THE JOURNAL ON TECHNOLOGY AND PERSONS WITH DISABILITIES

# User Experience of Voice Assistants by People with Visual Disabilities

Hyung Nam Kim
North Carolina A&T State University
hnkim@ncat.edu

## Abstract

As technology advances, people with visual disabilities increasingly rely on emerging technologies (e.g., Siri, VoiceOver, and Microsoft's Seeing AI application optimized for use with VoiceOver). Those technologies are powered by voice user interfaces that assist users with reading information, controlling systems, and communicating with others. Yet, there is limited research on how people with visual disabilities interact with the voice assistants, focusing on gesture commands, voice commands, and relevant user interfaces. To address the knowledge gap, this study investigated how people with visual disabilities interact with the iPhone's voice assistant features and an assistive technology app accessible through the voice assistant features. This study found that people with visual disabilities had a poor user experience, and design recommendations were provided.

## Keywords

Siri, VoiceOver, Seeing AI, user interfaces, usability, accessibility.

**Introduction**

The prevalence of visual disabilities, including visual impairment and blindness, is significant in the United States. Over one million Americans live with blindness, and over 8 million Americans are visually impaired by uncorrected refractive error (Varma et al.). The Centers for Disease Control and Prevention (CDC) has recently released a report that 93 million American adults are at high risk for serious vision loss, and only half of them were likely to visit an eye doctor. The economic cost related to vision issues is estimated to increase to $373 billion by 2050 (CDC).

Today's technological advances have brought highly accessible consumer products to people with visual disabilities in a variety of domains – healthcare, education, rehabilitation training, and Internet access (Kim). For example, a survey study by Crossland et al. found that over 80% of 132 respondents with visual disabilities used a smartphone for phone calls, texting, reading, browsing the Internet, and identifying objects. A longitudinal survey study (WebAIM) was conducted in 2013 and 2018 to assess the technology use among people with visual disabilities. The number of respondents using iOS increased from 43.1% to 64.3%, and the respondents using Android also increased from 18.1% to 23.8%. Nearly 61% of the respondents using iOS and 35.7% of the respondents using Android reported that they used voice assistants (e.g., VoiceOver, Siri) *very* or *somewhat frequently*.

Many users with visual disabilities take advantage of assistive applications (apps) such as Microsoft's Seeing AI (Dockery and Krzystolik). The Seeing AI app (powered by an artificial intelligence technology) can identify people, colors, currency, scenes, objects, and texts, and then audibly describe them for users with visual disabilities. The Seeing AI app is fully accessible with VoiceOver. Yet, when VoiceOver is on, standard touchscreen gestures (for sighted users) will have different effects, and additional gestures will become available to operate the iPhone and apps.

However, little is known about how users with visual disabilities understand and use the voice assistant features. Wong and Tan conducted a single case study to investigate how an individual with visual impairment, named Bill (aged 45 years) learned and used an iPhone. Bill encountered challenges with some apps that are not running well with VoiceOver. When errors occurred, there were no verbal prompts to help Bill to exit from the app. Celusnak, as a blind rehabilitation specialist, acknowledged that VoiceOver is a useful tool for users with visual disabilities in using an iPhone. Yet, Celusnak also argued that the learning curve is steep and very challenging; for example, the gesture command *Split Tap* was suggested to be one of the most difficult concepts for those with visual disabilities to understand and execute properly. Leporini et al. conducted a user study with a large sample (n = 55 participants with blindness) to examine the user experience of VoiceOver and found a range of usability problems; however, they merely relied on a passive approach, i.e., an online survey. There are only a handful of publications on user experience with VoiceOver (Grussenmeyer and Folmer; Park et al.; Smaradottir et al.). There is still a lack of in-depth understanding of what circumstances cause poor user experience associated with voice assistants, especially in terms of finger gesture commands, voice commands, and relevant user interfaces.

To address the knowledge gap, this study conducted observations on people with visual disabilities interacting with VoiceOver and the Seeing AI app accessible through VoiceOver.

## Methods

### *Participants*

A convenience sample of eight individuals with visual disabilities participated in this study. The inclusion criteria included 18 years of age or older and visual acuity worse than 20/70 with the best possible correction (World Health Organization). Table 1 shows the detailed characteristics of the participants.

Table 1. Characteristics of the participants

| **Participant Demographics** | **N = 8** |
|---|---|
| Visual acuity - Between 20/200 and 20/400 | 5 |
| Visual acuity - Between 20/400 and 20/1200 | 2 |
| Visual acuity - Less than 20/1200 | 1 |
| Age | 65.71±12.33 (years) |
| Gender - Male | 3 |
| Gender - Female | 5 |
| Race/Ethnicity - African American | 5 |
| Race/Ethnicity - European American | 3 |
| Education - High school or equivalent | 3 |
| Education - Associate | 4 |
| Education - Masters | 1 |
| Do you use a smartphone? | Yes = 7, No = 1 |
| How long have you used the smartphone? | 4.90±3.20 (years) |
| How frequently do you use the smartphone? | Very frequently = 4<br>Frequently = 1<br>Very rarely = 2<br>Never = 1 |

*Materials*

An iPhone 12 mini was used in teaching participants how to use VoiceOver and the Seeing AI app, and their interactions were video recorded for further analysis. The tutorial was prepared based on the official User Guide of iPhone VoiceOver (Apple) and Microsoft's Seeing AI app (Microsoft). None of the participants had prior experience with an iPhone and the Seeing AI. User experience was measured with a System Usability Scale (SUS) – a 10-item questionnaire with a five-point Likert scale, ranging from 1 (*strongly disagree*) to 5 (*strongly agree*).

*Procedure*

A one-to-one individual tutorial was offered to each participant in which they learned how to use VoiceOver and the Seeing AI app. Afterward, participants completed a System Usability Scale (SUS) questionnaire to measure user experience with VoiceOver and the Seeing AI app. Participants' responses to the SUS questionnaire were analyzed using descriptive statistics. The video recordings of user interactions were analyzed using open, axial, and selective coding.

## Results

*SUS of VoiceOver*

The mean score of the participants' summed responses was 52.5±16.11, which is considered a poor user experience (Bangor et al.). The responses were broken down into positive and negative items (See Figures 1 and 2). The mean score of the five positive items was 3.52±0.52. The participants appreciated it that VoiceOver was equipped with various functions, and they would like to use it frequently. However, their confidence was low.

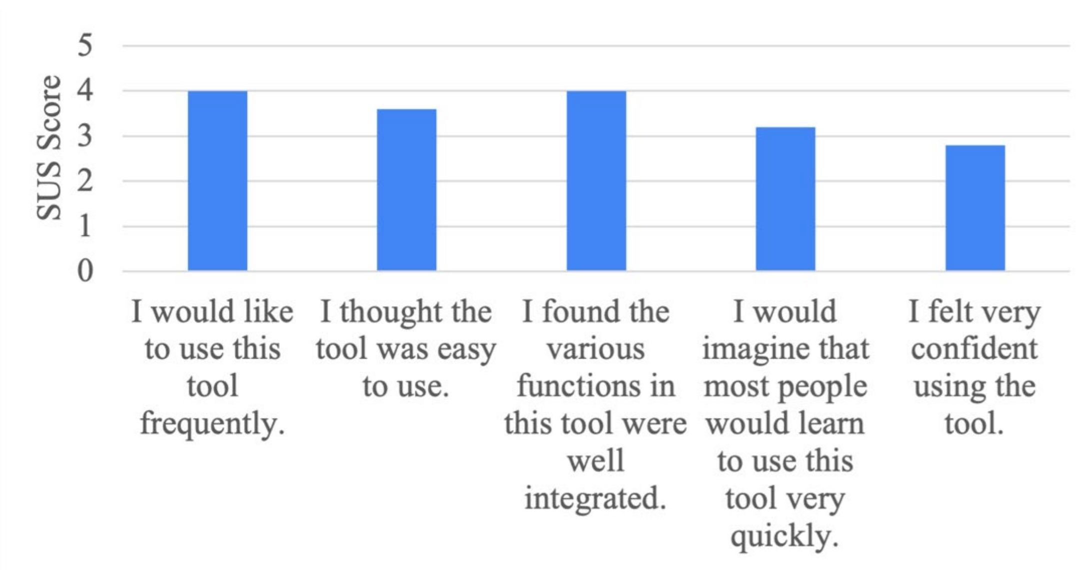


Fig. 1. Mean scores of positive items for VoiceOver.

The mean score of the five negative items was 3.32±1.06. The participants perceived that user interfaces of VoiceOver were well designed to ensure consistency but there were many things for them to learn to operate it properly, such that they would likely rely on a technical person.

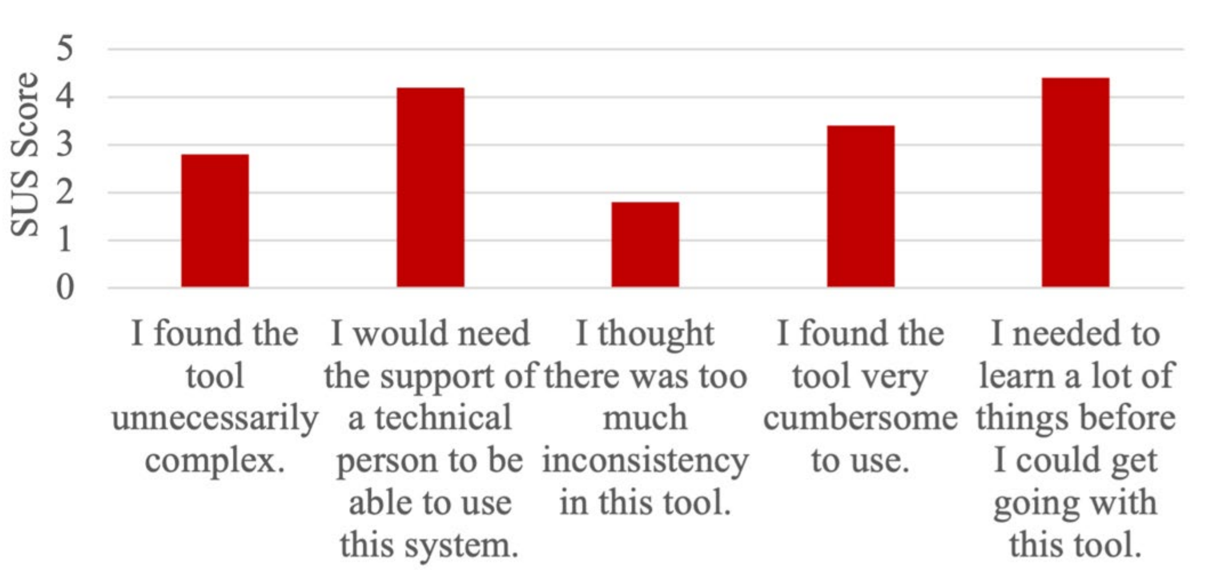


Fig. 2. Mean scores of negative items for VoiceOver.

*SUS of the Seeing AI App*

The mean score of the participants' summed responses was 74.50±12.17, which is considered a good user experience (Bangor et al.). The responses were broken down into positive and negative items (See Figures 3 and 4). The mean score of the five positive items was 4.24±0.46. The mean score of the five negative items was 2.28±1.18. The participants were satisfied with the Seeing AI app overall.

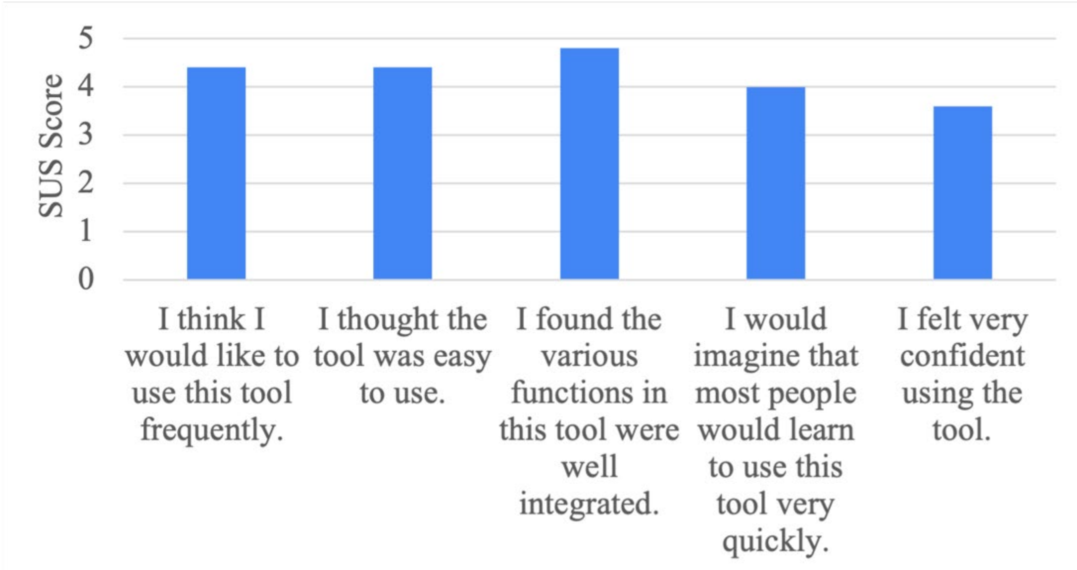


Fig. 3. Mean scores of positive items for Seeing AI.

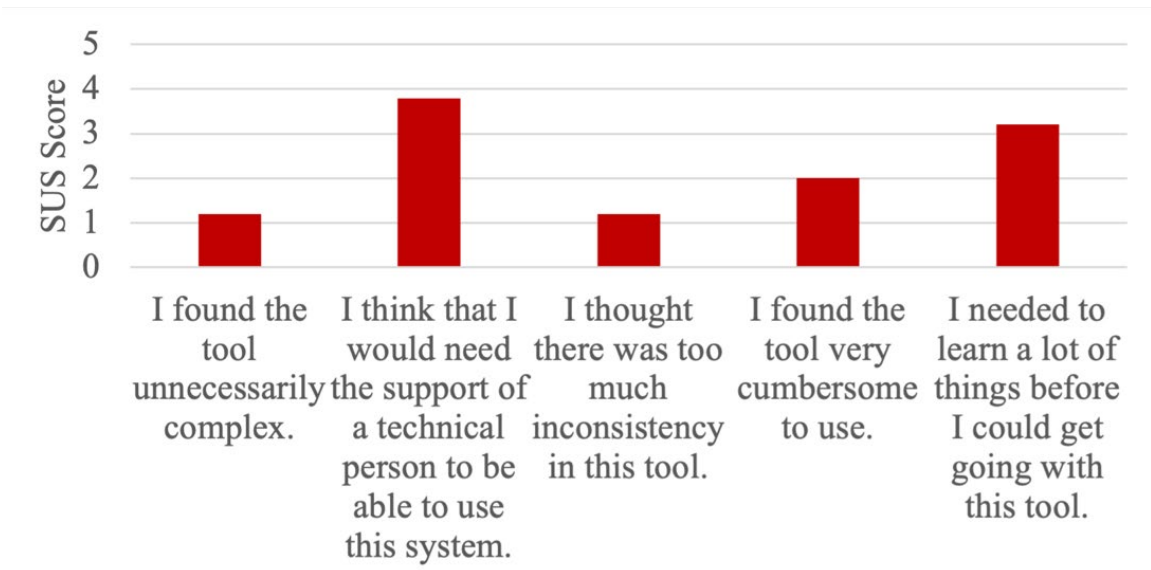


Fig. 4. Mean scores of negative items for Seeing AI.

*User Interactions with VoiceOver*

The participants were observed to have poor user experience, the cases of which were summarized under three categories: gesture commands, voice commands, and user interfaces (UI) (see Tables 2, 3, and 4). Tables 2 - 4 include adequate recommendations to address each case of poor user experience.

Table 2. Gesture-related incidents and recommendations.

| **Incidents** | **Recommendations** |
| --- | --- |
| Scroll down one page: Wrong direction | Users should be given the option to switch the scrolling direction based on their preference. |
| Scroll down one page: The wrong number of fingers | Users should be given the option to reset the gesture command (e.g., one, two, or three fingers). |
| Scroll down one page: Fingers apart gradually | The iOS should distinguish the finger gestures for "zoom in" and "scroll down." Users should be given the option to reset a minimum distance threshold between fingers. |
| Select, speak an item: Touch but not hold long enough | Once users touch an item, a status message should be provided, e.g., "*You tapped one of four buttons. There are three more*." |
| Select the next item: Swipe very slowly | Users should be given the option to change the moving speed for the swiping gesture. |
| Select an item: Lack of mental models of spatial layouts | Users should be given the option to easily change spatial layouts (e.g., a calendar view for day, week, month, or year). |
| Quick actions menu: Press-and-hold for too short or long | Users should be given audio feedback. If they press and hold long enough to see the Quick Action menu, a beep sound is generated. |
| Double tap: Taping on wrong targets | Users should be given voice advice (e.g., "*Double tap on the TIME and DATE texts to schedule an event*"). |

Table 3. UI-related incidents and recommendations.

| **Incidents** | **Recommendations** |
|---|---|
| Affordance: Against Fitts' Law | Fitts' Law should be applied to improve the user interfaces, e.g., a small *plus* sign is currently located at the top right of the screen, which is too far from the place where users enter a new schedule in the calendar. The *plus* sign should be relocated, or there should be an alternative way for users to enter a new schedule more easily. |
| Affordance: Irremovable notifications | Users should be given an alternative option (e.g., a physical button) to close the Notification Center. |
| Affordance: Inconsistent direction for a slider | User interfaces should be designed to be consistent with natural hand gestures, e.g., if a slider is designed to move horizontally (left and right), a gesture must be designed to move horizontally instead of vertically (up and down). |
| Graphic UI: Images/texts vs. clickable buttons | Images/texts and clickable buttons should be easily differed. For example, the VoiceOver should inform users whether interface components (e.g., images, texts, and buttons) are clickable. |
| Graphic UI: Low color contrast | A strong color contrast should be employed to distinguish between an active button and a disabled button. |
| Graphic UI: Identical buttons | Different user interface designs should be assigned to buttons for different functions. |

Table 4. Voice-related incidents and recommendations

| **Incidents** | **Recommendations** |
|---|---|
| Message: Unclear message | In case Siri does not understand users' voice commands, users should immediately be allowed to repeat it without saying "Hey, Siri" again. |
| Wake phrase "*Hey, Siri*": Overwhelmed with the wake phrase by repetition | Users should be given the option to keep having a conversation with Siri for a certain period without repeatedly using the wake phrase. |
| Voice commands: Unlisted commands | Siri should be able to guess voice commands that are not programmed yet (or users can easily add new commands). |

*User Interactions with the Seeing AI App*

As compared to VoiceOver, the Seeing AI app resulted in a fewer number of poor user experience cases (see Table 5).

Table 5. The app-related incidents and recommendations.

| Incidents | Recommendations |
|---|---|
| Camera covered by fingers | The app should be designed to alert users when the camera is blocked (e.g., beep sound). |
| Camera not aiming at the entire paragraph | Users should receive audible notifications when the entire paragraph is not captured. |

## Discussion

The participants had a better user experience with the Seeing AI app (mean of overall SUS scores > 74) compared to VoiceOver (< 53). The SUS score of 70 is considered a cut-off point for determining good user experience (Bangor et al.). Given the norm, it can be interpreted that the participants in this study encountered many usability and accessibility problems while using VoiceOver. The observation in this study also confirmed qualitatively that VoiceOver caused more cases of poor user experience. Hence, this study suggested alternative designs to address the poor user experience (see Tables 2, 3, and 4).

The research findings also infer that the Apple's User Guide has limitations. It does not offer alternative formats for users with visual disabilities as it simply consists of texts and images. Furthermore, it does not convey detailed instructions about "dynamic" finger gestures. For example, the participants were not able to learn "how long" they were supposed to press and hold their finger on an app icon in order to have a Quick Action menu (i.e., a pop-up menu with shortcuts for additional actions). Due to such an incomplete user guide, the participants ended up with holding for too long, such that an "X" mark was shown next to the app icon, ready to be deleted. If they accidentally touched the "X" mark, the app would be deleted against their will.

The primary responsibility of a user guide (or manual) is to help users learn how to use a new application; however, a user guide becomes useless if a user cannot understand the instructions as they are poorly written. Regardless of the product quality, users are likely to be unsatisfied due to the lack of understanding about the product. It is well documented that there is a significant relationship between the quality of the user guide and the perceived product quality (Gök et al.). Allwood and Kalén, for example, observed that a user-friendly user guide could help users to spend less time on the tasks and make fewer errors. Byrne tested *Iconic Linkage* (i.e., the use of the same words to present the same information multiple times in a text) to improve the usability of a user guidebook. He revealed that Iconic Linkage contributed to a shorter time to complete a task, improved retention of information, fewer mistakes, and higher user satisfaction with the product. Besides the quality of instructions per se, the medium to deliver the instructions could be another critical factor for good usability (Gök et al.). Alexander compared the effectiveness of print- versus video-based user guides and found that the video format was more likely to result in positive outcomes as users who were given the video format made fewer errors and completed tasks with more accuracy. For users with visual disabilities, alternative formats are recommended such as tactile, haptic, or audio formats for user guides.

The participants also showed a lack of understanding of how voice commands worked. With Siri, users can use many accessibility features, make and receive phone calls, hear notifications, and so forth. The participants seemed to treat Siri as a human-like voice assistant. Hence, they kept talking to Siri without repeating the wake phrase “Hey Siri”; had a long conversation instead of a short voice command; and used voice commands that were not programmed in the Siri system (e.g., *“Hey, Siri. Go to 100.9*”, which was a radio channel number). Ghosh et al. also administered the SUS survey with sighted users of Siri, and the mean of overall SUS scores was 54.17, which is greater than that of the participants with visual disabilities in this study. A voice assistant such as Siri is critical for users with visual disabilities to fully use various features of smart technologies, such that more user-friendly voice user interfaces should be provided (e.g., more flexible and natural voice commands).

The participants showed a higher level of satisfaction with the Seeing AI app. As an assistive technology app, the Seeing AI app is also executed based on voice user interfaces as does VoiceOver. Yet, the Seeing AI app has more intuitive user interfaces, i.e., users simply use a built-in camera to capture what they want to identify, and it reads out loud for users, leading to a higher level of user satisfaction. It infers that VoiceOver should be redesigned to be equipped with better user-friendly interfaces and interactions for those with visual disabilities.

## Conclusion

Today, many mainstream technologies are accessible to people with visual disabilities via assistive technologies such as Siri, VoiceOver, and Seeing AI app. The assistive technologies can help users with visual disabilities to control a system, obtain information, and communicate with others without barriers, leading to independence in everyday life. However, little is known about the user experience of those applications (Siri, VoiceOver, and Seeing AI). This study conducted in-person observations and found that participants with visual disabilities encountered many poor user experience cases. To address them, this study suggested a set of design recommendations.

## Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 1831969.

**Works Cited**

Alexander, Kara Poe. “The usability of print and online video instructions.” *Technical Communication Quarterly* 22.3 (2013): 237-59. Print.

Allwood, Carl Martin, and Tomas Kalén. “Evaluating and improving the usability of a user manual.” *Behaviour & Information Technology* 16.1 (1997): 43-57. Print.

Apple. “iPhone user guide.” 2023. Web. February 03 2023.

Bangor, Aaron, Philip Kortum, and James Miller. “Determining what individual sus scores mean: Adding an adjective rating scale.” *Journal of usability studies* 4.3 (2009): 114-23. Print.

Byrne, Jody. “Evaluating the effect of iconic linkage on the usability of software user guides.” *Journal of technical writing and communication* 35.2 (2005): 155-78. Print.

Celusnak, Brian M. “Teaching the iPhone with voiceover accessibility to people with visual impairments.” *Journal of Visual Impairment & Blindness* 110.5 (2016): 369-72. Print.

Centers for Disease Control and Prevention. “Fast facts of common eye disorders.” 2022. Web. January 27 2022.

Crossland, Michael D, Rui S. Silva, and Macedo, Antonio F. “Smartphone, tablet computer and e-reader use by people with vision impairment.” *Ophthalmic and Physiological Optics* 34.5 (2014): 552-57. Print.

Dockery, Dominique, and Magdalena Krzystolik. “The use of mobile applications as low-vision aids: A pilot study.” *Rhode Island Medical Journal* 103.8 (2020): 69-72. Print.

Ghosh, Debjyoti, et. al. “Assessing the utility of the system usability scale for evaluating voice-based user interfaces.” *Proceedings of the Sixth International Symposium of Chinese CHI*. 2018. Print.

Gök, Osman, Pervin Ersoy, and Gülmüş Börühan. “The effect of user manual quality on customer satisfaction: The mediating effect of perceived product quality.” *Journal of Product & Brand Management* (2019). Print.

Grussenmeyer, William, and Eelke Folmer. “Accessible touchscreen technology for people with visual impairments: A survey.” *ACM Transactions on Accessible Computing (TACCESS)* 9.2 (2017): 1-31. Print.

Kim, Hyung Nam. “User experience of mainstream and assistive technologies for people with visual impairments.” *Technology and Disability* 30.3 (2018): 127-33. Print.

Leporini, B., Buzzi, M. C., & Buzzi, M. “Interacting with mobile devices via Voiceover: Usability and accessibility issues”. *Proceedings of the 24th Australian Computer-Human Interaction Conference*. 2012. ACM. Print.

Microsoft. “Seeing AI in New Languages.” 2023. Web. February 03 2023.

Park, Kyudong, Taedong Goh, and Hyo-Jeong So. “Toward accessible mobile application design developing mobile application accessibility guidelines for people with visual impairment.” *Proceedings of Hci Korea*. 2014. 31-38. Print.

Smaradottir, B. F., Martinez, S. G. and Håland, J. A. “Evaluation of touchscreen assistive technology for visually disabled users.” *IEEE Symposium on Computers and Communications (ISCC)*. 2017. IEEE. Print.

Varma, Rohit, et al. “Visual impairment and blindness in adults in the United States: Demographic and geographic variations from 2015 to 2050.” *JAMA ophthalmology* 134.7 (2016): 802-09. Print.

WebAIM. “Survey of users with low vision #2 results.” 2018. Web. January 30 2023.

Wong, Meng Ee, and Stacey SK Tan. "Teaching the benefits of smart phone technology to blind consumers: Exploring the potential of the iPhone." *Journal of Visual Impairment & Blindness* 106.10 (2012): 646-50. Print.

World Health Organization. "Change the Definition of Blindness." 2008. Web. October 4 2022.