some to disable the feature, while others experienced anxiety when unable to perform tasks or faced with unfamiliar notifications like update requests. Excessive text read-outs or surrounding noise overwhelmed some participants.

4.3. Technical challenges specific to VoiceOver

VoiceOver occasionally failed to fully read on-screen content due to apps and websites not being universally designed for screen reader use, often requiring assistance from more knowledgeable individuals or those without impairments.

Users found the autocorrect function in VoiceOver challenging, as it was often imprecise and difficult to use for correcting misspelled words, leading some to delete and retype words entirely. Moreover, adding words to the dictionary was difficult for many participants.

Users often found it imprecise and cumbersome to use the screen keyboard with VoiceOver, specifically the default setting to double tap for each letter, and to remember to shift focus away from the keyboard to conduct other tasks.

The VoiceOver rotor function was often triggered accidentally due to incorrect execution of other gestures. This frequently resulted in unintended changes, such as switching languages and altering reading speech. Lastly, VoiceOver's grouping behavior option confused users by consolidating all menu items under one single heading, which resulted in skipping the reading of some options. Changing settings for VoiceOver was frequently challenging and confusing as many users were unaware of the accessibility settings and how to modify them.

4.4. User guidance and educational challenges

The participants desired more opportunities to practice. Although Apple provides a practice area within the VoiceOver settings, it is somewhat hidden, assumes users already have basic knowledge of the available gestures and lacks pedagogical support. The practice area consists of a plain canvas that only verifies and verbalizes performed gestures without offering explanations or guides. This often led to confusion, and several participants abandoned the feature,

Although manuals exist [1,4,18], many participants found them lengthy and inaccessible. Learning about screen readers often paradoxically requires using the same screen reader. They preferred a short, clear user guide or "cheat sheet." Such a "cheat sheet" could list key functions, explaining which fingers are used for different gestures, ideally presented in an accessible format with audio descriptions.

The lack of suitable feedback mechanisms in the practice area left participants often uninformed or unaware when an error occurred. Auditory feedback often did not clarify why gestures did not produce the expected outcome. Moreover, there was no straightforward way to report functionality VoiceOver issues to Apple.

Many municipalities reportedly lack sufficient digital skill training ICT courses [19]. Despite being mandated by law, the availability of dedicated experts to assist and support individuals with (visual) impairments varies. At the same time, many visually impaired individuals are reluctant to ask for help, fearing loss of dignity.

5. Key traits and facilitators for learning gesture-based screen readers

Using gesture-based screen readers requires precision and patience for accurate gesture execution and attentive feedback listening. Without patiently listening to auditory feedback correct gesture execution cannot be confirmed, and erroneous gesture execution cannot be detected. Instructors suggested that participants could practice precision and patience by selecting apps for daily use such as checking the weather. Repeated practice, suggested with the screen curtain on to avoid reliance on residual vision, aims to commit gestures to muscle memory.

Participants favored instructor-led training for its guidance and interactive nature. Seeing even experts struggle at times alleviated their tech and performance anxieties, and they appreciated skilled and didactically proficient instructors. Likewise, participants favored one-on-one instruction to avoid overwhelming beginners or under-stimulating advanced users in mixed-level groups. Casual break-time discussions were also seen as valuable for skill development.

6. Feasibility and design of a gesture-practice app

While participants favored personalized training, they expressed curiosity towards a gesture practice app as a supplementary self-study tool following formal training. They saw potential in such an app as a reference source, containing key information and tutorials to reinforce their practical skills. Recognizing the benefits of repetition, they also stressed the necessity of universal design within the app.

Participants proposed features for a gesture-practice app based on Apple iOS's current practice area, emphasizing universal design, accessibility, and usability for all, including those with minimal experience with gesture-based screen readers. They recommended high contrast, readable text, adjustable font sizes, and read-aloud functionality. Apple's plain canvas approach for free gesture practice could be extended by adding detailed explanations and guides of available gestures, clarifying how and where to use them, and making them searchable. Participants also requested a comprehensible and searchable gesture glossary, or "cheat sheet," accessible both in the practice area and during general device use, possibly via a dedicated on-screen button activated by a gesture or voice command. The practice area could offer instructional guidance, including step-by-step walkthroughs for specific gestures and options for

repeated practice, incorporating examples of specific actions like pressing buttons or opening links. Participants suggested using auditory or vibrational feedback for correct finger placement and gesture execution. Additionally, an optional feedback system should be available to clarify unsuccessful gestures and guide correct performance. Incorporating gamification and reward elements into the app could counter monotony in repetitive gesture practice. Participants advocated for a gradual learning approach, progressing from simple to complex gestures with milestones along the way. Challenges like timed navigation to specific app parts or error-free gesture execution could make learning interactive. Auditory badges or music snippets could serve as a sense of achievement.

Participants proposed a stepwise approach for a possible gesture practice app, progressing from basic to complex screen reader gestures, as categorized by participants and instructors. They suggested four levels based on relevance, frequency, and complexity: basic beginner, practical usage, advanced shortcuts, and expert users, with many advanced gestures replaceable with simpler combinations. The gestures listed below were identified as the most essential for VoiceOver [5,11,12]. Similar functions could be chosen for other screen readers.

1. Level One - Beginner's basics:
    - Turn VoiceOver on/off: Triple-click the "Home" button or the side/top button. The initial setup through settings is required.
    - Turn the screen curtain on/off: three-finger triple tap.
    - Mute/unmute VoiceOver: Three-finger double tap.
    - Home gesture: Swipe up from the screen's bottom edge to return to the home screen on newer devices without a "Home" button.
2. Level Two - Practical usage:
    - Select the previous/next item: Swipe left or right with one finger.
    - Activate a selected item; Double tap with one finger.
3. Level Three - Advanced shortcuts:
    - Scroll left/right: Three fingers to the right and left, respectively.
    - Scroll up/down: Three fingers down and up, respectively.
    - Back: two-finger Z gesture.
4. Level Four - Expert techniques:
    - Choose a new rotor function: Two-finger rotation to the left/right.
    - Change value of rotor setting: Swipe up or down.
    - Turn dictation on/off: Two-finger double tap in a text field.

7. Discussion

While previous research has discussed the benefits of gesture-based screen readers for visually impaired individuals [6–9,13], this study focused on effective learning for these tools including educational best-practices, challenges, and facilitators. In our study, we chose to focus on Apple products due to their popularity among visually impaired users in Norway[2] [1,20]. Nevertheless, our findings are applicable to other gesture-based screen readers, such as Android's TalkBack. In our opinion, best practices can be transferred between brands with minor adjustments. Our study suggests standardizing

[2] As reported by the NAPB, iOS products, favored for their accessibility and ease of use, are preferred by most screen reader users in Norway. However, it is recognized that in other regions, the preference might lean towards Android products, largely due to cost considerations.

teaching methods for gesture-based screen readers through a blend of one-on-one training, self-study, and peer support from visually impaired experts. Participants favored this approach over pure self-study using an app, valuing group practice and expert guidance for basic setup and gestures. After initial training, they found independent app practice beneficial.

Teaching gesture-based screen readers is challenging due to the inherent difficulty of learning new technology and the possible lack of general digital skills among visually impaired individuals. This issue is intensified as visual impairments often correlate with aging, a group that generally struggles with new technology. Hence, instructors often begin with general ICT skills, and then carefully balance this with specific screen reader instruction to avoid overwhelming the learner.

Our study highlights the urgent need for municipalities to invest more in the digital skill training of visually impaired individuals. Many Norwegian municipalities, despite their responsibility, fail to provide digital skills training due to a lack of experts and resources restricting their ability to supplement this training - a challenge also faced by civil-society organizations [19]. Therefore, it is crucial for research to propose concrete solutions, such as organizing workshops and awareness seminars for municipality employees, as well as improving skills and training for relevant personnel. We specifically recommend municipalities to prioritize raising awareness about the needs of visually impaired individuals and enhancing their digital training programs. Engaging politicians in this issue is also essential as they hold decision-making power over the allocation of funds. Improved digital skill training not only meets legal requirements but also fosters inclusion, curbs exclusion, and could be cost-effective, aligning with prior research [17,19]. Investment in training boosts visually impaired individuals' independence, digital participation, and social engagement, reducing both their assistance needs and potential loneliness.

In the meantime, developers could co-create a gesture-practice app together with visually impaired users. One idea is to use an open-source approach to avoid relying on enterprise resources that are difficult to access. Finally, developers and designers must ensure the adherence of their solutions to legal standards and guidelines, such as the Web Content Accessibility Guidelines (WCAG) [21], given that the lack of universal design in websites poses a key challenge for all screen readers.

8. Conclusion

Screen reader training is essential for enabling visually impaired users to interact with mobile technology. Our research underscores the need for a safe practice space, feedback, and support as key to successful training. This study can serve as a foundation for the standardization of instruction methods for gesture-based screen readers. It also proposes a gesture practice app setup to complement personalized training and general online resources. Our study indicates potential interest in such a training app among both instructors and users, as it supplements one-on-one training by allowing repeated practice to improve precision, repetition, and patience.

References

[1] Apple, Chapter 1. Introducing VoiceOver, (2024). https://www.apple.com/voiceover/info/guide/_1121.html. (accessed June 18, 2024).

[2] Apple, Turn on and practice VoiceOver on iPhone, Apple Support. (2024). https://support.apple.com/guide/iphone/turn-on-and-practice-voiceover-iph3e2e415f/ios. (accessed June 18, 2024).
[3] Google, Get started on Android with TalkBack, (2024). https://support.google.com/accessibility/android/answer/6283677. (accessed June 18, 2024).
[4] Apple, VoiceOver User Guide, (2021). https://support.apple.com/en-gb/guide/voiceover/welcome/mac. (accessed June 18, 2024).
[5] Apple, Learn VoiceOver gestures on iPhone, Apple Support. (2022). https://support.apple.com/en-gb/guide/iphone/iph3e2e2281/16.0/ios/16.0.
[6] A. Rodrigues, K. Montague, H. Nicolau, and T. Guerreiro, Getting Smartphones to Talkback: Understanding the Smartphone Adoption Process of Blind Users, in: Proc. 17th Int. ACM SIGACCESS Conf. Comput. Access., ACM, New York, NY, USA, 2015: pp. 23–32.
[7] V. Luthra, and S. Ghosh, Understanding, Evaluating and Analyzing Touch Screen Gestures for Visually Impaired Users in Mobile Environment, Univers. Access Hum.-Comput. Interact. Access Interact. (2015) 25–36.
[8] M.C. Buzzi, M. Buzzi, B. Leporini, and A. Trujillo, Exploring Visually Impaired People's Gesture Preferences for Smartphones, in: Proc. 11th Biannu. Conf. Ital. SIGCHI Chapter, Association for Computing Machinery, New York, NY, USA, 2015: pp. 94–101.
[9] L. Anthony, R.-D. Vatavu, and J.O. Wobbrock, Understanding the consistency of users' pen and finger stroke gesture articulation, in: Proc. Graph. Interface 2013, Canadian Information Processing Society, CAN, 2013: pp. 87–94.
[10] D.A. Norman, Natural user interfaces are not natural, Interactions. 17 (2010) 6–10.
[11] Apple, Learn advanced gestures to interact with iPad, Apple Support. (2024). https://support.apple.com/en-gb/guide/ipad/ipadab6772b8/ipados (accessed June 18, 2024).
[12] Deque University, VoiceOver Gestures on iOS | Screen Reader Keyboard Shortcuts and Gestures, (2022). https://dequeuniversity.com/screenreaders/voiceover-ios-shortcuts (accessed June 18, 2024).
[13] B.M. Celusnak, Teaching the iPhone with VoiceOver Accessibility to People with Visual Impairments, J Vis Impair Blind. 110 (2016) 369–372.
[14] U. Oh, S.K. Kane, and L. Findlater, Follow that sound: using sonification and corrective verbal feedback to teach touchscreen gestures, in: Proc. 15th Int. ACM SIGACCESS Conf. Comput. Access., Association for Computing Machinery, New York, NY, USA, 2013: pp. 1–8.
[15] G.F. Lobo, D. Gonçalves, P. Pais, T. Guerreiro, and A. Rodrigues, Using Games to Practice Screen Reader Gestures, in: Proc. 23rd Int. ACM SIGACCESS Conf. Comput. Access., Association for Computing Machinery, New York, NY, USA, 2021: pp. 1–4.
[16] B. Leporini, M.C. Buzzi, and M. Buzzi, Interacting with mobile devices via VoiceOver: usability and accessibility issues, in: Proc. 24th Aust. Comput.-Hum. Interact. Conf., Association for Computing Machinery, New York, NY, USA, 2012: pp. 339–348.
[17] K.S. Fuglerud, S. Tunold, and K. Kjæret, Social Contact for Older People with Visual Impairment Through Mastery of Smartphones: Barriers and Suggested Solutions, in: I. Verma (Ed.), Stud. Health Technol. Inform., IOS Press, 2021. doi:10.3233/SHTI210417.
[18] Statlig spesialpedagogisk tjeneste (Statped, State special education service), Lær å bruke iPad og iPhone for synshemmede. (Learn how to use iPad and iPhone for the visually impaired), (2022). In Norwegian, https://www.statped.no/laringsressurser/syn/opplaring-i-mobil-og-nettbrett/ipad-og-iphone-for-synshemmede/. (accessed June 18, 2024).
[19] K.S. Fuglerud, J.T. Simon-Liedtke, T. Halbach, K. Kjæret, and E.E. Skråmestø, Motvirkning av ensomhet gjennom inkludering i informasjonssamfunnet. (Counteracting loneliness through inclusion in the information society), Report 1061, ISBN 978-82-539-0571-6, Norsk Regnesentral, Oslo, 2023. In Norwegian.
[20] N. Griffin-Shirley, D.R. Banda, P.M. Ajuwon, J. Cheon, J. Lee, H.R. Park, and S.N. Lyngdoh, A Survey on the Use of Mobile Applications for People who Are Visually Impaired, J Vis Impair Blind. 111 (2017) 307–323.
[21] World Wide Web Consortium (W3C), Web Content Accessibility Guidelines (WCAG) 2.1, (2018). https://www.w3.org/TR/WCAG21/. (accessed June 18, 2024).

Acknowledgments

This project has been supported with funds from The Norwegian Directorate of Health and the Research Council of Norway with additional in-kind contributions from Norges Blindeforbund and Norsk Regnesentral.