University of Glasgow | School of Computing Science

Honours Individual Project Dissertation

Accessible Solutions: A Low-Cost Assistive Web App for Deaf and Hard-of-Hearing Users Utilising Web Speech API

Amy Eden
October 14, 2023

Abstract

This dissertation explored the development and implementation of a cost-effective, assistive web application tailored for users who are Deaf or Hard-of-Hearing [1]. The project leveraged the capabilities of the Web Speech API to create an inclusive platform that addressed the unique needs of this user demographic. The research focused on the integration of speech-to-text, text-to-speech, and audio playback features to facilitate seamless communication and information accessibility. Through a user-centric design approach, the web-app aimed to provide practical solutions for the diverse challenges faced by individuals with hearing loss.

The study delved into the technical aspects of utilising the Web Speech API, evaluating its effectiveness in real-time scenarios, and considering the user experience and usability of the developed features. It was found that this web app was seen to be 17.4% more effective than current assistive technology. Additionally, all requirements were met, ensuring that the project fulfilled its objectives comprehensively.

By presenting a comprehensive and affordable solution, this dissertation contributed to the advancement of assistive technologies, fostering a more inclusive digital landscape for individuals within the Deaf and Hard-of-Hearing Community.

[1] Link to the repository: `https://github.com/deden3791/L4Project`

Education Use Consent

I hereby grant my permission for this project to be stored, distributed and shown to other University of Glasgow students and staff for educational purposes. **Please note that you are under no obligation to sign this declaration, but doing so would help future students.**

Signature: Amy Eden Date: 07 March 2024

Contents

1 Introduction

1.1 Motivation

More than 430 million individuals, comprising over 5% of the global population, need rehabilitation for their hearing loss. By 2050, it is projected that this number will surpass 700 million, affecting approximately one in every ten people worldwide. Currently, approximately 80% of individuals with hearing loss reside in countries classified as low- and middle-income (WHO 2023).

Numerous frameworks aim to advance the implementation of assistive technologies in low- and middle-income countries. For instance, the Global Cooperation on Assistive Technology highlights the significance of incentive product development through programs that encourage the creation of cost-effective assistive products (Tangcharoensathien et al. 2018). Similarly, the United Nations Convention on the Rights of Persons with Disabilities (CRPD) (Nations) provides an international framework dedicated to supporting implementation efforts and monitoring progress in assistive technology. Unfortunately, these crucial mandates have faced substantial neglect primarily due to inadequate financial backing, leading to minimal advancements in implementation (Borg et al. 2011).

Additionally, the cost-of-living crisis in the UK significantly affects individuals in poverty who cannot afford essential assistive technology. Driven by factors such as wage reductions and benefits cuts, this crisis exacerbates existing inequalities. For those in poverty, the inability to access vital assistive devices amplifies the impact on their daily lives, hindering the management of health conditions and essential communication. This heightened stress can impact mental health, emphasising the urgent need for targeted policies to address the specific challenges faced by this vulnerable population during economic upheaval (Broadbent et al. 2023).

Assistive technology serves as a life-changing tool in mitigating the impact of hearing loss and enhancing an individual's participation in daily activities. Among these technologies, hearing aids stand out as the most prevalent clinical intervention, experiencing an estimated compound annual growth rate (CAGR) of 6.4% and generating approximately $7.5 billion in revenue in 2021. Projections suggest a remarkable growth trajectory for the global hearing aids market, reaching a revenue size of $10.2 billion by 2026, predominantly driven by the continuous technological advancements in this field (Firm). From this, we can see that the hearing aid market is dominated by many companies, resulting in high prices due to restrictive contractual agreements between insurers and manufacturers. A lot of funding, like Medicare, the main health insurance provider for people aged 65 and older, does not cover the cost of purchasing or maintaining hearing aids Mahmoudi et al. (2018).

Hearing loss is the most prevalent sensory deficit in society (Sheffield and Smith 2019). It can lead to challenges in interpreting speech, resulting in diminished capacity for effective communication, delays in language acquisition, economic and educational disadvantages, social isolation, stigmatisation, and a compromised quality of life (Mathers et al. 2000; Lawrence et al. 2020; Davis et al. 2007). Hearing aids represent a front-line intervention for the majority of individuals experiencing hearing loss. This technology has demonstrated efficacy in alleviating many of the negative factors mentioned earlier, underscoring the essential role of such technology in addressing the needs of individuals with hearing loss (Borre et al. 2023; Jiménez-Arberas and Díez 2021).

Their impact spans various dimensions, including physical, social, emotional, and mental well-being, while significantly bolstering an individual's capacity to engage in effective communication with others. Combined, these factors lead to social isolation, withdrawal, depressive symptoms, and, again, a compromised quality of life (Davis et al. 2007). Despite the substantial advantages they offer, one in three individuals who could benefit from hearing aids do not have access to them. Furthermore, the rate of non-use can be as high as 24%, indicating a significant gap between the potential benefits of these devices and their actual utilisation (Ferguson et al. 2017).

Moreover, epidemiological evidence establishes an association between hearing loss and an escalated susceptibility to dementia. Failure to attend to hearing loss dismisses a critical opportunity for targeted intervention that holds the potential to serve as a cornerstone in the prevention of dementia (Powell et al. 2021; Lin et al. 2011; Livingston et al. 2020; Loughrey et al. 2018). Furthermore, recent studies have shown that hearing aids and mortality are linked (Choi et al. 2024). Specifically, an investigation based on a nationally representative sample of US adults revealed a significant association between audiometry-measured hearing loss and an increased risk of all-cause mortality. This study demonstrated a dose-response relationship, indicating that more severe levels of hearing loss are correlated with higher risks of mortality. Moreover, among individuals with hearing loss, the study found that regular hearing aid users exhibited a lower risk of mortality compared to those who never used hearing aids. These findings not only underscore the importance of addressing hearing healthcare disparities but also suggest a potential long-term benefit of regular hearing aid use in mitigating adverse health outcomes associated with hearing loss.

1.2 Project Aim

Building upon the motivation outlined in Section 1.1, the primary objective of this dissertation project is to develop an affordable, open-source web application specifically designed to assist individuals who are Deaf or Hard-of-Hearing. Leveraging the React framework, the project aims to provide a versatile platform where users can access and interact with various assistive features on various devices, essentially functioning as a multi-modal hearing aid.

From this we outline the following research questions:

RQ.1 What are the socio-economic impacts of assistive technologies for the Deaf and Hard-of-Hearing community and their effect on access/adoption rates?

RQ.2 How effective is inclusive design in addressing diverse needs among Deaf and Hard-of-Hearing users?

RQ.3 What strategies can be used to prioritise user needs, preferences, and feedback in developing assistive technologies?

RQ.4 How do usability features impact the accessibility/usability of assistive technologies for varying hearing loss levels?

RQ.5 To what extent does community engagement contribute to the success of assistive technology platforms, and what factors influence participation?

RQ.6 How do different devices affect the user experience/accessibility of assistive technology platforms, and what design considerations are necessary?

We use these research questions to illustrate in the following aims:

PA.1 **Affordable**: The application's primary aim is to offer its services entirely free of charge, lessening any financial barriers that could limit access to essential assistive technologies for the Deaf and Hard-of-Hearing community.

PA.2 **Inclusive**: By design, the application aims to cater to individuals across the spectrum of hearing loss, acknowledging and addressing the diverse needs and variations among users. Its features are tailored to be adaptable and accommodating for all levels of hearing loss.

PA.3 **User-Centric**: Throughout the development and evolution of the application, the primary focus is to prioritise user needs, preferences, and feedback. Regular updates and feature enhancements stem from user engagement studies, ensuring that the platform evolves in direct response to user input.

PA.4 **Usable**: The application is engineered with usability at its core, ensuring users of all abilities can utilise the web-app. This includes adherence to universal design principles, variability for text features, being able to access the web-app on various devices, and integrating intuitive interfaces to accommodate a wide range of users.

PA.5 **Collaborative**: The platform fosters a sense of community and engagement, encouraging the Deaf and Hard-of-Hearing community to actively participate in its upkeep, suggest improvements, report issues, and collectively contribute to its maintenance and growth.

PA.6 **Versatile**: Recognising the diversity in device usage, the application is designed to be versatile, ensuring seamless functionality across various devices and operating systems, including smartphones, tablets, laptops, Raspberry pi, and desktops. This ensures users can access the platform conveniently regardless of their preferred device.

These aims collectively underscore the commitment to improve access to assistive technologies, embracing inclusivity, and prioritising user satisfaction and community engagement throughout the development and evolution of the multi-modal hearing aid application. However, it's crucial to evaluate the necessity of this project and examine the sociological implications of assistive technologies, like this project, on potentially segregating the Deaf community.

1.3 Summary

This chapter acts as the catalyst for the project and delineates its objectives. Subsequently, the dissertation is structured as follows:

Chapter 2 initiates a comprehensive exploration of the research pertinent to the present project, commencing with an in-depth examination of the historical context surrounding Deafness and its intersection with assistive technology. Subsequently, the discussion delves into an analysis of contemporary assistive technologies.

Chapter 3 explores use cases and requirements of the web-app, outlining various scenarios for diverse users needs. It highlights the inclusive design approach, detailing how the app accommodates users with various levels of hearing loss, those seeking customisation, multilingual users, collaborative users, users across different devices, and new users. Additionally, the chapter presents functional and non-functional requirements categorised based on priority using the MoSCoW method, emphasising critical features like speech-to-text, text-to-speech, intuitive user interface, and device compatibility, among others.

Chapter 4 delves into the conceptualisation and strategical considerations that lay the foundation for the design phases. Without delving into specific coding intricacies, we present an overview of the architectural decisions, user experience considerations, and the general design philosophy that guides the development of our solution. This chapter serves as a roadmap for the ensuing sections, where we will elaborate on the specific design components and their functionalities.

Chapter 5 delves into the tangible realisation of the proposed solution. Building upon Sections 1.1 and 4 discussed earlier, the implementation phase aimed to bring the envisioned assistive web application for the Deaf and Hard-of-Hearing community to life.

Chapter 6 provides an in-depth examination of two surveys designed to collect data regarding the affordability and usability of existing assistive technologies.

Chapter 7 describes an evaluation which assesses the usability of the web application designed for users who are Deaf or Hard-of-Hearing. After each task, participants answered questions to provide feedback on their experience with the application based on the requirements outlined in Sections 4.3 and 4.4. The evaluation identified strengths and areas for improvement in the application's design and functionality. Data from participant responses will be analyzed to guide further enhancements to the web app.

Chapter 8 provides a concise summary of the key findings and insights obtained throughout the project. It will reiterate the significance of the research objectives and discuss how the findings contribute to addressing the identified problem and research aims. Additionally, this chapter reflects on the implications of the project's results, offers recommendations for future research or practical application, and emphasises the importance of the work in the broader context of the field. Finally, we will conclude with a closing statement that reinforces the project's contribution and significance.

2 Background

Understanding and developing assistive technology for the Deaf and Hard-of-Hearing community necessitates a comprehensive grasp of historical perspectives on hearing loss, its evolving sociological implications, and the nuanced medical delineation of hearing loss. Acknowledging the historical context of hearing loss provides crucial insights into the origins, societal responses, and early attempts at addressing Deaf and Hard-of-Hearing individuals. Furthermore, exploring the sociological impact of hearing loss across different epochs highlights the varying societal attitudes, stigmas, and accommodations afforded to the Deaf and Hard of Hearing community.

Burrows outlines two fundamental conceptual frameworks that underpin our understanding of Deafness and hence our understanding of assistive technology. The cultural perspective delineates deafness as an identity distinct from disability, emphasising sign language as a pivotal linguistic conduit. In contrast, the infirmity or medicalized model perceives Deafness as a modifiable impairment through advanced technological interventions. These divergent perspectives significantly shape the trajectory of assistive technology for the Deaf and Hard-of-Hearing community. The infirmity model steers the development of technologies aimed at rectifying the Deaf body, while the cultural model fosters the creation of technologies that facilitate the harmonious coexistence of Deaf and hearing experiences (Burrows 2022).

Prior to delving into the diverse range of existing technologies, it is imperative to establish a clear definition of hearing loss. Hearing loss is commonly diagnosed as an individual experiencing a diminished capacity to perceive sound, characterised by hearing thresholds exceeding 20 decibels (dB) or better in both ears (WHO 2023; Britannica 2024). However, the concept of Deafness encompasses a socially constructed perspective regarding its significance, shaped by the perceptions of both Deaf individuals and those who are hearing (Burrows 2022).

2.1 Historical Context

The earliest evidence of hearing loss, dating back over 10,000 years in the Shanidar Caves, reveals skeletal remains with ear canal bony growths causing conductive hearing difficulties. The Ebers Papyrus from 1550 BC provides the first documented treatment for "Auditory Dysfunction," involving a mixture of olive oil, red lead, ant larvae, bat appendages, and goat urine. Moving into the 10th century, Plato and Aristotle linked hearing loss to intellectual deficiency, marking the beginning of societal alienation for the Deaf and Hard-of-Hearing community (CalHearing 2023).

California Hearing (CalHearing 2023) further details the origins of sign language's, tracing back to the 10th century monks of ancient Burgundy, not as an aid for the Deaf but as a means of communication among vow-of-silence individuals. In the 13th century, people with hearing loss used hollowed-out animal horns as rudimentary hearing aids. The 18th-century ear trumpet, despite its bulkiness, gained popularity, followed by other devices like stethoscopes and speaking tubes (Valentinuzzi 2020).

Alexander Graham Bell's telephone in 1876 paved the way for Miller Reese Hutchison's amplified electronic hearing aids in 1889 (CalHearing 2023; Valentinuzzi 2020). Vacuum tube technology

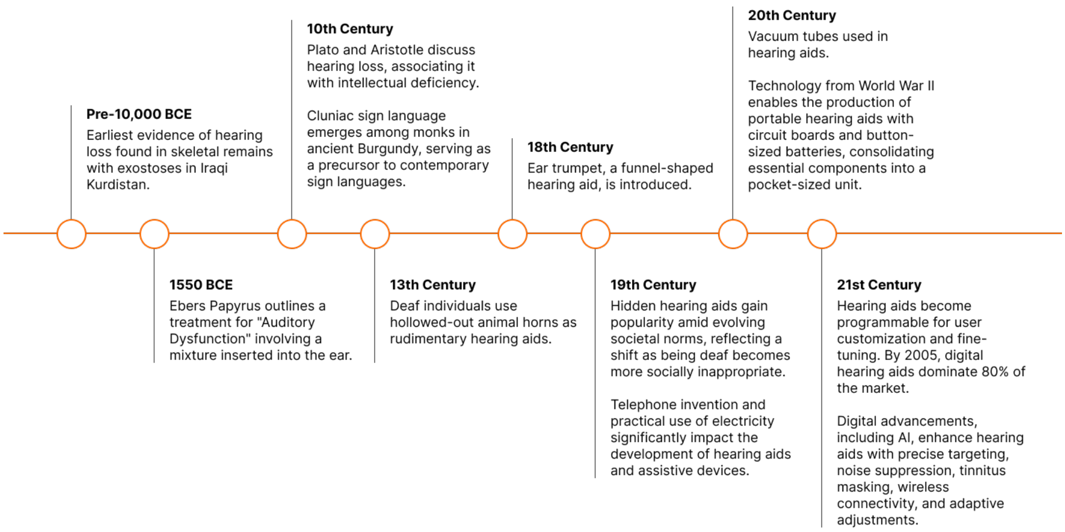


Figure 2.1: *Evolution of Hearing Aid Technology: A historical journey from rudimentary devices in the 13th century to 21st-century digital advancements, including AI integration, transforming the landscape of assistive technology for the Deaf and Hard of Hearing community.*

in the 1920s improved portability until the mid-1940s, replaced by transistors developed during World War II. Microprocessors in the '70s and '80s led to lighter, more potent analog hearing aids. In the 1990s, digital technology marked a milestone, miniaturizing and enhancing the potency of hearing aids. Present-day digital advancements include precise targeting of individualized hearing loss, noise suppression, tinnitus masking, wireless connections, and AI adaptation to various environments (CalHearing 2023). Figure 2.1 displays the evolution of assistive technology throughout history.

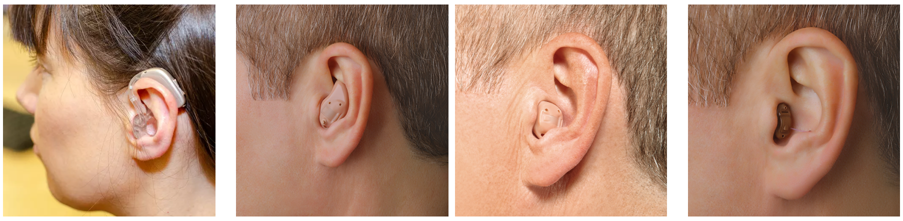
Notably, the prevalence of hereditary Deafness on Martha's Vineyard during the 19th century was strikingly high, with an estimated rate far exceeding the national average. Sign language became essential for daily communication among both Deaf and hearing islanders, leading to a fully integrated community where communication barriers were rare. Despite the absence of a separate "Deaf" society, the seamless integration of Deaf individuals into everyday life showcased the island's inclusive and cohesive cultural ethos (Groce 2016). This displays what Burrow's defined as a culteral persepctive onto Deafness, which we must take into consideration when developing assistive technology (Burrows 2022).

2.2 Existing Technology

2.2.1 Hearing Aids

Since the 2000s, hearing aids have been programmable, allowing user customisation and flexibility. By 2005, 80% of the market was dominated by digital devices, which continued to shrink in size due to advancements in transistor fabrication using silicon. These modern hearing aids seamlessly adapt to different environments and connect with electronic devices like computers and telephones, enhancing accessibility in public spaces (Valentinuzzi 2020).

A hearing aid consists of three components: a microphone, amplifier, and speaker. It captures sound, converts it into electrical signals, amplifies them, and relays them to the ear (NIDCD 2022). Designed to alleviate the impacts of hearing loss, hearing aids focus on enhancing speech sounds to improve users' capacity to engage effectively in various situations (Ferguson et al. 2017).



(a) Behind the ear (BTE) hearing aid *(b) In the ear (ITE) hearing aid* *(c) In the canal (ITC) hearing aid* *(d) Completely in canal (CIC) hearing aid*

Figure 2.2: *Comparison of Hearing Aid Styles: From Behind-the-Ear (BTE) to In-the-Canal (ITC), illustrating the diverse options for users with varying preferences and hearing needs.*

The National Institute on Deafness and Other Communication Disorders categorise three primary styles of hearing aids: Behind-the-ear (BTE), In-the-ear (ITE), and Canal aids, which inclue in-the-canal (ITC) and completely-in-canal (CIC). Each addresses different degrees of hearing loss and preferences, offering versatility to users and are displayed in Figure 2.2.

Hearing aids utilise analog or digital systems. Analog aids amplify sound waves, with adjustable and programmable options. Digital aids, using numerical codes, provide precise programming, adaptability to frequencies, and advanced features like directional focus. Digital technology, applicable across all types of hearing aids, offers significant customization options for users (NIDCD 2022).

As previously mentioned in Section 1.1, the hearing aids market is anticipated to continue its upward trajectory (Firm). This highlights the significance of hearing aids as a vital clinical intervention. However, exploring the landscape of auditory healthcare prompts a critical examination of the accessibility and affordability of these essential devices, particularly considering hearing aid prices detailed in the following tables. This raises important questions about their economic feasibility for a broader demographic.

Multimodal hearing aids represent a transformative advancement in auditory assistive technologies, offering a crucial solution for individuals facing diverse hearing challenges. These next-generation devices seamlessly integrate auditory cues with complementary visual or tactile stimuli, transcending traditional auditory augmentation. They play a pivotal role in restoring intelligibility and reducing cognitive load in noisy environments (Shah et al. 2022). However, their real-time implementation demands a delicate balance between high data rates, low latency, low computational complexity, and robust security measures (Adeel et al. 2020). By strategically combining sensory modalities, multimodal hearing aids aim to provide a more nuanced understanding of the acoustic environment, catering to users with varying degrees of hearing loss.

Table 2.1: *Hearing Aid Prices in 2024*

Hearing Aid Style	Hearing Aid Price Each (£)	Hearing Aid Price Pair (£)
BTE hearing aid	595	895
RIC hearing aid	595	895
ITE hearing aid	695	1195
CIC hearing aid	696	1195
ITC hearing aid	695	1195
IIC hearing aid	695	1195
BICROS hearing aid	795	NA
CROS hearing aid	795	NA

Table 2.2: Hearing Aid Price Range in 2024

Hearing Aid Budget	Hearing Aid Price Range Singular (£)	Hearing Aid Price Range Pair (£)
Low Budget	345-695	345-695
Basic Budget	695-895	1295-1395
Mid-level Budget	1095-1195	1595-2095
Advanced Budget	1245-1395	2295-2895
Premium Budget	1595-1695	2895-3190

In the context of hearing aid affordability, Tables 2.1 and 2.2 provide an overview of hearing aid prices in 2024. Table 2.1 details the individual and pair prices for various hearing aid styles. Table 2.2 further categorises hearing aid budgets, outlining the price ranges for singular and pair purchases across different budget tiers: Low Budget, Basic Budget, Mid-level Budget, Advanced Budget, and Premium Budget. This categorisation provides insights into the affordability spectrum within the hearing aid market (Harrison 2024).

There are many other forms of hearing aids, including implantable hearing aids, such as the middle ear implant (MEI) and bone-anchored hearing aid (BAHA), aim to enhance sound transmission directly into the inner ear. The MEI, attached to a middle ear bone, moves bones to reinforce sound vibrations, while the BAHA, affixed to the bone behind the ear, transmits sound vibrations through the skull, bypassing the middle ear. BAHAs are often used for middle ear issues or unilateral Deafness. However, the need for surgical implantation raises concerns among specialists who assess potential benefits against associated risks (NIDCD 2022).

2.2.2 Speech Recognition

The history of speech recognition, and automated speech recognition (ASR), traces its roots back to the 1920s, where "Radio Rex," a celluloid dog animated by a spring release triggered by the vocal cue "Rex", which is 500 Hz acoustic energy. This toy demonstrated the first innovations in speech recognition technology (David and Selfridge 1962). From this, ASR has evolved into a sophisticated Natural Voice-Voice Interface (VUI). Present-day applications span a spectrum of utilities, including communication with smart home appliances, personal assistants, and cellphones, particularly catering to the elderly (Song et al. 2022). Furthermore, ASR plays a pivotal role in general transcription, generating automatic captions for audio or video text, and significantly contributes to augmentative communication for individuals with disabilities (Semary et al. 2024).

Jurafsky, (Jurafsky and Martin), provides an insightful exploration of ASR dimensions, including vocabulary size, speaker interactions, channel and noise conditions, and speaker-class characteristics. Complexities escalate with open-ended tasks like transcribing videos or human conversations, variations in channel and noise, and the nuanced impact of speaker characteristics such as regional accents. Furthermore, Jurafsky outlines the encoder-decoder architecture of speech recognition, beginning with the user speaking and producing an initial audio signal. This signal is then captured, undergoing pre-processing for noise reduction and enhancement. Relevant features are extracted using techniques like MFCCs, transformed into feature vectors for analysis. Acoustic modeling, often with HMMs or neural networks, establishes the relationship between features and phonetic units. Language modeling adds context, refining word sequence likelihood. Decoding combines acoustic and language models, utilizing algorithms like Viterbi decoding. Post-processing refines text using statistical language models and grammatical constraints, ensuring accuracy. The final output is transcribed text or commands, representing the user's spoken input. Figure 2.3 displays a simplified version of the ASR architecture.

Table 2.3 presents a comprehensive overview of Word Error Rates (WER%) and Character Error Rates (CER%) across diverse speech recognition tasks in American English and Mandarin Chinese. Reflecting the accuracy of ASR systems, these rates fluctuate across tasks, with lower

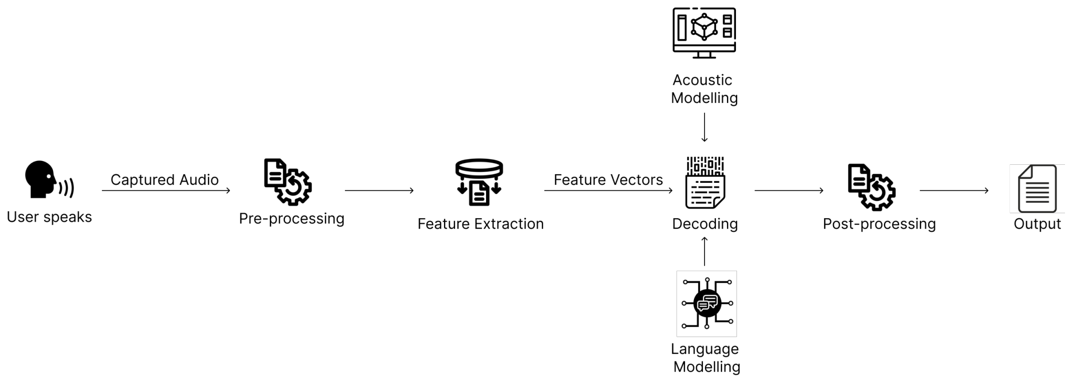


Figure 2.3: *Flow Diagram of Speech Recognition Process: From User Speech Input to Accurate Transcription*

percentages indicative of heightened accuracy and efficiency in transcription (Jurafsky and Martin). This consolidated analysis underscores the evolving landscape of ASR and its crucial role in contemporary technological applications.

Table 2.3: *Word Error Rates (WER%) and Character Error Rates (CER%) across diverse speech recognition tasks in American English and Mandarin Chinese*

English Tasks	**WER%**
LibriSpeech audiobooks 960hour clean	1.4
LibriSpeech audiobooks 960hour other	2.6
Switchboard telephone conversations between strangers	5.8
CALLHOME telephone conversations between family	11.0
Sociolinguistic interviews, CORAAL (AAL)	27.0
CHiMe5 dinner parties with body-worn microphones	47.9
CHiMe5 dinner parties with distant microphones	81.3
Chinese (Mandarin) Tasks	**CER%**
AISHELL-1 Mandarin read speech corpus	6.7
HKUST Mandarin Chinese telephone conversations	23.5

Advancements in speech recognition technology are empowering individuals, particularly those with disabilities, to seamlessly control devices, access information, execute tasks, and interact with digital platforms. This progress not only fosters independence but also enhances overall quality of life by providing newfound convenience and accessibility. For people with disabilities, such as those with mobility or visual disabilities, speech recognition serves as a vital tool, enabling them to navigate digital interfaces and perform tasks that may have been challenging or impossible otherwise. By offering voice-controlled alternatives to traditional input methods, speech recognition technology enables individuals with disabilities to participate more fully in various aspects of daily life, bridging the gap between accessibility and inclusivity in the digital realm (Semary et al. 2024). However, there are still some minority biases within speech recognition systems.

Many individuals become frustrated using technology such as Siri and even stop using them due to its Americanisation, which made it unable to recognise their speech. This illustrates the challenges faced by individuals with non-American names or accents. Many studies corroborates these concerns, revealing significant racial disparities in ASR technologies like Siri, with non-white speakers being disproportionately misunderstood (Koenecke et al. 2020). This bias not only erases cultural and linguistic identities but also poses challenges for individuals with disabilities who rely on voice recognition tools. Addressing these biases demands a multifaceted approach, including incorporating diverse training data, extensive product testing, and fostering diversity within tech companies' workforces. Moreover, tech companies need to recognise the influence of market-driven decisions on the prioritisation of inclusive technology development to mitigate these biases effectively (Lopez-Lloreda 2020).

2.2.3 Speech Synthesis

The historical narrative of speech synthesis technology traces back to Wolfgang von Kempelen's pioneering efforts in the 18th century, where he developed rudimentary methods using delicate bellows, springs, and resonance boxes to synthesize simple words. Despite these early attempts, the intelligibility of the synthesized speech remained poor. However, von Kempelen's innovations laid the groundwork for modern text-to-speech (TTS) synthesis, which involves mapping text to acoustic waveforms. Today, TTS technology plays crucial roles in conversational agents, aids for Blind individuals, and communication tools for individuals with neurological disorders (Jurafsky and Martin; Klatt 1980).

Jurafsky continues to explain the contemporary objective of speech synthesis, often referred to as text-to-speech or TTS, involves the inverse process compared to ASR. It aims to transform textual inputs into corresponding waveforms, a technological imperative with applications spanning dialogue systems, gaming interfaces, and educational tools. TTS systems, like their ASR counterparts, predominantly adopt the encoder-decoder architecture, leveraging either Long Short-Term Memory networks (LSTMs) or Transformers.

A distinctive characteristic in the training paradigm distinguishes TTS from ASR. While ASR systems prioritise speaker independence, necessitating training on diverse datasets with contributions from numerous speakers, TTS systems often embrace speaker dependence. For instance, the LJ Speech Corpus, a widely used dataset, comprises 24 hours of recordings from a singular speaker, emphasizing the feasibility of consistent voice generation with comparatively smaller datasets.

The TTS process unfolds in two pivotal components: an encoder-decoder model for spectrogram prediction, mapping text to mel spectrographs, and a vocoder, responsible for generating waveforms from mel spectrograms. In parallel, TTS systems undergo a crucial first pass of text normalization preprocessing to handle non-standard words, including numbers and abbreviations. This initial normalization step is vital given the verbalization differences compared to their spelled counterparts. Although modern end-to-end TTS systems exhibit the capacity to learn some normalization during training, the inherent limitation in training data size necessitates a distinct normalization step. This normalization can be achieved through rule-based approaches or an encoder-decoder model.

Crucially, the evaluation of speech synthesis systems remains a domain predominantly influenced by human listeners. This assessment involves playing synthesized sentences to listeners, who provide mean opinion scores (MOS) or participate in AB tests. The human-centric evaluation approach underscores the nuanced and subjective nature of assessing the quality of synthesized speech, prompting ongoing research in search of automated metrics to complement or potentially replace human evaluations.

In tandem with the advancements in speech recognition, speech synthesis has undergone rapid evolution propelled by new models, particularly those rooted in deep learning (Klatt 1980). These developments are progressively closing the gap, aiming to achieve speech synthesis that is virtually indistinguishable from human speech (Tan et al. 2024).

2.3 Other Technology

The landscape of assistive technology has been significantly shaped by the integration of accessibility features in contemporary smartphones, exemplified by prominent brands such as Apple, Android/Google Pixel, and Samsung Galaxy.

Apple has been a trailblazer in this domain, incorporating a robust set of accessibility features in its iOS devices, including VoiceOver for screen reading, magnification gestures, AssistiveTouch for motor challenges, and their beta feature for live captions (Support). Unfortunately, the majority of these advanced features are only available on iOS17, rendering them incompatible with iPhone models preceding the iPhone 8. This presents a significant barrier to access, effectively pricing out users who wish to utilise these capabilities, as they would be compelled to invest in a new phone that supports the latest iOS version.

Android, particularly on Google Pixel devices, has followed suit with its commitment to inclusivity, offering features like TalkBack, Live Transcribe, Sound Amplifier, and being able to connect hearing aids to your device to enhance usability for individuals with diverse needs (Google). Samsung Galaxy phones have also made strides in accessibility, featuring Voice Assistant, Screen Reader, Hearing Aid Compatibility, and sound amplification to cater to users with visual and hearing loss (Samsung 2023). The collective efforts of these technology giants underscore a commitment to creating inclusive digital ecosystems, ensuring that individuals of varying abilities can seamlessly integrate smartphones into their daily lives.

3 Use Cases & Requirements

3.1 User Cases

This section details varied hearing loss categories and how the web app accommodates them. Acknowledging complete Deafness limitations, it highlights alternatives like speech-to-text and text-to-speech for effective communication. Through concise user cases, it ensures a nuanced understanding of the app's inclusivity and usability across different hearing loss.

User with Partial Hearing Loss

- Utilises speech-to-text for improved communication
- Integrated text-to-speech feature with speech-to-text for clearer communication
- Accesses the hearing aid functionality for audio playback
- Engages with text-to-speech for communication if uncomfortable with speech

User with Profound Hearing Loss

- Utilises speech-to-text for captions
- Engages with text-to-speech for communication

User Seeking Customisation

- Adjusts hearing aid settings based on personal preferences
- Customises the user interface for a personalised experience
- Logging into user account to save individualised settings

User in Multilingual Settings

- Utilises language customisation in speech recognition and synthesis
- Engages with the application in their preferred language

Collaborative Users

- Engages in collaborative features, such as sharing messages
- Provides feedback and contributes to the improvement of the platform

Users on Various Devices

- Accesses the web app seamlessly across smartphones, tablets, laptops, smartwatches, Raspberry Pi, and desktops
- Ensures consistent functionality on different operating systems

New Users

- Navigates through an intuitive user interface designed for user-centric experience
- Benefits from user-friendly onboarding processes

3.2 Functional and Non-Functional Requirements

Throughout the project's evolution, initial requirements aimed at real-time sound amplification, alongside edge computing for additional functionalities like user information management, speech recognition, and synthesis. However, as our research deepened and time constraints became apparent, we made strategic decisions to streamline our focus.

We recognised the necessity to refine our objectives, prioritising core functionalities while leaving behind more ambitious features that could be taken forward in future projects. Consequently, we opted for a simplified approach, substituting real-time amplification with a more manageable frequency cutoff mechanism along-side gain and Q value adjustment, enabling users to manipulate sound waves effectively through their earbuds. Additionally, instead of delving into intricate edge computing processes, we simplified the architecture to incorporate essential elements through API calls.

By aligning our project goals with practical considerations, we ensure efficient resource utilisation and a more targeted development process. This strategic pivot allows us to deliver a robust solution within our constraints while maintaining a high standard of functionality and usability.

The requirements for the web-app were categorised by employing the MoSCoW method (Consortium), an Agile Business Consortium prioritization technique. This method involves labeling requirements based on their priority:

- **Must Have:** Critical requirements essential for creating a minimum viable solution.
- **Should Have:** Important requirements that are not critically essential.
- **Could Have:** Desirable requirements acceptable if not implemented.
- **Won't Have this Time:** Requirements out of the current project's scope but could be considered for future implementation.

Functional requirements are denoted by (**F**), while non-functional requirements are indicated by (**NF**).

Must have (MH)

MH.1 Speech-to-text (**F**): The web app must support speech-to-text functionality for generating captions.

MH.2 Text-to-speech (**F**): It must include text-to-speech features for effective communication.

MH.3 User interface (**F**): The app must have an intuitive and accessible user interface.

MH.4 Hearing aid (**F**): Audio playback in the web app is a critical feature for users with partial hearing loss.

Should Have (SH)

SH.1 Customisable Settings (**F**): Implement customisable variables to all audio outputs.

SH.2 User Feedback (**F**): Include an option for users to provide feedback on the app's usability, allowing continuous improvement.

SH.3 Device Compatibility (**NF**): Ensure compatibility with various devices, including smartphones, tablets, and desktops.

SH.4 User Profiles (**F**): The web app should offer user profiles to save personalised settings.

SH.5 Language Settings (**F**): Allow for multiple language options for speech recognition and synthesis.

Could Have (CH)

CH.1 Integration with Other Technologies (**F**): Explore integration with additional assistive technologies for a comprehensive user experience.

CH.2 Collaboration Features (**F**): Allow users to share experiences or tips on using the app, fostering a sense of community.

CH.3 Regular Updates (**F**): Continuous improvement through regular updates based on user feedback.

Won't Have this Time (WH)

WH.1 Proprietary Hardware Support (**F**): Integration with proprietary, expensive hardware for hearing aid support is out of scope.

WH.2 Complex Features (**F**): Features requiring extensive resources, such as speaker recognition and siren detection, are not prioritised.

4 Design

4.1 Architectural Design

In designing the architecture for the assistive web application, careful consideration was given to ensuring our project aims outlined in Section 1.2 are met. The following details the architectural components and decisions made during the design phase to meet the requirement MH.3 aforementioned in Section 3.2.

The client-side architecture, built on React TypeScript (TSX) [1], fosters a dynamic and responsive user interface. Users can interact directly with the interface or connect an audio device to unlock additional audio features. Leveraging React TSX's component-based design, development remains modular, enhancing both reusability and maintainability. On the server side, operations are driven by a Node.js server, delivering an event-driven and scalable backend. This server hosts the React TSX application and manages requests, facilitating seamless communication between client and server. Secure storage of user preferences in a database enables personalized experiences across the platform.

Communication between the client and server is achieved through RESTful APIs (Red-Hat 2020). This approach simplifies data exchange and ensures interoperability. Users experience a smooth interaction with the application, and future integrations are facilitated by the adherence to RESTful principles.

Security stands as a paramount concern within the architectural design, given the prevalent risks associated with potential eavesdropping on conversations through device microphones, owing to their system access permissions (Kröger and Raschke 2019). HTTPS secures communication, preventing unauthorized access to sensitive data. User preferences will be encrypted before storage, and robust authentication and authorization mechanisms are implemented to protect user accounts.

The deployment strategy involves running the React TSX application on a server. GitHub Pages encapsulates the application and its dependencies, providing a consistent environment. This ensures ease of deployment and scalability. Future scalability needs are addressed through horizontal scaling, leveraging load balancing techniques as displayed in Figure 4.1.

[1] `https://www.typescriptlang.org/docs/`

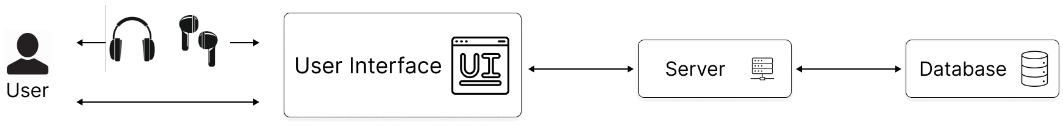


__Figure 4.1:__ Architectural Design Overview: A visual representation of the client-server architecture and communication protocols underlying the development of the assistive web application.

4.2 Affordability

As outlined in Section 1.2, the primary aim of the application is to provide its services completely free of charge, aligning with Project Aim PA.1, with only minimal hardware required. By offering essential assistive technologies for the Deaf and Hard-of-Hearing community without extortionate financial barriers, the platform ensures accessibility to all individuals, regardless of their economic status as mentioned in Section 1.1. However, as outlined in Section 2.2, contemporary technology, particularly hearing aids, often carries a heavy price tag, with the most affordable option starting at £345. Upon navigating through the Currys website [2], we identified significantly more budget-friendly alternatives that can access the outlined assistive web-app. The lowest-priced smartphone equipped with browser accessibility is priced at just £59 [3], while the least expensive earbuds are available for a mere £3.97 [4]. Remarkably, this combination totals £282.03 less than the cheapest hearing aids on the market.

4.3 User Interface (UI) Design

Although assistive technologies have many beneficial features, there are instances where users discontinue their usage. For hearing aids, research has indicated abandonment rates reaching up to 78% (Scherer 1996). Factors contributing to this rate of abandonment encompass issues associated with the device itself, including malfunctions or aesthetic considerations (Jiménez-Arberas and Díez 2021). Hence why this project aims to be inclusive and user-centric in it's UI design.

In crafting the User Interface (UI) for the assistive web application, paramount consideration was given to ensure user-friendliness and accessibility. The design principles not only prioritise ease of use but also cater to the unique needs of Deaf and Hard-of-Hearing users. Particularly being mindful of the fact that around 25% of individuals aged over 60 years grapple with hearing loss, underscoring the significance of this challenge among older adults (Haile et al. 2021).

In software development, UI design holds a crucial role, serving as a pivotal element in meeting user expectations and maximizing the software's functionality. A well-crafted UI not only enhances user acceptance of the software but also plays a vital role in educational contexts. For instance, an effective UI in educational software can significantly improve users' engagement with and absorption of knowledge (Chu and Chan 1998). Users rely on the feedback provided by the UI interface to shape their learning experiences, emphasizing the substantial impact of UI design on the overall user interaction with educational software.

In adherence to the foundational principles of User Interface Design Basics (Usability.gov 2014a), the wireframe design for the web application was meticulously crafted with a focus on key elements:

UI.1 **Simplicity in Interface:** The decision to employ a single-page web application was deliberate, aiming to ensure all features remain straightforward and user-friendly.

UI.2 **Consistency with Common UI Elements:** Consistent CSS styling for common interface elements was implemented, fostering a sense of comfort and familiarity for users across the web application.

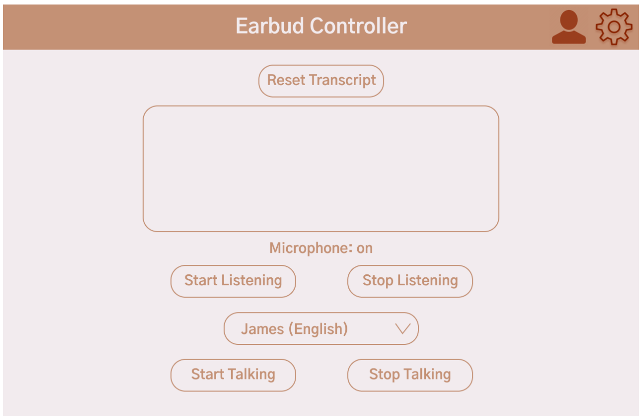
UI.3 **Purposeful Page Layout:** The centralization of features within the layout strategically directs attention to vital information, enhancing overall user experience.

UI.4 **Strategic Use of Color and Texture:** Thoughtful selection of colors, appealing to users and promoting a comfortable interaction with the web application, potentially contributing to user retention.

[2]Currys operates as a retailer specialising in technology products: `https://www.currys.co.uk/`

[3]`https://www.currys.co.uk/products/xiaomi-redmi-a2-32-gb-black-10250273.html`

[4]`https://www.currys.co.uk/products/skullcandy-jib-headphones-black-10178565.html`



(a) Wireframe design for the single-page web application, meticulously crafted for a user-centric and inclusive experience

(b) Thoughtfully chosen color palette for the web application, ensuring a visually appealing and user-friendly interface for diverse users.

Figure 4.2: *A comprehensive wireframe design for the single page web-app (a), coupled with a carefully curated color palette (b), harmonizing aesthetics and usability for an inclusive user experience.*

UI.5 **Typography for Hierarchy and Clarity:** Incorporation of varying font sizes for text boxes, especially considering potential users from the elderly demographic, enhances hierarchy and readability.

UI.6 **Clear Communication of System Status:** Implementation of error communication strategies and provision of a user guide contribute to effective user-system communication.

UI.7 **Consideration of Defaults:** The establishment of default settings caters to the needs of new users, streamlining their initial interaction with the web application.

Leveraging the capabilities of Figma, we crafted the wireframe that serves as the blueprint for the single-page web application as seen in Figure 4.2a. Adhering to the established principles of User Interface Design (Usability.gov 2014a), the wireframe design strategically centers features to guide user attention effectively. The color palette, depicted in Figure 4.2b, was meticulously curated with three warm tones – brown, off-white, and pink – to ensure a visually comfortable and user-friendly experience. Grounded in color-emotion associations and preferences, these warm hues align with positive sentiments of friendliness, comfort, reassurance, and liveliness (Liao and Hsu 2023). The selection of brown, off-white, and pink is further justified by their favorable rankings in Liao's research, reinforcing the intent to create an emotionally resonant and engaging web-app interface.

4.4 User Experience (UX) Design

UX design revolves around a deep understanding of user needs, values, and limitations, aligning with overarching business objectives. In crafting the UX for the assistive web application, the focus is on a user-centric approach. This ensures a seamless and inclusive interaction tailored to varying degrees of hearing loss. Key principles include accessibility, ease of use, and overall satisfaction, fostering a positive and engaging user experience throughout (Usability.gov 2014b).

In aligning with the foundational principles of the User Experience Honeycomb (Morville 2004), the web application's design is meticulously tailored, emphasizing key elements:

UX.1 **Useful:** The web app's content is meticulously curated to provide affordable and effective solutions for the diverse needs of the Deaf and Hard-of-Hearing community. Features such as speech-to-text, text-to-speech, and adaptive settings contribute to the practicality and usefulness of the application.

UX.2 **Usable:** Through the implementation of a single-page design, the web app prioritizes simplicity and intuitive interaction. The user interface is thoughtfully structured, allowing users to navigate seamlessly between features, ensuring a user-friendly experience for individuals of varying technological proficiencies.

UX.3 **Desirable:** The design incorporates visually appealing elements, such as warm and comforting colors, ensuring a positive emotional impact on users. The user interface is crafted to be aesthetically pleasing, fostering a sense of familiarity and comfort during interaction.

UX.4 **Findable:** The web app emphasizes navigability, with clear and well-organized menus and sections. Users can easily locate and access desired features, ensuring a straightforward and efficient search and interaction process.

UX.5 **Accessible:** Accessibility is a core principle, with features designed to cater to users with diverse hearing loss. The web app includes options for customization, font-size adjustments, and other accessibility settings to enhance usability for all users.

UX.6 **Credible:** The web app establishes credibility through transparent communication, clear error handling, and user guides. By providing reliable and accurate information, the application builds trust and confidence among users.

UX.7 **Valuable:** The web app's value proposition lies in its commitment to delivering meaningful and accessible assistive solutions. Regular updates, user engagement studies, and responsiveness to user feedback ensure continuous improvement aligned with the mission of enhancing the user experience and meeting the unique needs of the target audience.

5 Implementation

5.1 Chosen Technologies

Figma was chosen for wireframe development due to its versatile and collaborative features. This cloud-based design platform enables real-time collaboration, concurrent editing, and seamless transitions from wireframing to prototyping. Its accessibility through a web browser ensures cross-platform compatibility, aligning with the project's collaborative and user-centric ethos.

React with TypeScript (TSX) has been chosen as the front-end framework for the development of this web application. Leveraging TypeScript's static typing ensures enhanced code quality and reduces the occurrence of runtime errors. The component-based architecture of React simplifies the process of user interface (UI) development, providing a dynamic and responsive user experience to fulfill the MH.3 requirement outlined in Section 3.2. Utilising React provides us with the capability to fulfill SH.3, as the web-app is accessible on any device with a web browser. We even verified this functionality by successfully testing it on a Raspberry Pi.

Node.js serves as the selected back-end framework, responsible for handling server-side operations and facilitating interactions with the chosen database. Its asynchronous, non-blocking I/O capabilities make it well-suited for managing concurrent connections, thereby ensuring optimal performance and responsiveness for user requests.

The database system utilised for storing user preferences and relevant data is Clerk. This choice aligns with the project's requirements, offering robust capabilities for user authentication. Clerk's implementation ensures a secure and seamless user experience, addressing the authentication needs of the web application effectively.

For audio processing, the project integrates two fundamental technologies: Web Speech API and Web Audio API. These APIs empower the web application with features such as speech-to-text, text-to-speech, and audio playback to fulfill the MH.1, MH.2, and MH.3 requirements aforementioned in Section 3.2. The Web Speech API facilitates accurate and efficient speech recognition, while the Web Audio API enhances audio processing capabilities, ensuring a seamless and high-quality user experience.

Several security measures have been implemented to safeguard user data and uphold a secure user experience. The utilisation of HTTPS ensures encrypted data transmission, mitigating potential security risks with easedropping (Kröger and Raschke 2019). Additionally, Clerk's robust user authentication mechanisms contribute to the overall security posture of the web application.

To establish the web application's public accessibility, the selection of an appropriate hosting platform was imperative. GitHub Pages emerged as the preferred choice due to its provision of free static website hosting, comparable to alternatives like Netlify, Cloudflare and Render, which were all contemplated during the project's planning phase. These platforms offer streamlined integration with Git version control, ensuring collaborative development and effective change tracking. GitHub Pages, integral to the GitHub ecosystem utilised for repository hosting, brings the advantage of seamless integration with other GitHub features, including issues and pull requests, streamlining the development workflow. Notably, GitHub Pages supports custom domain usage, enabling the utilisation of a personalised domain name for the website. Furthermore, GitHub fosters a sense of community where users can collaborate and leave feedback. This technology allows for the project to fulfill the CH.2 and CH.3 features outlines in Section 3.2.

During the course of this project, the WH.2 requirement outlined in Section 1.2 was thoroughly researched and initial implementation efforts commenced utilising Picovoice Eagle (Wordmark). Despite the availability of numerous examples for implementation, akin to web development practices, the model provided failed to function as expected. Subsequently, upon contacting the company to address a potential bug in the model, no response was received. Although this feature can potentially be a stepping stone for real-time captioning to distinguish between speakers, the priority of this feature was downgraded within the project's requirements priorities due to the bugs present.

5.2 Technological Stack

It is important to note that, throughout the project, I implemented verbal code reviews with fellow Computing Science students after developing each feature. This helped to catch bugs and improve code readability. It also provided valuable learning opportunities, allowing reviewers to familiarise themselves with the codebase and learn new technologies and techniques. Overall, code reviews fostered a collaborative environment for continuous improvement and skill development to produce a better project (Atlassian).

5.2.1 Web Speech API

Following our aims and requirements of the web-app in Sections 1.2 and 3.2, we identify Web Speech API to be the ideal method of providing precise and accurate speech recognition and speech synthesis. The Web Speech API empowers the integration of voice data into web applications, consisting of two primary components: SpeechSynthesis for Text-to-Speech capabilities and SpeechRecognition for Asynchronous Speech Recognition (Docs 2023).

The functionality of the Web Speech API's speech recognition involves capturing speech through a device's microphone. This input undergoes scrutiny by a speech recognition service against a predefined grammar list, essentially the vocabulary designated for recognition in a specific application. Upon successful recognition of a word or phrase, the result is returned in the form of a text string, triggering subsequent actions.

This capability facilitates the real-time conversion of spoken language into written text, offering precise and responsive speech-to-text capabilities within the web application. Users can employ this feature for immediate captioning, enhancing effective communication. The access to speech recognition is streamlined through the SpeechRecognition interface, which interprets voice content from an audio input, usually through the device's default speech recognition service. The interface's constructor is utilised to instantiate a new SpeechRecognition object, equipped with diverse event handlers for detecting speech input from the device's microphone. Additionally, the SpeechGrammar interface serves as a container for a specific grammar set that the application should recognize, defined using JSpeech Grammar Format (JSGF).

The Web Speech API seamlessly integrates speech synthesis, facilitating the conversion of text into spoken language. This essential functionality empowers the web application with text-to-speech capabilities, enabling users to receive clear, natural, and customisable auditory feedback, fostering vocal communication. The SpeechSynthesis interface acts as the primary controller for this feature, representing a text-to-speech component that enables programs to articulate their text content. This is achieved through the device's default speech synthesizer. Various voice options are captured by SpeechSynthesisVoice objects, while distinct sections of text designated for speech are encapsulated by SpeechSynthesisUtterance objects. The initiation of spoken output is carried out through the SpeechSynthesis.speak() method. Additionally, the API leverages the

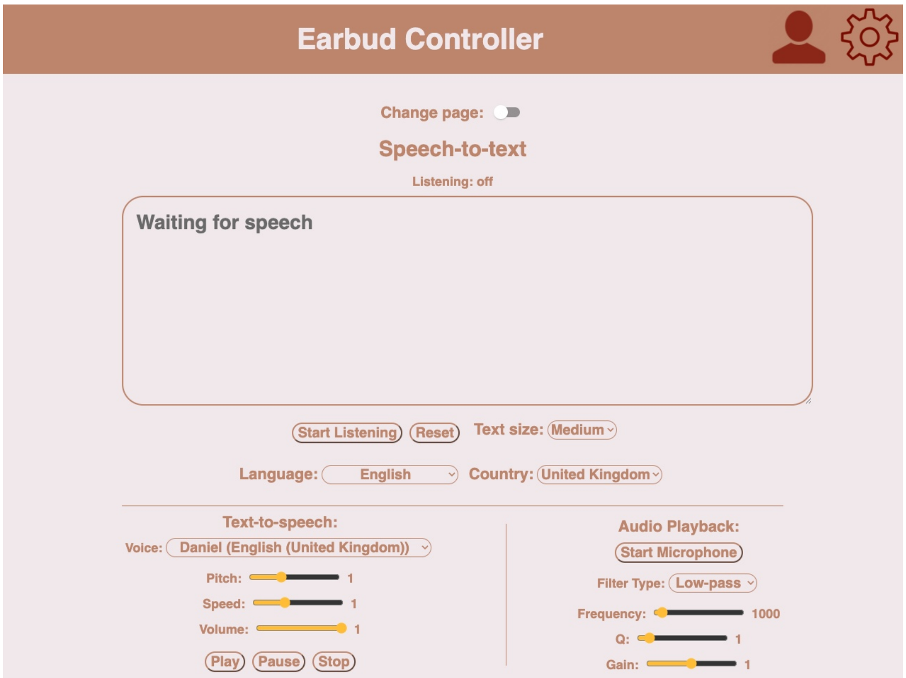


Figure 5.1: This screenshot displays the speech-to-text page of the web-app.

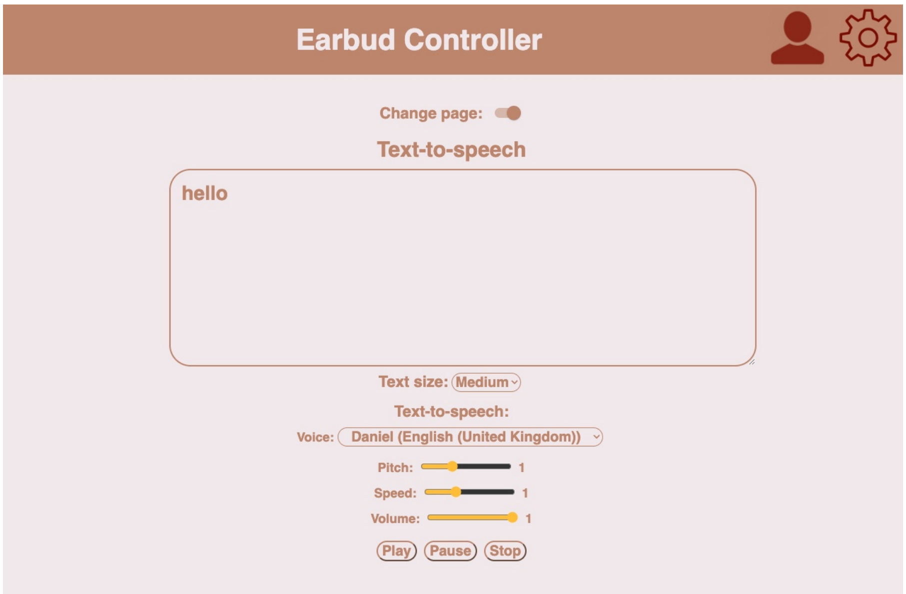
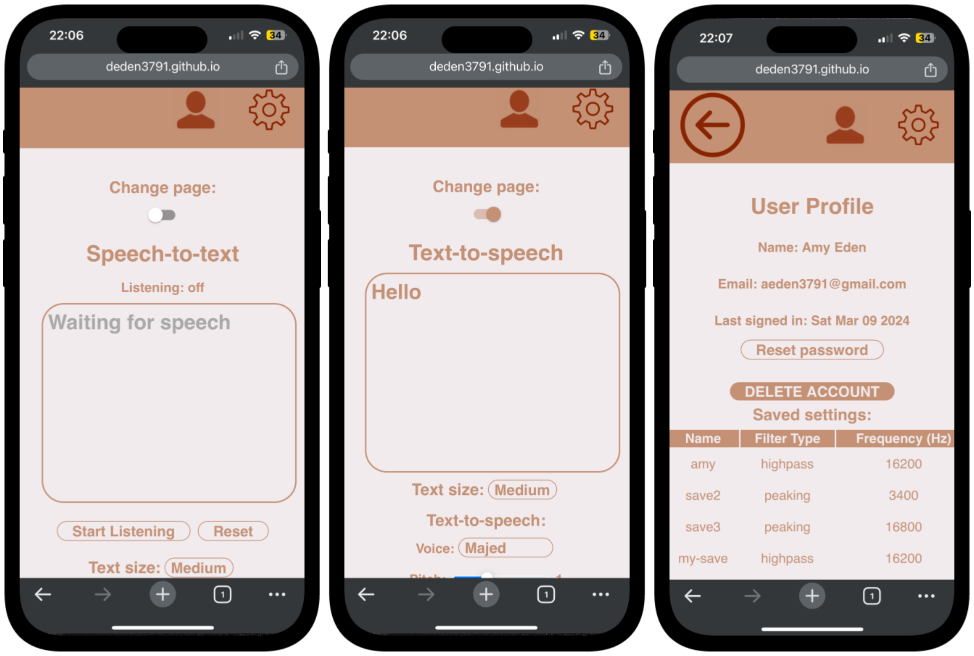


Figure 5.2: This screenshot displays the text-to-speech page of the web-app.



(a) This screenshot displays the speech-to-text page of the web-app on iPhone. *(b) This screenshot displays the text-to-speech page of the web-app on iPhone.* *(c) This screenshot displays the user profile page of the web-app on iPhone.*

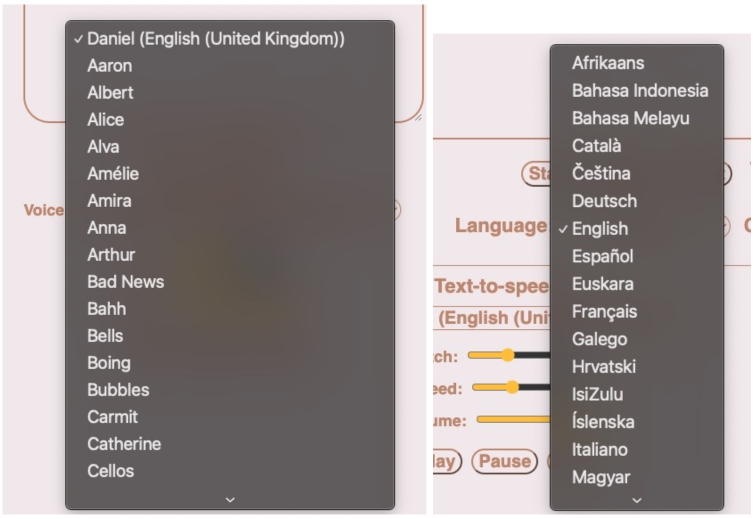
Figure 5.3: *Displays the web-app on an iPhone to fulfill requirement SH.3 in Section 3.2.*

speech synthesis systems inherent in most operating systems, ensuring compatibility and optimal performance for this task. These speech recognition interfaces are utilised to create the captions page as seen in Figure 5.1. Along with these screenshots we can see Figure 5.3 for each page on a mobile device.

One of the key advantages of leveraging the Web Speech API is its flexibility in customisation. The assistive web application incorporates user preferences for speech recognition and synthesis, allowing individuals to tailor the experience according to their specific needs. This includes adjustments in speech rate, pitch, and language selection utilising Web Audio API, contributing to a personalised and adaptive environment to achieve requirement SH.1 aforementioned in Section 3.2. This is featured in the text-to-speech page seen in Figure 5.2 along with the speech synthesis methods displayed in Figure 5.4a.

Another key advantage of leveraging the Web Speech API is its extensive language support. This feature enables web applications to cater to a diverse audience by providing speech recognition and synthesis capabilities in a wide range of languages to satisfy the SH.5 requirement outlined in 3.2 as displayed in 5.4b. Users can interact with the application using their preferred language, enhancing accessibility and inclusivity to achieve the PA2 outlined in 1.2. This multilingual support expands the reach of the web application to global users, fostering communication and engagement across different linguistic backgrounds, specifically those in countries classified as low- and moddle-income as aforementioned in Section 1.1. Additionally, the ability to recognise and synthesise speech in various languages contributes to the versatility of the application, making it adaptable to different cultural contexts and user preferences. This comprehensive language support underscores the Web Speech API's efficacy in facilitating effective communication and interaction within web applications.

A drawback of utilising this API is its dependency on an internet connection. Throughout this



(a) This screenshot displays just a fraction of the various voices you can choose from for the text-to-speech feature.

(b) This screenshot displays the range of languages the web-app supports.

Figure 5.4: *This figure displays the various voices and languages we implemented in the web-app from the Web Speech API.*

project, we explored a range of technologies, one of which was the Python library called vosk (PyPI). The code snippet for this library can be found in listing ??, and we initially considered integrating it into the web application. However, the process of integrating Python into React proved to be cumbersome and ultimately not worthwhile due to the significant inaccuracies in the output generated by this method.

5.2.2 Web Audio API

The Web Audio API provides developers with a comprehensive toolset for manipulating audio on the web. At its core, it operates within an audio context, where audio operations are orchestrated through a modular routing system. Central to this system are audio nodes, which are interconnected to form an audio routing graph. These nodes represent various audio sources, including oscillators, audio buffers, media elements, and streams (MozDevNet).

The audio graph allows for dynamic processing of audio samples, enabling developers to apply effects, modify volume levels, and spatialize audio. Effects nodes, such as reverb, filters, panners, and compressors, can be inserted into the graph to shape the audio according to desired characteristics. This is how we achieved the audio playback requirement (MH.4), outlined in 3.2.

Ultimately, the processed audio is routed to a destination, typically the system speakers or headphones, allowing users to hear the audio output. The precise timing and low latency of the Web Audio API enable accurate event response and sample targeting. This capability makes it suitable for applications like drum machines and sequencers, but we are here to test its ability to function as an assistive feature. Additionally, the API offers spatialization controls, facilitating the placement of audio sources within a virtual environment relative to a listener. This versatile functionality empowers developers to create immersive and interactive audio experiences on the web.

We use the Web Audio API's workflow to create an audio context, which serves as the foundation for all subsequent audio operations. Within this context, various audio sources are instantiated,

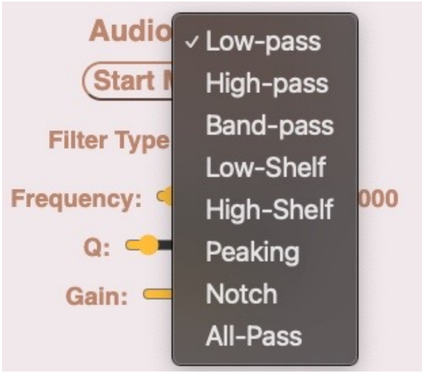


Figure 5.5: *This screenshot displays the various filter types utilised in the BiquadFilterNode method to provide the audio playback feature.*

ranging from elements like <audio> tags to programmatically generated oscillators and streaming audio sources.

Next, we can enhance and manipulate the audio by introducing effects nodes into the audio graph. These nodes, such as reverb, biquad filters, panners, and compressors, allow for a wide range of audio processing possibilities, from spatialization to dynamic tonal shaping. For this context, we are utilising the biquad filters.

Once the audio has been appropriately modified, we designate the final destination for the audio output, typically the system speakers or headphones. Finally, the sources are connected to the effects nodes, and the effects nodes are connected to the destination, establishing the desired audio flow within the context. This structured approach facilitates the creation of rich and dynamic audio experiences on the web, with precise control over every aspect of the audio signal's journey.

The BiquadFilterNode, established via the createBiquadFilter method within the BaseAudio-Context, serves as a pivotal element of the Web Audio API. Functioning as an AudioNode, it embodies various filtering operations, basic tone control, and graphic equalisation. Crucially, each BiquadFilterNode maintains a uniform structure comprising a single input and output, ensuring consistency within audio processing configurations.

With the implementation, users can effortlessly choose from various filter types, as displayed in Figure 5.5, using the Web Audio API. The user guide - on the web-app and GitHub - details how each variable influences sound output. Users can email the contact in the user guide to provide any feedback to satisfy the SH.2 requirement aforementioned in Section 3.2. In signal processing, these filters serve distinct roles. A low-pass filter allows frequencies below a set cutoff, attenuating higher frequencies, while a high-pass filter does the opposite. The band-pass filter selectively allows a range of frequencies, and a notch filter attenuates a narrow band. Meanwhile, a peaking filter adjusts gain within a specific frequency range. Moreover, there are two more types: the all-pass filter, which allows all frequencies to pass through unchanged but alters their phase relationship, and the notch filter, which attenuates a narrow band of frequencies centered around a specified frequency (MozDevNet 2023). This versatile toolkit, integrated through the Web Audio API, proves essential for shaping signals in applications like audio processing and telecommunications.

5.2.3 Clerk

Clerk, an advanced cloud-based authentication service, is specifically tailored for React applications, presenting a comprehensive and developer-friendly solution. Its seamless integration capabilities extend to various front-end frameworks, including React, Next.js, Remix, React Native, and

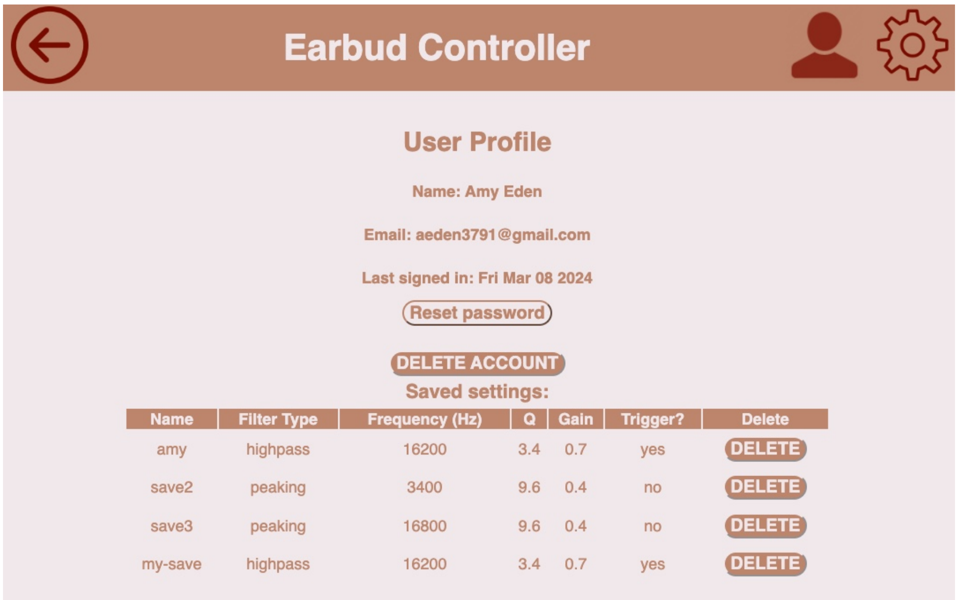


Figure 5.6: *This screenshot displays the profile page, including the various save instances for the audio playback feature utilising Clerk.*

Expo, ensuring broad accessibility within the developer community. Addressing the intricacies of authentication and user management, Clerk boasts a feature-rich environment encompassing user management, organisations, email and SMS authentication, password-less authentication, social login, multi-factor authentication (MFA), multi-sessions, device management, and password leak protection for a save user experience (Grant 2023). Notably, Clerk distinguishes itself through unique features like Organizations, Multi-Sessions, and Device Management, making it a preferred choice for applications requiring these advanced functionalities.

The React API offered by Clerk significantly enhances developer productivity by providing pre-built components and a hook-based API. These components, including SignIn, SignUp and UserButton, offer ready-made solutions for common authentication tasks such as sign-up, sign-in, and password reset forms (Clerk). This feature streamlines the development process, reducing the overhead associated with creating custom UI components to upkeep requirement UI.2 outlined in Section 4.3. Moreover, Clerk's React API supports themes, facilitating easy integration with existing website styles, while allowing extensive customization to align with specific design preferences. The combination of user-friendly features, seamless React integration, and robust security measures establishes Clerk as a compelling choice for developers seeking a sophisticated authentication solution for React applications.

In addition to its robust authentication features, Clerk offers a seamless solution for storing and managing saved instances for users, enhancing the overall user experience. Leveraging Clerk's capabilities, developers can implement a secure and efficient system for users to save and select instances as needed within their applications. The versatility of Clerk extends beyond authentication, providing a well-integrated approach to user data storage. By utilising Clerk's functionality, developers can design and implement features that allow users to conveniently save their instances, facilitating a personalized and tailored experience. This capability aligns with Clerk's commitment to providing a comprehensive environment for developers, further solidifying its position as a versatile and sophisticated choice for React applications seeking both authentication and data management solutions.

This solution for storing and managing saved instances for users is how we implemented a secure

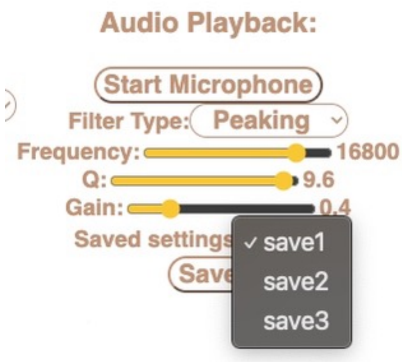


Figure 5.7: *This screenshot displays how the playback feature changes, so they can save instances, when the user is logged in utilising Clerk.*

system, enabling users to save and select playback instances for the hearing aid feature within the application. Moreover, users have the ability to assign any saved setting as a trigger word. This functionality allows the speech-to-text feature to automatically switch to the corresponding saved instance when detecting the name associated with it. For instance, if a user has saved an audio instance named 'Amy' for a friend named Amy, and the speech-to-text detects the name 'Amy' (e.g., if the user says "Hello Amy"), the audio playback feature will seamlessly transition to the 'Amy' saved instance.

This versatility extends beyond authentication, aligning with our project's commitment to providing a comprehensive environment for users, solidifying Clerk as a sophisticated choice for our React web-app, addressing both authentication and data management needs. We, also, utilised Clerk to create the user profile page displayed in Figure 5.6 with the ability to save audio playback instances displayed in Figure 5.7. These components allowed us to achieve SH.4 outlined in Section 3.2.

6 User Study

To delve deeper into the perspectives of the Deaf and Hard of Hearing community regarding assistive technology, I conducted two surveys to fulfill the PA.4 requirement outlined in Section 1.2. The first survey aimed to gather insights into individuals' expenditure on assistive technology, their satisfaction levels with existing solutions, and introduced a new assistive web-app, inviting feedback on desired features and improvements. The second survey focused on assessing the overall usability of assistive technology, extending the evaluation to include the renowned System Usability Scale (SUS) Brooke (1995).

In this scale, prompts with positive connotations contribute to the score by subtracting one from the scale position, while prompts with negative connotations contribute by subtracting the scale position from 5. The sum of these scores is then multiplied by 2.5 to determine the overall value of SUS. Typically, the scale ranges from 0 to 100, but we included additional questions related to the project, which ranges from 0 to 150. We ensured to calculate the official System Usability Scale (SUS) using only the 10 questions outlined in the paper, while the modified SUS incorporates the entire survey.

Throughout the engagement with the Deaf and Hard of Hearing community, it became evident that there's a prevailing sentiment of fatigue with researchers constantly approaching them for surveys. Many expressed feeling like they're treated as exhibits in a zoo, where researchers come in, conduct surveys, and leave without meaningful engagement or follow-up. It's crucial to acknowledge and respect these sentiments, recognising that Deaf individuals are not mere subjects for research but individuals with their own agency and perspectives. Referring to the outlined project aim PA.3 in Section 1.2 we must take this perspective into consideration and question if this project is even needed in this society.

These surveys aim to address the scarcity of research on new assistive technologies by directly involving the Deaf and Hard of Hearing community. By gathering feedback, we seek to ensure that new technologies meet the actual needs and preferences of users. This collaborative approach prioritizes inclusivity and moves away from traditional researcher-subject dynamics. Ultimately, our goal is to empower the community by amplifying their voices and perspectives in the development and evaluation of assistive technologies, creating solutions that are technologically advanced, accessible, and inclusive for all users.

6.1 Affordability and Satisfaction of Assistive Technology

This survey aims to gather information on Deaf and Hard-of-Hearing individuals' expenditure on assistive technology, their satisfaction levels with these technologies, and to introduce the new assistive web-app.

The survey received 47 responses, predominantly comprising individuals aged between 45 and 64. This reflects a notable demographic skew towards an older population. This demographic trend suggests a pertinent focus on the needs and preferences of middle-aged and older individuals regarding assistive technologies, as aforementioned in 4, making it imperative to tailor the user interface and experience (UI/UX) of the web-app to cater to the needs and preferences of older individuals. Furthermore, the survey indicates a prevalent degree of hearing loss among respondents, with a substantial 66% experiencing moderate hearing loss.

The survey reveals a widespread reliance on assistive technologies among respondents, with nearly half utilising hearing aids. Despite this widespread adoption, satisfaction levels vary across different assistive technologies. Notably, while 29.5% express high satisfaction with hearing aids, 41.7% report significantly less satisfaction with speech-to-text capabilities, with 75% report even less satisfaction with text-to-speech capabilities. These varying satisfaction levels highlight the importance of understanding user preferences and expectations to enhance the effectiveness of assistive technologies.

A noteworthy finding from the survey pertains to the funding status of assistive technologies, with a slight majority of respondents receiving funding support. However, satisfaction levels differ between funded and non-funded individuals, with funded users expressing higher levels of satisfaction. Interestingly, while a considerable portion of respondents spend nothing on assistive technology, those who do report spending over £1,000 exhibit varying levels of satisfaction. These findings emphasise the complex interplay between funding, costs, and user satisfaction in the adoption and utilisation of assistive technologies.

The survey delves into the diverse expectations of assistive technologies by users, including clear sound quality, accurate speech-to-text capabilities, ease of use, and customisation options. Respondents also express preferences for discounted pricing, long battery life, and compatibility with other devices. However, concerns regarding privacy, data security, and the effectiveness of proposed products in noisy environments are also apparent. While approximately 28% of respondents find the proposed product useful, various suggestions and apprehensions highlight the importance of addressing user concerns and preferences in the development of assistive technologies. For instance, a participant proposed a feature for the web-app, suggesting the incorporation of an arrow indicating the direction from which sounds, and consequently speech, emanate. This would aid the user in identifying the speaker. Due to the project's defined scope, this feature was not implemented. However, it warrants consideration for future iterations of the web-app.

In conclusion, the survey offers valuable insights into the usage patterns, satisfaction levels, and expectations regarding assistive technologies among individuals with hearing loss. These findings underscore the importance of user-centric design, cost considerations, and compatibility with existing technologies in the development and implementation of assistive solutions. By addressing user needs and concerns, developers and policymakers can ensure the creation of more effective and inclusive assistive technologies to enhance the quality of life for individuals with hearing loss.

6.2 Usability of Assistive Technology

This survey aims to gather information on individuals' thoughts on the usability of assistive technology by analysing broth quantitative and qualitative data. The survey garnered 24 responses, with the majority of participants being in the 35-44 age range, providing a broad perspective on the topic. A significant portion of participants reported profound hearing loss, highlighting the prevalence of severe impairment within the surveyed population. The average, modified System Usability Scale (SUS) score of 81.04, alongside an official SUS score of 59.69, indicates generally positive usability experiences with assistive technologies, suggesting room for improvement in specific areas. This data is visualised in Figure 6.1.

During the qualitative analysis, we found that many respondents expressed frustration with phone communication due to inaccurate transcription apps and unreliable transcribing services during calls. This poses significant obstacles in both professional and social interactions. This is why we tested the speech-to-text feature during a phone call, and it effectively transcribed the conversation.

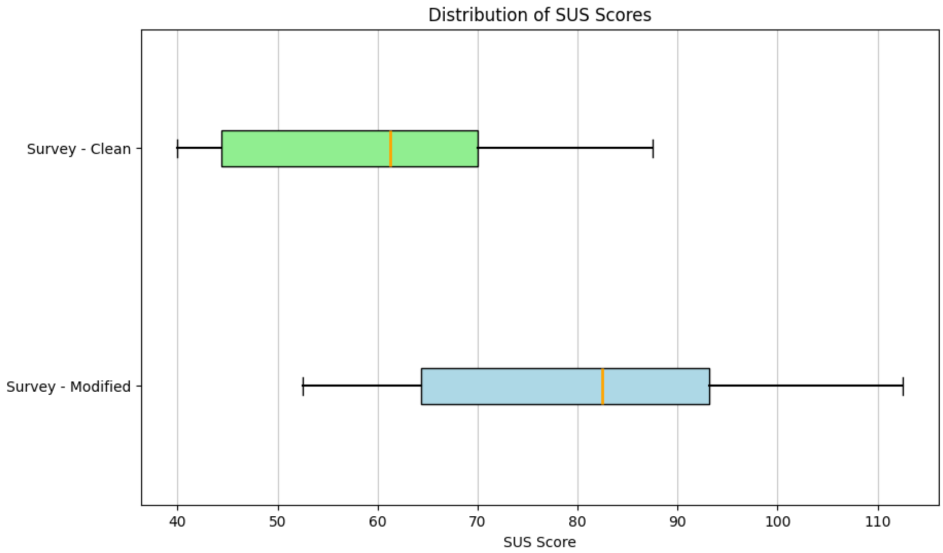


Figure 6.1: *This figure compares responses to 10 original questions with those to 2 additional questions probing perspectives on current assistive technology solutions.*

Issues with hearing aid functionality were also highlighted, including vulnerability to moisture, intrusive sound dampening features, and limitations in streaming capabilities. However, there were positive experiences shared regarding the effectiveness of assistive technology in empowering Deaf children, fostering independence, and facilitating educational and social development.

Despite advancements, current assistive technologies still face limitations, including reliance on Wi-Fi connectivity, challenges in noisy environments, and the need for a holistic communication approach beyond technology alone. Some respondents also highlighted the financial burden associated with acquiring assistive devices, emphasising the need for better accessibility and support programs to alleviate costs, which this project aims to do.

Additionally, respondents emphasised the significance of accommodating Deaf individuals' language preferences and recognising the limitations of spoken language-centric technologies. One respondent articulated this sentiment by questioning why Deaf individuals should be expected to communicate in a second language, while hearing individuals often neglect to learn Deaf individual's first language. They challenged assumptions about Deaf people's desire to communicate in English and questioned the underlying assumption that solutions are needed to fix perceived problems. Instead, they highlighted that technology should serve to facilitate communication between hearing and Deaf individuals, rather than aiming to fix Deaf individuals themselves. This perspective sheds light on the concept of audism and underscores the importance of understanding and respecting Deaf culture and language.

In summary, while there are positive aspects to current assistive technologies, there remain significant challenges and areas for improvement in usability, communication options, financial accessibility, and design considerations to better meet the diverse needs of users with hearing loss.

7 Evaluation

The aim of this project, as outlined in 1.2, is to evaluate a proof of concept for an affordable solution tailored to the needs of the Deaf and Hard-of-Hearing (DHH) community, leveraging the Web Speech API. By incorporating features described in Section 3.2, the web application aims to foster inclusivity for individuals with various types of hearing loss. The primary objective is to provide cost-effective accessibility solutions while accommodating the diverse needs of the DHH community for the web-app to be usable. Furthermore, the project endeavours to create an open-source platform, enabling ongoing development and improvement beyond my own dissertation.

Evaluating assistive technology (AT) involves assessing the extent to which ATs effectively enhance participation, address loss, and alleviate health-related limitations. In essence, it entails gauging the overall effectiveness and usefulness of assistive technologies (Tao et al. 2020). This is why we chose to employ the System Usability Scale (SUS) (Brooke 1995), as we used in Section 6 to evaluate if requirement PA.4 outlined in Section 1.2 has been achieved. Furthermore, leveraging the findings from the analysis conducted in Section 6, we will undertake a comparative evaluation of the usability between the existing assistive technology and the newly introduced assistive web application. To evaluate the requirements outlined in Sections 4.3 and 4.4, we expanded upon the original 10 System Usability Scale (SUS) statements, creating 17 statements while adhering to the same scoring methodology (i.e., subtracting one for positive statements and subtracting the score from 5 for negative statements). We will analyse the original SUS score along with our modified version. In addition to this both quantitative analysis, we also utilise the qualitative data for a deeper analysis.

The assessment was distributed to individuals with and without hearing loss to complete 8 sections. Hearing individuals are encouraged to use noise-canceling earbuds or headphones, when testing the web-app. This is intended to provide a modest simulation to assess the effectiveness of the web-app.

7.1 Overview

Each participant was first encouraged to familiarise themselves with the app's interface before proceeding to the five sections, each containing a distinct task outlined in the provided user guide [1]. Upon completion of each task, the participant is kindly asked to answer questions sharing their experience and feelings regarding the utilisation of the feature.

The evaluation intends to assess each feature against qualities such as:

- Intuitive: does each feature feel intuitive?
- Accurate: does each feature perform accurately?
- Comfortable: does each feature feel comfortable to use?
- Usable: is each feature usable?

[1] `https://github.com/deden3791/L4Project/blob/main/main/UserGuides/UserGuide.md`

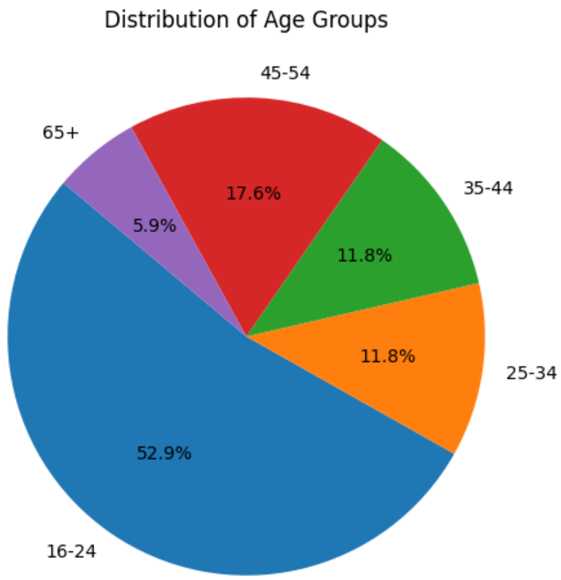


***Figure** 7.1: This pie chart displays the age distribution of participants.*

The evaluation received 17 responses, with the majority of participants falling within the 16-24 age range and exhibiting no signs of hearing loss. Nonetheless, Figures 7.1 and 7.2 illustrates the distribution of ages and hearing loss among our participants, highlighting the diversity of the demographic. When analysing both the quantitative and qualitative data of the responses, we meticulously distinguish between the perspectives of the hearing and the Deaf and Hard-of-Hearing community.

For each task, we utilised a Likert scale ranging from 'Strongly Agree' to 'Strongly Disagree'. However, we recognise the inherent challenges in analysing Likert scales (Harpe 2015; Pornel and Saldaña 2013). To address this, we convert the responses into numerical values and standardise them on a 0-4 scale, similar to how we handle the System Usability Scale (SUS). This approach helps mitigate common pitfalls associated with Likert scale interpretation.

7.2 Speech-to-text Feedback

For this task, the participants were asked to navigate to the speech-to-text section. They were then asked to click the "Start Listening" button to initiate the conversation to engage in verbal communication as desired. When the conversation is finished, click the "Stop Listening" button again to conclude the web-app's listening mode.

Both hearing and non-hearing participants unanimously agreed that the speech-to-text feature felt intuitive, accurate, and comfortable, with the lowest average being 3.2 as seen in Figure 7.3 However, some participants pointed out certain limitations, such as the feature's occasional failure to recognise acronyms like 'LOL', often mistaking them for 'Hello'. Additionally, one participant reported compatibility issues with the 'Ark' browser, attributed to restrictions within the Web Speech API that render it suitable for specific browsers. Furthermore, a few participants observed minor inaccuracies in the transcription, attributing them to their strong Yorkshire accent. The persistent challenge posed by accent variations in speech recognition systems is a notable issue, as discussed in Section 2.2.2, warranting further attention and awareness.

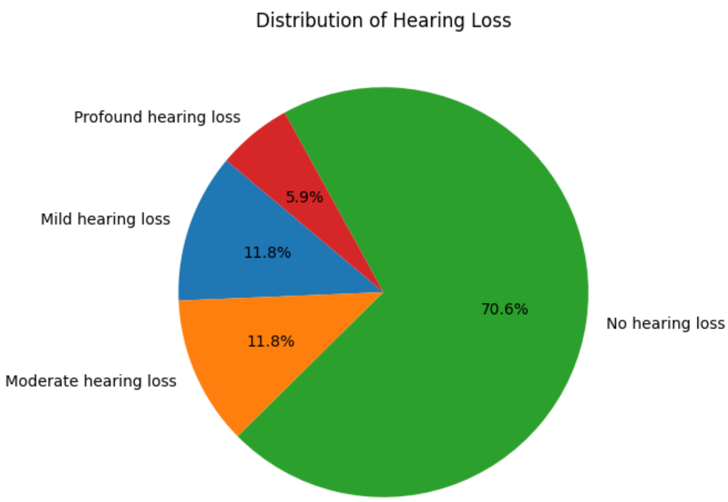


Figure 7.2: This pie chart displays the hearing distribution of participants.

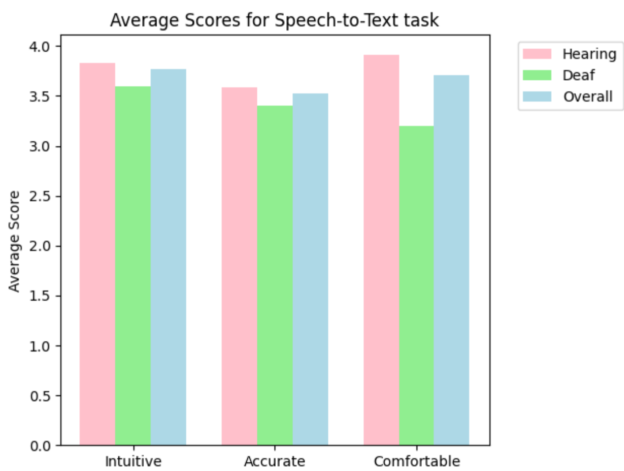


Figure 7.3: This bar chart displays the scores for the Speech-to-Text task on a 0-4 scale.

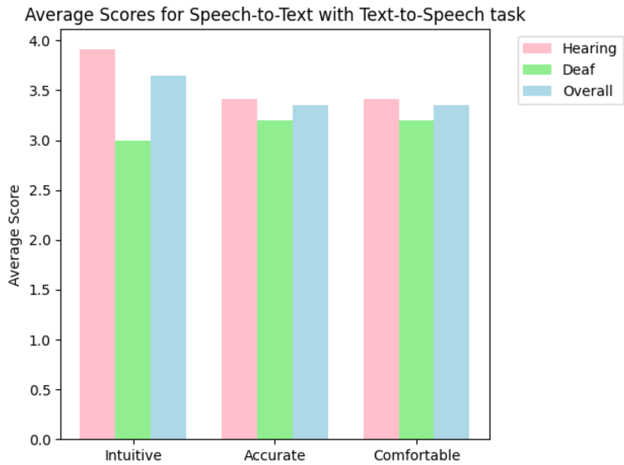


Figure 7.4: *This bar chart displays the scores for the Speech-to-Text with Text-to-Speech task on a 0-4 scale.*

7.3 Speech-to-text with Text-to-speech Feedback

For this task, participants are instructed to begin by using earbuds or headphones. They then proceed to access the speech-to-text feature as outlined in the previous task. Following this, they have the option to click the "Play" button on the text-to-speech feature to listen to the converted text. Additionally, participants are encouraged to modify the settings within each feature, such as languages for speech-to-text and speed, pitch, volume, and voice for text-to-speech. They can also pause or stop the speech-to-text and text-to-speech features as necessary during their interaction.

For hearing participants, the speech-to-text feature, along with the text-to-speech feature, was unanimously perceived as intuitive, with a rating of 3.9 as seen in Figure 7.4. However, we do see a 0.9 drop in the rating for our Deaf and Hard-of-Hearing participants. It is important to note that, there is a 0.5 drop in hearing participants finding the feature accurate and comfortable to use. This may be due to them experiencing no hearing loss so these features are not comfortable for them to use. For Deaf and Hard-of-Hearing individuals, the rating for accuracy and comfortability raised by 0.2.

Several participants noted that the accuracy of the speech-to-text feature varied significantly depending on the microphone used. For example, they found it more accurate when using headphones connected via wire or when using the laptop's built-in microphone instead of headphones. Others expressed frustration with default settings, mentioning that their devices would revert to default settings, requiring them to change language or voice settings repeatedly, which they found cumbersome. Additionally, some participants found the text-to-speech output to be choppy and unnatural, lacking the fluidity of a normal conversation. This aspect of speech synthesis is an ongoing area of improvement, as discussed in Section 2.2.3, and warrants further attention.

7.4 Audio Playback Feedback

For this task, participants were instructed to start by using earbuds or headphones for optimal audio experience. Next, they were guided to navigate to the audio playback section and encouraged to familiarise themselves with the available adjustable variables. Participants were informed that they

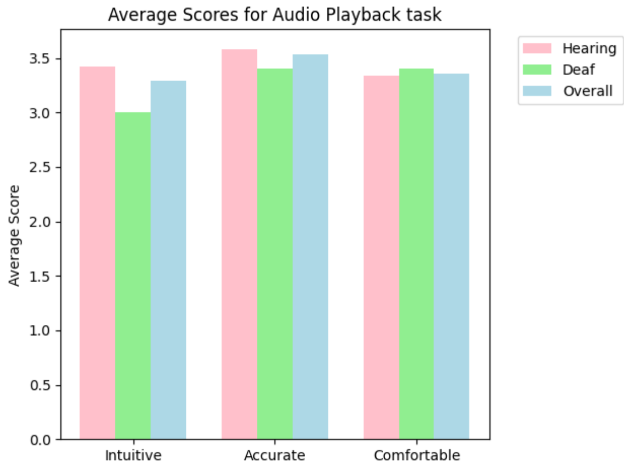


Figure 7.5: *This bar chart displays the scores for the Audio Playback task on a 0-4 scale.*

could access additional details by referring to the settings icon in the header. Once participants felt comfortable with their chosen settings, they were prompted to click the "Start Microphone" button to initiate the conversation. Additionally, participants were reminded of their flexibility to modify settings at any point during the interaction.

Hearing participants appeared to perceive the feature as intuitive, evidenced by a 3.4 rating as depicted in Figure 7.5. This rating slightly decreased by 0.4 for Deaf and Hard-of-Hearing participants. Nevertheless, all Deaf and Hard-of-Hearing participants unanimously considered the feature both accurate and comfortable, averaging a rating of 3.4. In contrast, hearing participants rated the accuracy slightly higher at 3.6, while their comfort rating was slightly lower at 3.3. Again, this discomfort experienced by some hearing participants could be attributed to their unaffected hearing capabilities, leading them to perceive the feature as uncomfortable due to its alteration of their finely tuned auditory experience. This contrasts the aid that this feature may give a Deaf user.

Some participants expressed confusion regarding the functionality of each variable in altering the surrounding sound, indicating a need for more emphasis on the user guide. Suggestions were made, including implementing pop-ups when hovering over the variables to provide clearer explanations. Additionally, although the abundance of variables was acknowledged as beneficial for usability, it was also noted that they could potentially overwhelm users and lead to feature abandonment. We acknowledge the significant abandonment rate associated with assistive technologies (Jiménez-Arberas and Díez 2021). Therefore, we have enhanced our user guide to provide greater clarity and understanding, aiming to mitigate this issue.

One participant highlighted the necessity for significant layout improvements within the web-app to be able to distinguish features with ease. They specifically noted the need for increased spacing between elements, noting frustration when using the web-app on a mobile devices due to oversized finger interactions with sliders. In response, we implemented borders between each feature and increased spacing between variables to address these usability concerns effectively.

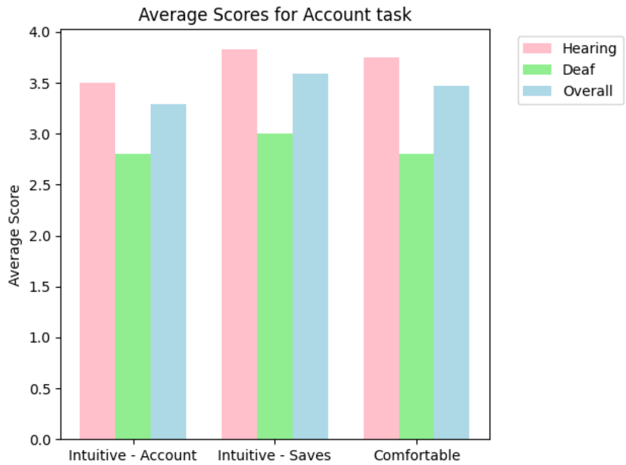


Figure 7.6: *This bar chart displays the scores for the Account task on a 0-4 scale.*

7.5 Creating an Account & Saving Audio Playback Instances Feedback

For this task, participants were directed to register by creating an account and explore the functionality of saving audio playback instances. They were instructed to click the account icon in the header and enter the required information for registration. After successfully creating their account, participants were prompted to navigate to the audio playback section. Here, they were encouraged to familiarise themselves with the options available for saving instances and experiment with any customisation features offered (i.e. changing the audio playback variables, saving this audio instance, and even making the save name a trigger word). If participants desired to delete any instances, they were instructed to access their user profile page using the account icon.

All hearing participants unanimously agreed that both the features and save instances felt intuitive and comfortable, with both aspects receiving an approximate rating of 3.8 as depicted in Figure 7.6. However, when it came to account creation, the rating for intuitiveness decreased by approximately 0.3. In contrast, for Deaf and Hard-of-Hearing participants, these ratings were notably lower, with account creation being perceived as intuitive at 2.8. Similarly, for save instances, only 3.0 found them intuitive. Comfort ratings were also lower among Deaf and Hard-of-Hearing participants, with a rating of 2.8. One participant noted that having the save button change colour when it is a saved setting might inform the user that a change has been made.

Despite many positive responses regarding the trigger word feature for naming saves, it was observed that the accuracy of save names depended heavily on how the speech recognition system formatted words. For instance, one participant named a save "passthrough," but the system outputted the text as "pass through," leading to the trigger word not being detected.

7.6 Text-to-speech Feedback

In this task, participants were guided to explore another feature within the application by clicking the toggle button to access another, separate text-to-speech functionality. They were instructed to enter a set of text or sentences into the designated input field and ensure they were comfortable with the content. Subsequently, participants were prompted to activate the text-to-speech

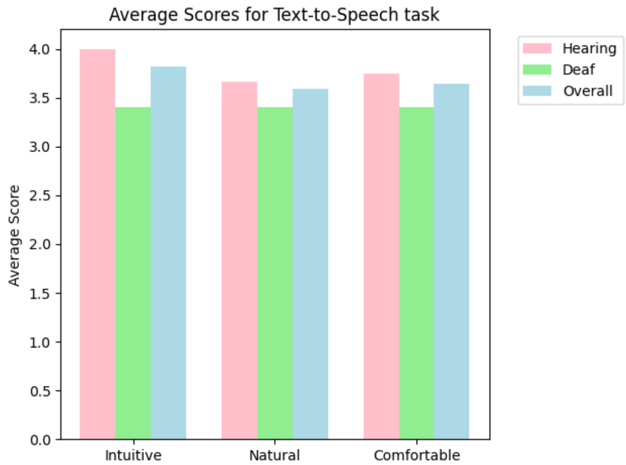


Figure 7.7: *This bar chart displays the scores for the Text-to-Speech task on a 0-4 scale.*

function by clicking the "Play" button and attentively listen to the synthesised speech output. They were encouraged to assess various aspects such as speech clarity, intonation, and overall satisfaction with the text-to-speech functionality.

All hearing participants unanimously expressed strong agreement that the text-to-speech feature felt intuitive, garnering a maximum rating of 4.0. Additionally, they concurred that the feature felt natural, receiving a commendable rating of 3.7. However, among Deaf and Hard-of-Hearing participants, we observed lower ratings for the perceived intuitiveness, accuracy, and naturalness of the feature, all averaging at 3.4.

Participants expressed the necessity of a save configuration option for the text-to-speech feature, as implemented for the playback feature. Additionally, some participants emphasised the need for a testing feature, enabling users to input a phrase for the system to repeat, facilitating adjustment of sliders for optimal settings. Along with these suggestions, participants reiterated their dissatisfaction with the choppiness of the voices.

7.7 Usability Feedback

As mentioned in Section 7, we modify the SUS to analyse if this web-app satisfies our project aims outlined in Section 1.2. Figure 7.8 shows the distribution of modified SUS scores to assess the PA.4 project aim, separated into hearing and Deaf and Hard-of-Hearing individuals. For our Deaf and Hard-of-Hearing participants, we calculated an average, modified SUS score of 94.4, with a higher 132.5 from our hearing participants. This score is out of a potential 170.

To assess whether the outlined requirements have been fulfilled, we employ a scoring system similar to the calculate of the System Usability Scale (SUS), ranging from 0 to 4. For a detailed mapping of questions to requirements, please refer to the appendix (Section A), tables A.1 and A.2. For an outline of the scores for overall, hearing and Deaf participants, please refer to Figure 7.9. When analysing the scores given by hearing participants and Deaf and Hard-of-Hearing participants, there exists an average deviation of 0.373 with a maximum deviation of 1.0. While this deviation may not be considered exceptionally large, it should still be deemed significant when justifying if this web-app will be effective to the Deaf and Hard-of-Hearing community.

As previously stated, each of the requirements outlined in Sections 4.3 and 4.4 is associated with the statements given in the evaluation. To analyse if these requirements were fulfilled, we took the average score for each statement and averaged the statements associated with that requirements,

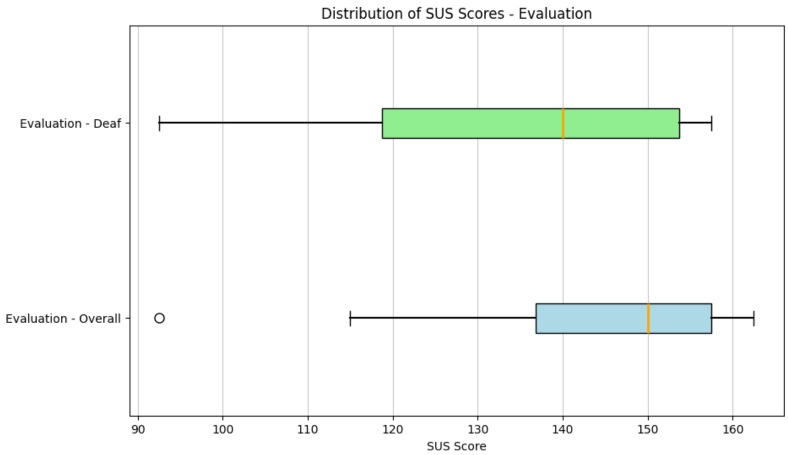


Figure 7.8: This shows the modified SUS scores for the evaluation, separated into hearing and Deaf and Hard-of-Hearing participants, that range from 0 to 170.

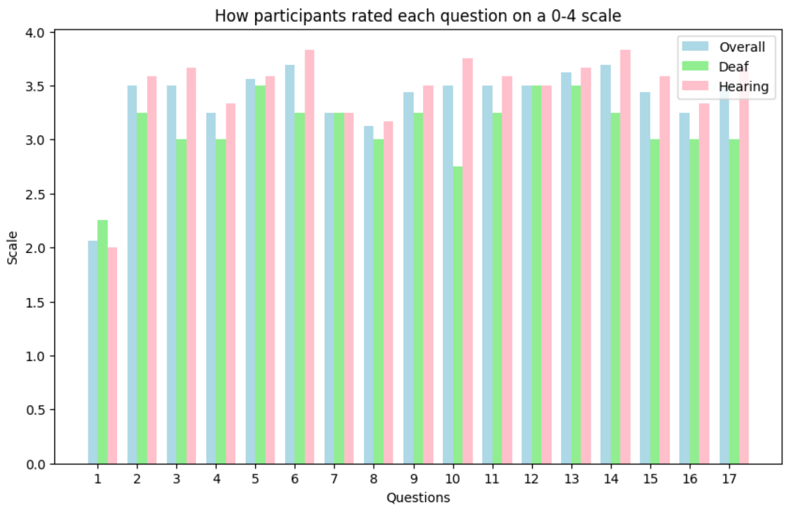


Figure 7.9: This shows the scores of each statement, separated into overall, hearing, and Deaf and Hard-of-Hearing participants, that range from 0 to 4.

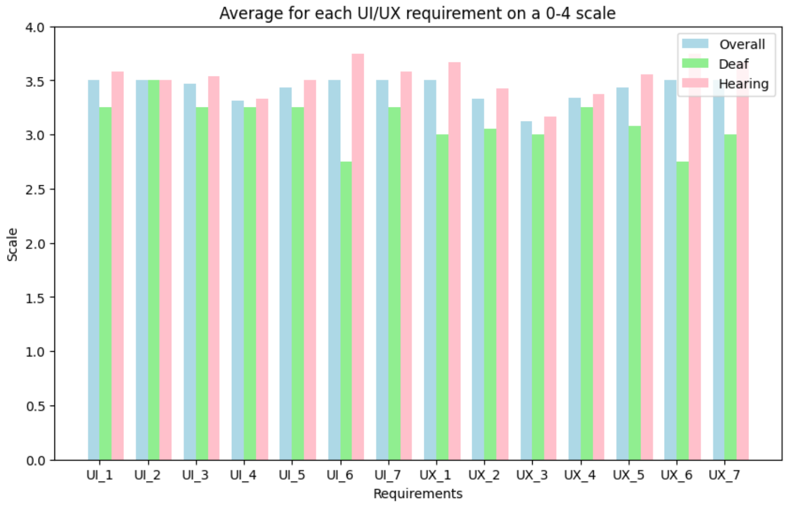


Figure 7.10: *This shows the average scores of each requirement, separated into overall, hearing, and Deaf and Hard-of-Hearing participants, that range from 0 to 4.*

as shown in Figure 7.10. The average deviation between hearing and Deaf and Hard-of-Hearing participants is calculated as 0.411, with a maximum deviation of 1. Again, while this deviation may not be considered exceptionally large, it should still be deemed significant when justifying if these requirements are fulfilled.

Upon examining the individual requirements, we found that each received a rating equal to or greater than 2.75 with an average of 3.12, indicating that all requirements were met to some extent. Notably, requirements UX.6 and UI.6, which pertain to error communication and user guides, respectively, emerged as areas in need of enhancement. Some participants encountered difficulties with the user guide, prompting us to revise it to provide clearer explanations of the system, thereby fostering better user understanding. The requirement with the highest rating was UI.2, reflecting the web-app's consistency. This positive feedback underscores the effectiveness of maintaining uniformity throughout the user interface, contributing to a smoother user experience

After removing our additional 7 statements and including only the original 10 SUS statements, we calculated an average SUS score of 75.6 from our hearing participants, with a slightly lower average SUS score of 70 from our Deaf and Hard-of-Hearing participants, as shown in Figure ??. Although this deviation of 5.6 isn't notably large, it does signify some areas for improvement in the web-app.

When comparing the SUS scores obtained in this evaluation to those obtained from the survey outlined in Section 6, a remarkable increase in scores is evident. In the survey, the SUS score obtained was 59.69, significantly lower than the score of 70 obtained from our Deaf and Hard-of-Hearing participants, as illustrated in Figure 7.11. This data strongly suggests that our web-app exhibits greater usability compared to the current assistive technologies utilised by the Deaf and Hard-of-Hearing community.

7.8 Limitations

While the evaluation provides valuable insights into the web-app's usability, many limitations should be considered when interpreting the results and making decisions based on them. For example, the evaluation's sample size of 17 participants is notably small, particularly with only 5

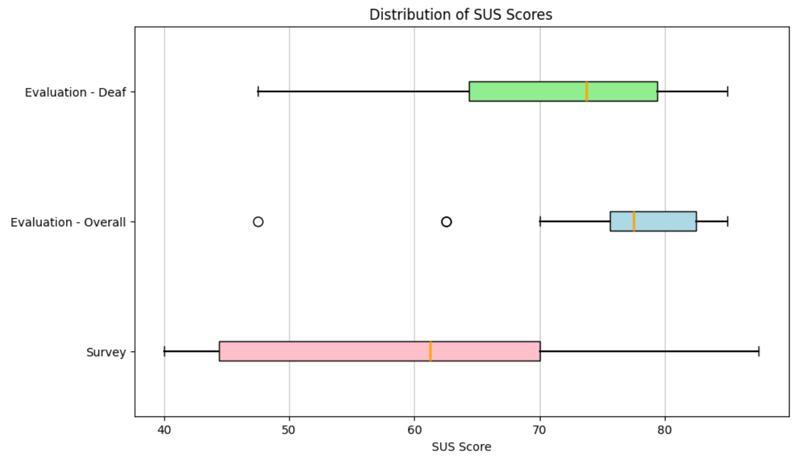


Figure 7.11: This shows the official SUS scores comparing the overall, Deaf and Hard-of-Hearing, and survey scores.

participants representing the Deaf and Hard-of-Hearing community. This limited representation may not capture the full spectrum of experiences and preferences within this user group, potentially impacting the ability to generalise the findings. Furthermore, it's important to note that several participants in the evaluation are family and friends, which could introduce bias into the results. Participants with personal connections to the developers or project team may provide feedback influenced by loyalty or familiarity, rather than unbiased assessments of usability.

Although efforts were made to simulate real-world scenarios and encourage natural interactions during the evaluation, the controlled nature of the study may still limit its applicability to diverse usage contexts. Individual preferences, prior experiences, and environmental factors can significantly influence usability in real-world settings, and these nuances may not have been fully captured in the evaluation. Efforts were made to recruit participants from diverse backgrounds by posting the evaluation on many Deaf and Hard-of-Hearing online groups, though the small sample size remains a constraint to keep in mind Expanding the participant pool to include a more diverse range of users, particularly from the target demographic of Deaf and Hard-of-Hearing individuals, would strengthen the validity and make it easier to generalise the findings.

8 Conclusion

This project was motivated by the gap in affordable assistive technology within the Deaf and Hard-of-Hearing community that legislation such as the United Nations Convention on the Rights of Persons with Disabilities (CRPD) (Nations) fails to bridge (Borg et al. 2011). Bridging this gap is specifically important due to the cost-of-living crisis (Broadbent et al. 2023), impacts of quality of life without assistive technology (Borre et al. 2023; Jiménez-Arberas and Díez 2021), with detrimental links to dementia (Powell et al. 2021; Lin et al. 2011; Livingston et al. 2020; Loughrey et al. 2018) and mortality (Choi et al. 2024).

Before deep-diving into developing assistive technology, it is important to understand Deaf culture throughout history to know what kind of assistive technology you develop: through cultural perspective or an infirmity model (Burrows 2022). From the rudimentary ear-trumpets constructed from skeletal remains (CalHearing 2023) to the current development of programmable multi-modal hearing aids (Shah et al. 2022), assistive technology has undergone rapid evolution. However, alongside these advancements, the associated costs have also escalated (Firm).

Throughout this project, we define many requirements to be fulfilled utilising Web Speech API to prove that low cost assistive technology can be made. The evaluation results show a usability rating of 70 out of a possible 100 from the Deaf and Hard-of-Hearing community. Alongside this, the requirements we outlined received an average rating of 3.12. This proves that the web-app fulfilled the requirements with room for improvement.

Future work could entail focusing on the initial plan of the project for real-time amplification and edge-computing. In relation to the current project, it is worthwhile delving deeper into speech recognition, particularly in terms of accent variability (Lopez-Lloreda 2020; Koenecke et al. 2020). This emerged as a notable challenge during the evaluation process. Additionally, there is a need to explore complex features outlined in requirement WH.2 in Section 1.2, such as speaker detection, siren detection, and offline capabilities, which were not fully implemented in this project. Although some progress was made towards speaker detection, no concrete feature was developed. However, this functionality holds promise as a foundational element for real-time captioning systems aimed at distinguishing between speakers. Moreover, as discussed in Section 6, an intriguing potential enhancement involves speaker detection leveraging directional sound cues. This feature would visually represent the speaker's orientation on the screen, potentially through an arrow indicator. Such advancements could significantly enhance the accessibility and usability of the web-app for users with diverse needs.

Although this project did delve into the Deaf and Hard-of-Hearing community for their opinions and guidance on assistive technology, it is important to note that this web-app is an infirmity model as the web-app's features aim to modify potential users hearing loss to accustom Deaf users to a hearing society. As aforementioned in Section 6, the Deaf and Hard-of-Hearing community are tired of the countless surveys being forced upon them with no significant outcome for their community. We should consider their opinions about replacing assistive technology with sign language to integrate both hearing and Deaf communities together, such as Deafness on Martha's Vineyard (Groce 2016), leading to a fully integrated community where communication barriers and feelings of isolation due to their hearing loss are rare.

A Appendices

Table A.1: *This table displays the scores for each question in relation to the UI requirements outlines in Section 4.3, from a scale of 0-4 utilising the same method outlined in Section 7.*

Questions	UI.1	UI.2	UI.3	UI.4	UI.5	UI.6	UI.7
1	–	–	–	–	–	–	–
2	✓	–	–	–	–	–	–
3	–	–	–	–	–	–	–
4	–	–	–	–	–	–	–
5	–	–	–	–	–	–	–
6	–	–	✓	–	–	–	–
7	–	–	✓	–	–	–	–
8	–	–	–	✓	–	–	–
9	–	–	–	–	✓	–	–
10	–	–	–	–	–	✓	–
11	–	–	–	–	–	–	✓
12	–	✓	–	✓	–	–	–
13	–	–	–	–	–	–	–
14	–	–	–	–	–	–	–
15	–	–	–	–	–	–	–
16	–	–	–	–	–	–	–
17	–	–	–	–	–	–	–

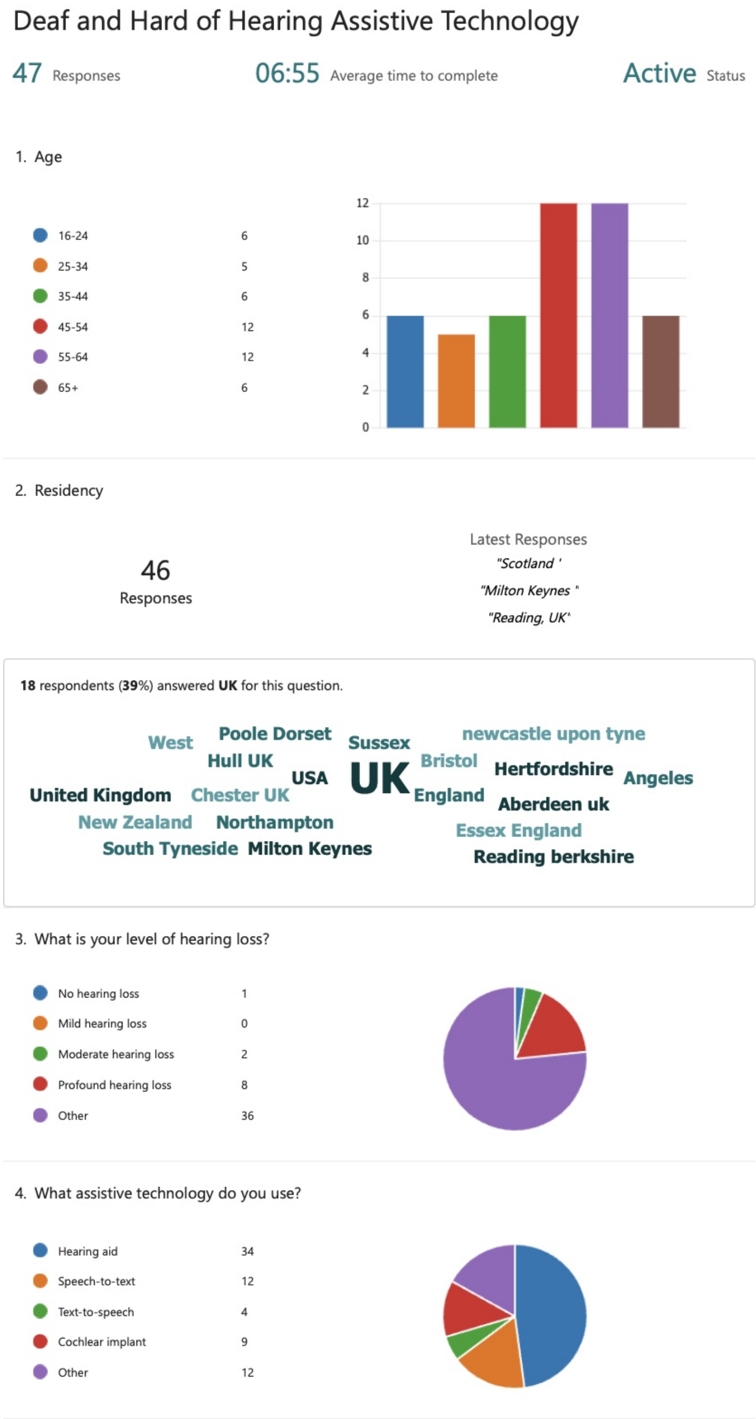


Figure A.1: This screenshots for the responses to the survey outlined in Section 6.1.

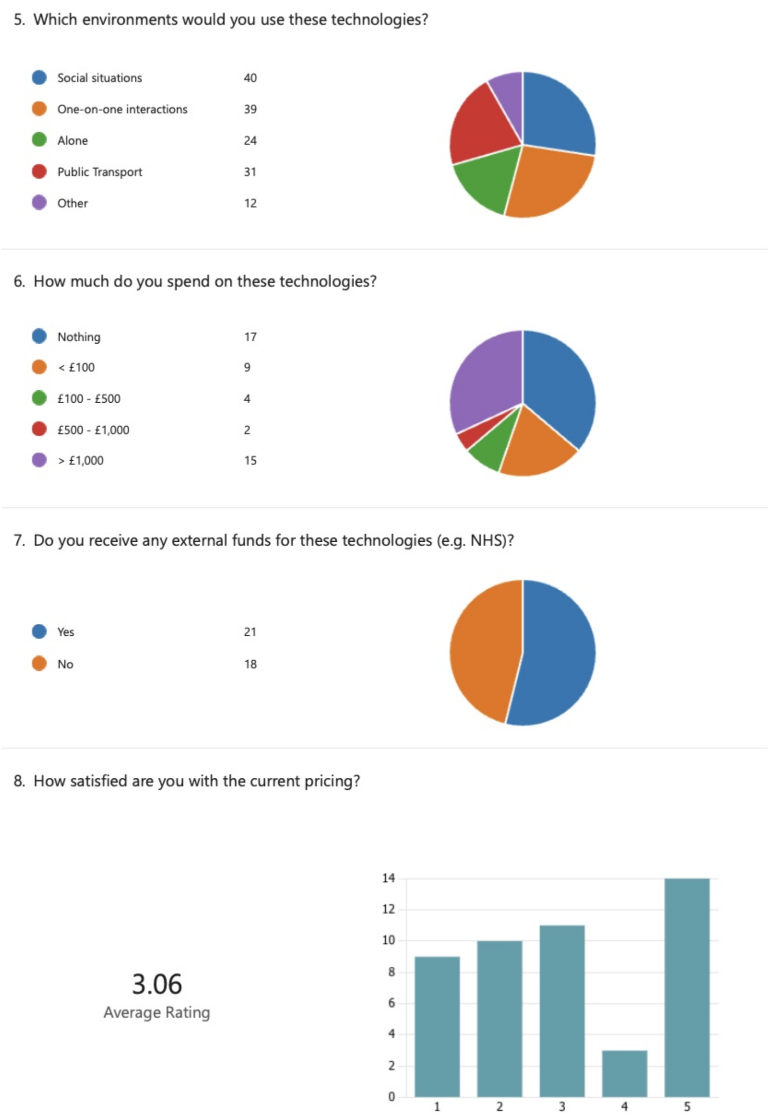


Figure A.2: *This screenshots for the responses to the survey outlined in Section 6.1.*

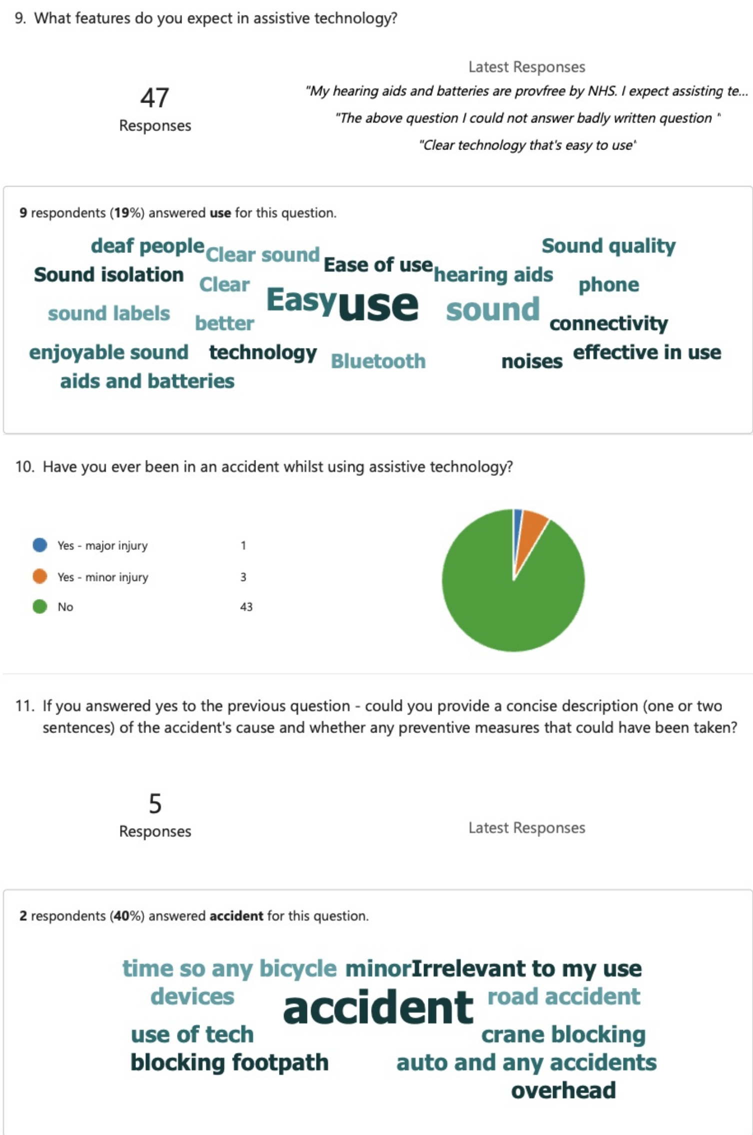


Figure A.3: *This screenshots for the responses to the survey outlined in Section 6.1.*

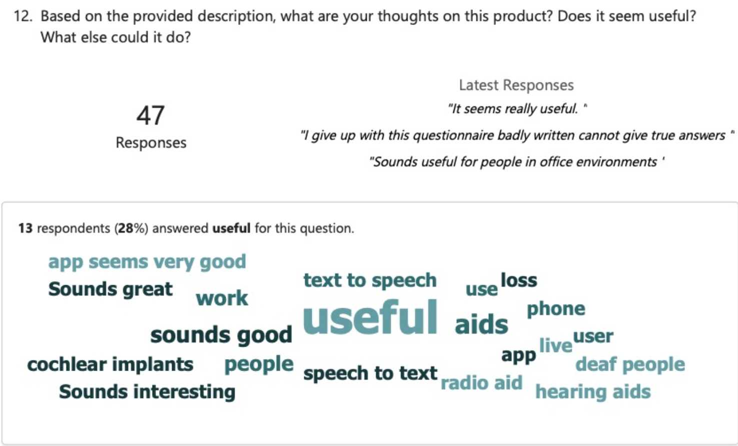


Figure A.4: This screenshots for the responses to the survey outlined in Section 6.1.

Table A.2: This table displays the scores for each question in relation to the UX requirements outlines in Section 4.4, from a scale of 0-4 utilising the same method outlined in Section 7.

Questions	UX.1	UX.2	UX.3	UX.4	UX.5	UX.6	UX.7
1	–	✓	–	–	–	–	–
2	–	✓	–	–	–	–	–
3	–	✓	–	–	–	–	–
4	–	✓	–	–	–	–	–
5	–	–	–	–	✓	–	–
6	–	✓	–	–	–	–	–
7	–	–	–	✓	–	–	–
8	–	–	✓	–	–	–	–
9	–	–	–	✓	–	–	–
10	–	–	–	–	✓	✓	–
11	–	–	–	–	–	–	–
12	–	–	–	–	–	–	–
13	–	✓	–	–	–	–	–
14	–	✓	–	–	–	–	–
15	–	✓	–	–	–	–	–
16	–	✓	–	–	✓	–	–
17	✓	–	–	–	–	–	✓

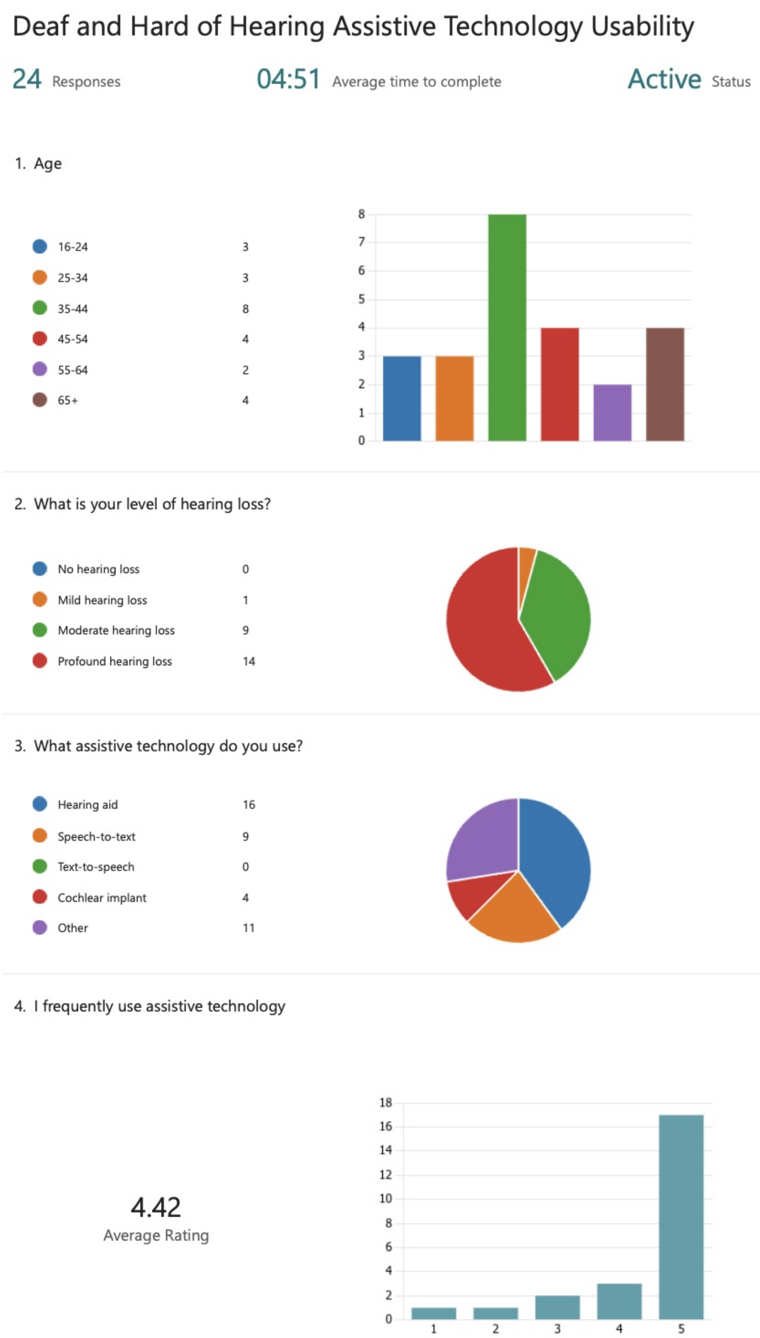
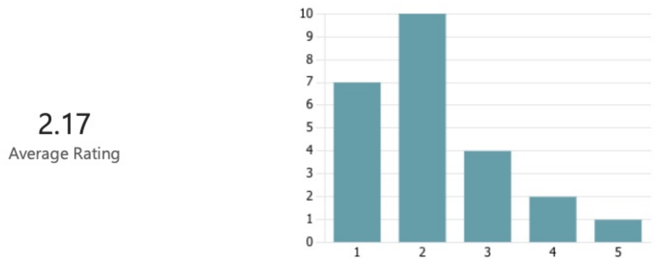
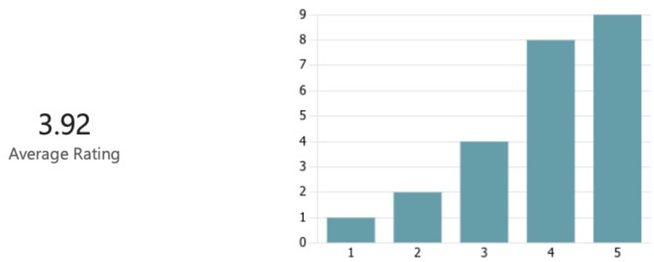


Figure A.5: *This screenshots for the responses to the survey outlined in Section 6.2.*

5. I find assistive technology to be unnecessarily complex



6. I find assistive technology easy to use



7. I often need the support of a technical person to be able to use assistive technology

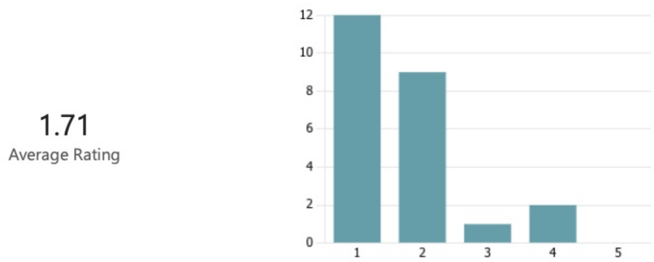
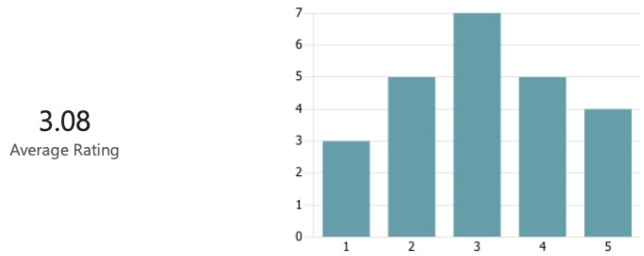
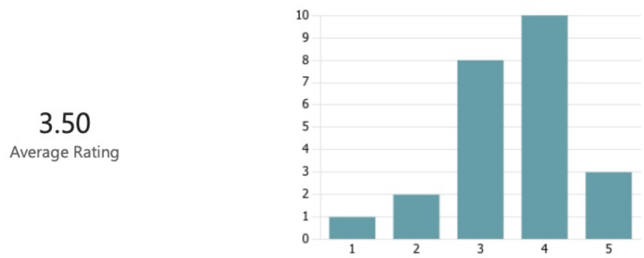


Figure A.6: *This screenshots for the responses to the survey outlined in Section 6.2.*

8. I find the various functions of assistive technology are well integrated



9. Assistive technology offers varying font sizes for text boxes to enhance readability, especially for potential users from the elderly demographic



10. I think there is a lot of inconsistency with assistive technology

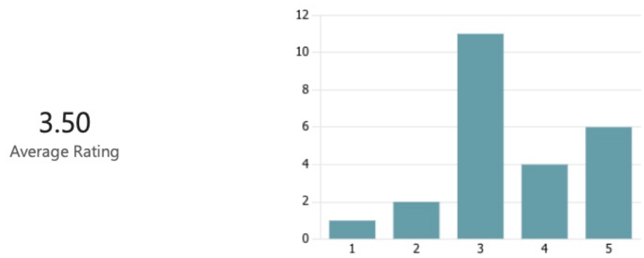


Figure A.7: *This screenshots for the responses to the survey outlined in Section 6.2.*

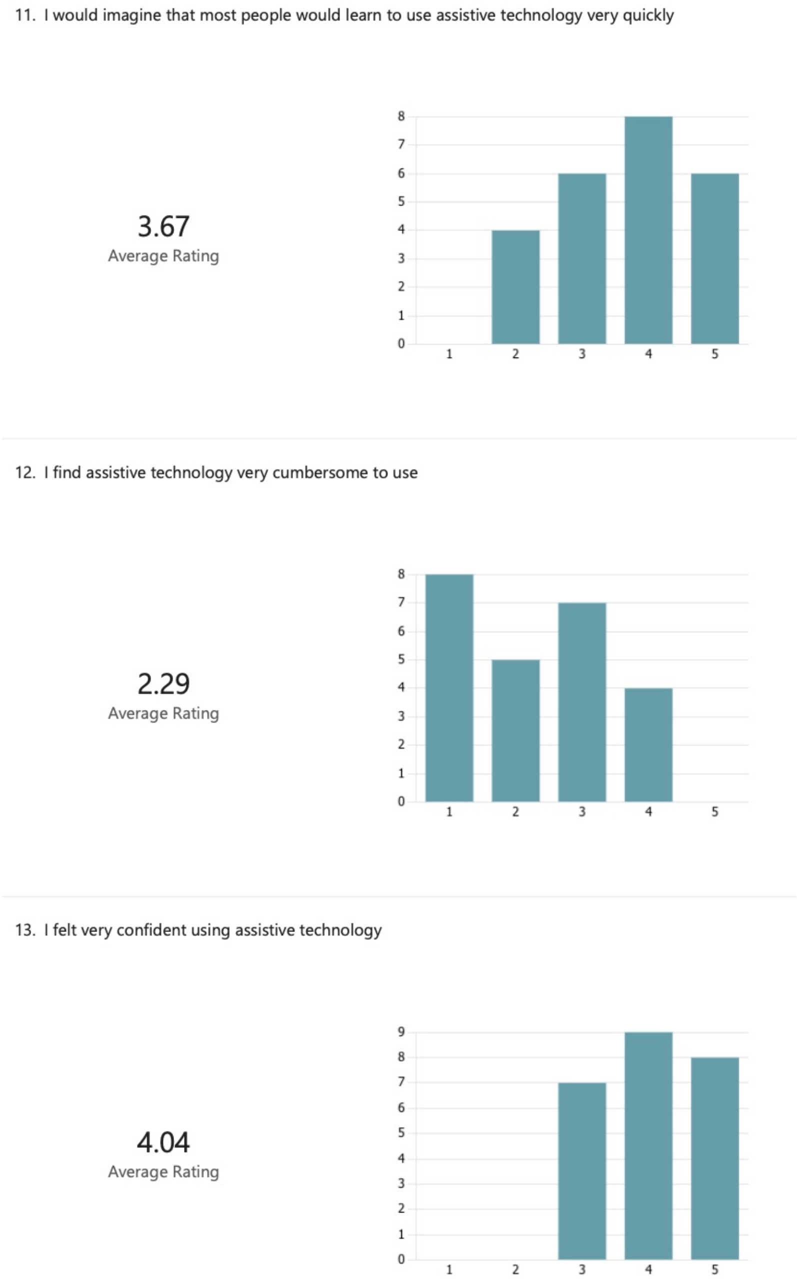
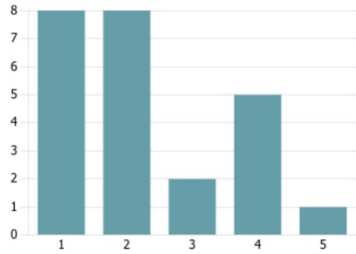


Figure A.8: *This screenshots for the responses to the survey outlined in Section 6.2.*

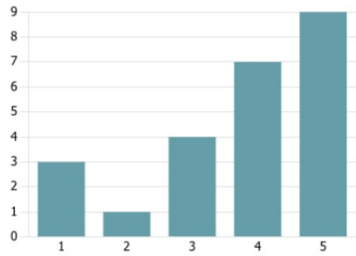
14. I needed to learn a lot of things before I could start using assistive technology

2.29
Average Rating



15. I think current assistive technology is useful in providing solutions for the Deaf and Hard-of-Hearing community

3.75
Average Rating



16. Please provide any other insights regarding assistive technology and their usability.

13
Responses

Latest Responses
"Brands vary alot between useful and boring - Sarabec for example bad ...

Figure A.9: *This screenshots for the responses to the survey outlined in Section 6.2.*

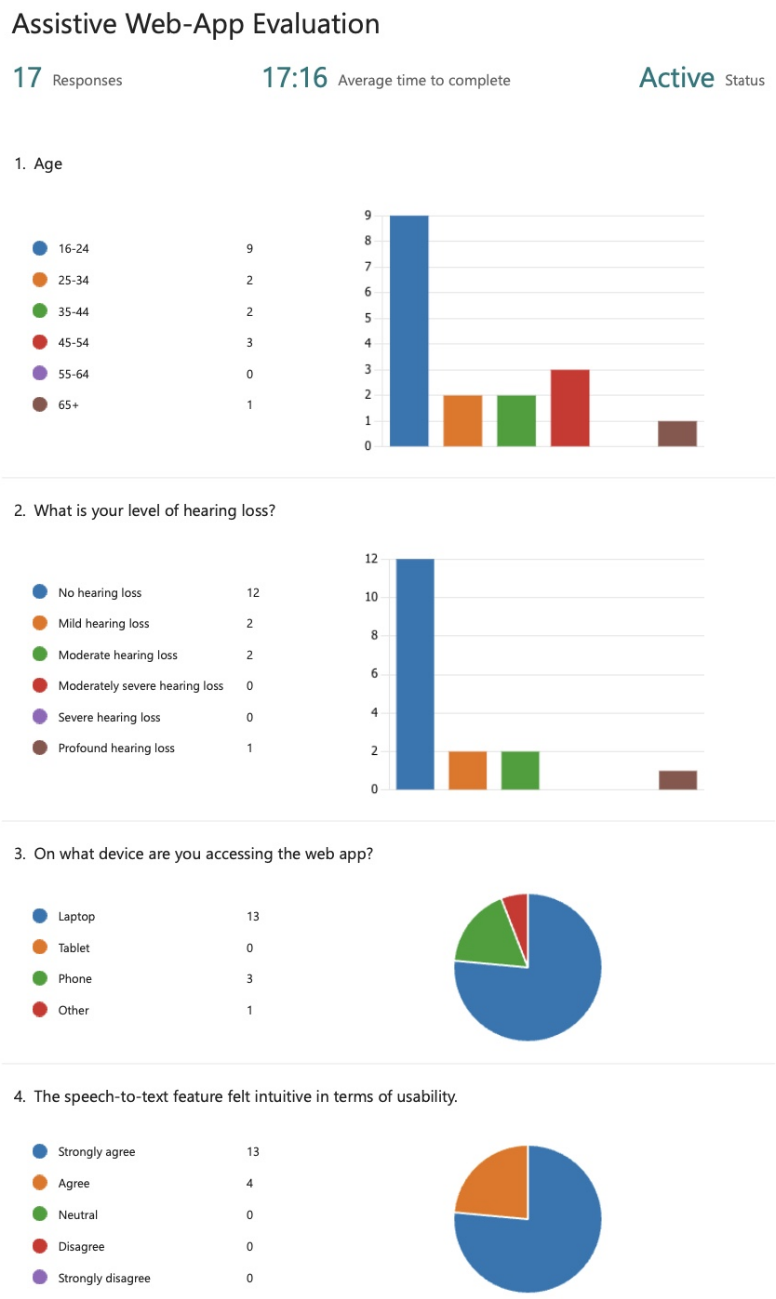


Figure A.10: *This screenshots for the responses to the survey outlined in Section 7.*

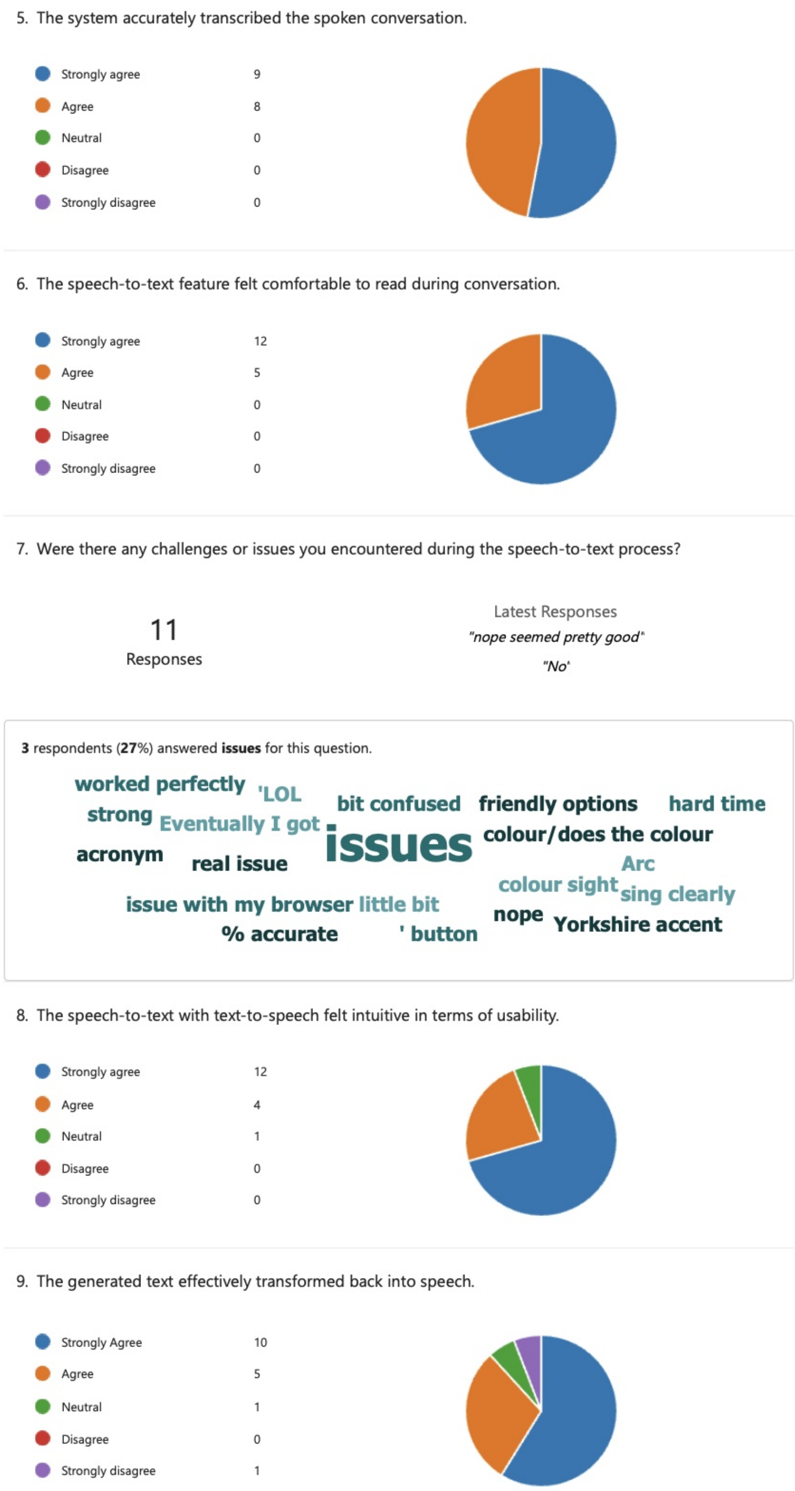


Figure A.11: This screenshots for the responses to the survey outlined in Section 7.

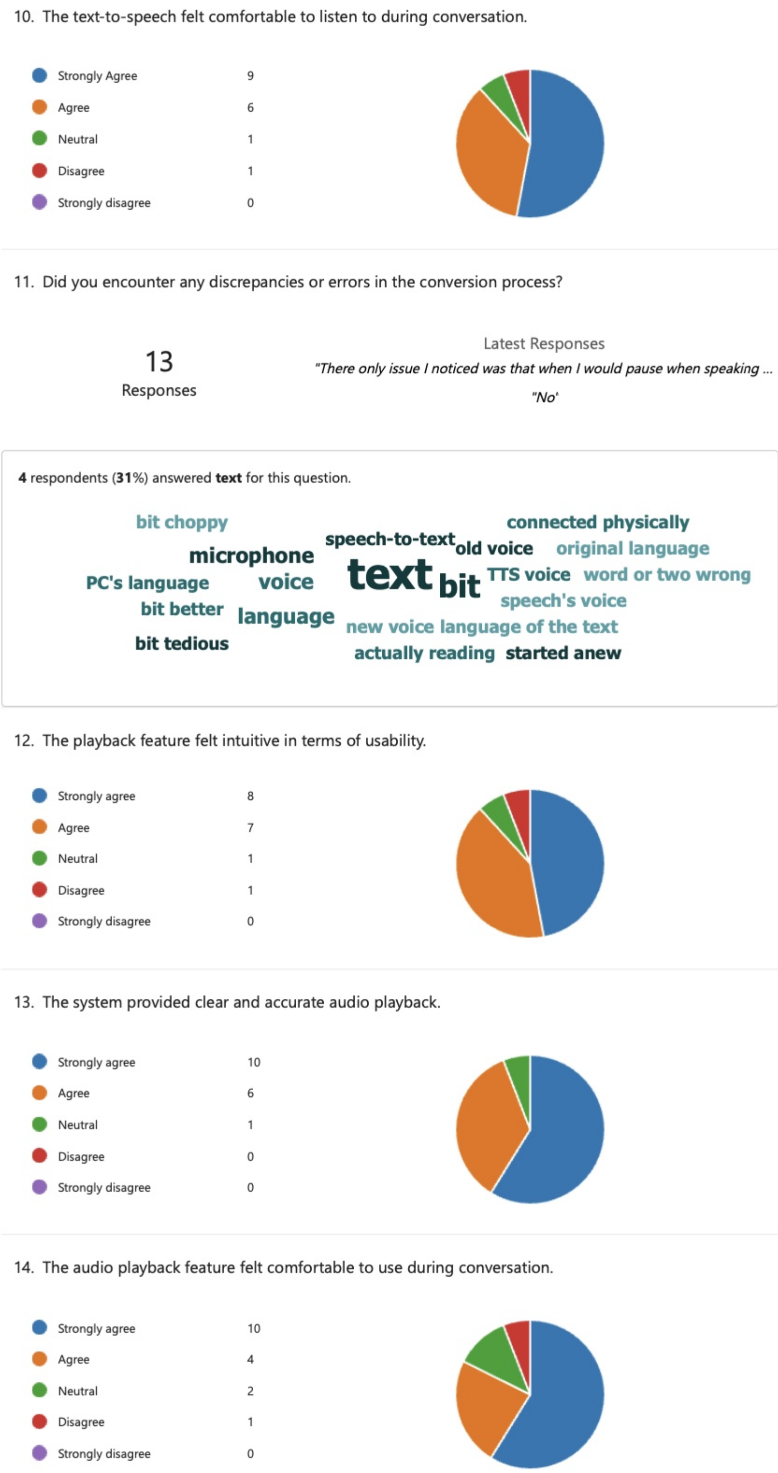


Figure A.12: *This screenshots for the responses to the survey outlined in Section 7.*

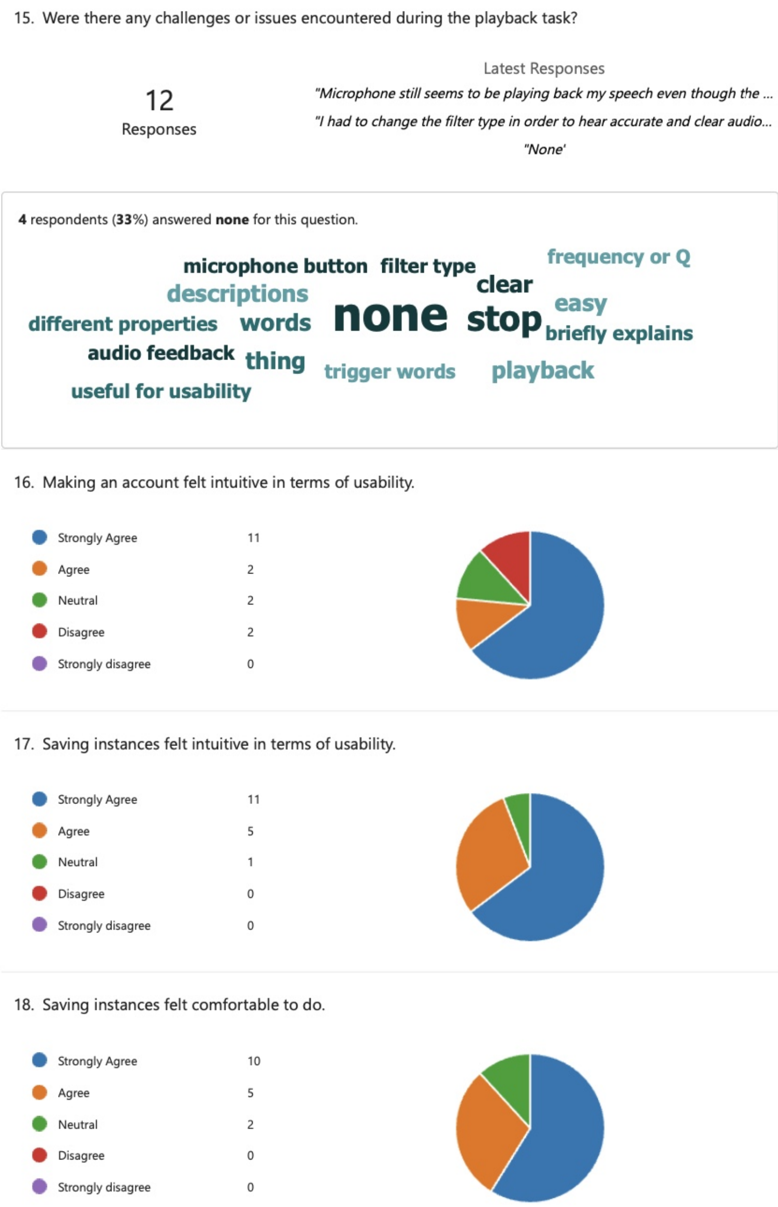


Figure A.13: *This screenshots for the responses to the survey outlined in Section 7.*

19. Were there any challenges or issues encountered during the sign-up and saving task?

9
Responses

Latest Responses
"Saving was easy, but at first I was searching for a "sign up" button befo...
"No'

4 respondents (**44**%) answered **saving** for this question.

20. The text-to-speech felt intuitive in terms of usability.

Strongly agree	15
Agree	1
Disagree	0
Strongly disagree	0
Neutral	1

21. The system produced clear and natural-sounding speech from the input text.

Strongly agree	11
Agree	5
Disagree	0
Strongly disagree	0
Neutral	1

22. The text-to-speech felt comfortable to use during conversation.

Strongly agree	11
Agree	6
Disagree	0
Strongly disagree	0
Neutral	0

Figure A.14: *This screenshots for the responses to the survey outlined in Section 7.*

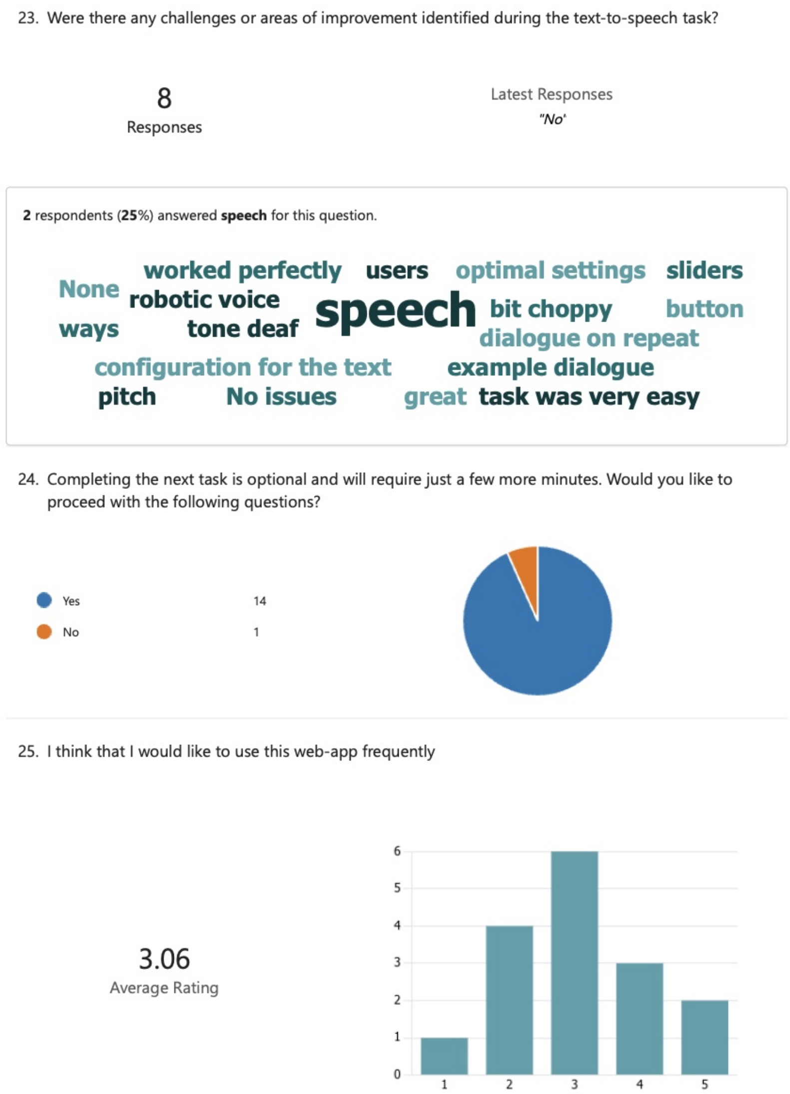
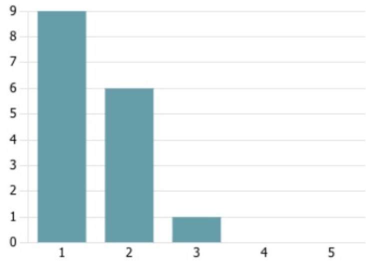


Figure A.15: *This screenshots for the responses to the survey outlined in Section 7.*

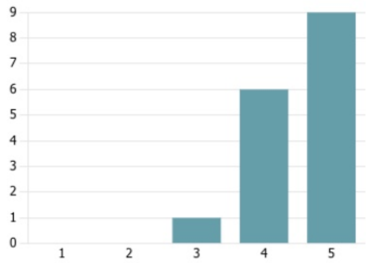
26. I found the web-app unnecessarily complex

1.50
Average Rating



27. I thought the web-app was easy to use

4.50
Average Rating



28. I found the web-app's design to be intuitive and user-friendly

4.25
Average Rating

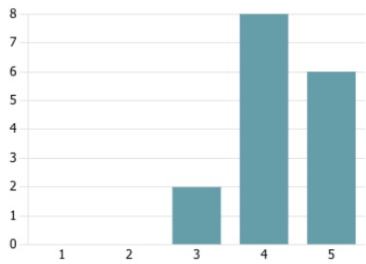
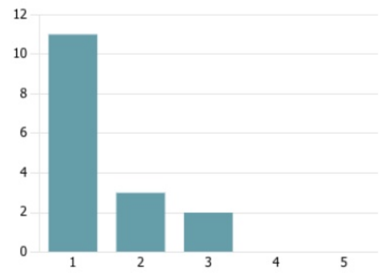


Figure A.16: *This screenshots for the responses to the survey outlined in Section 7.*

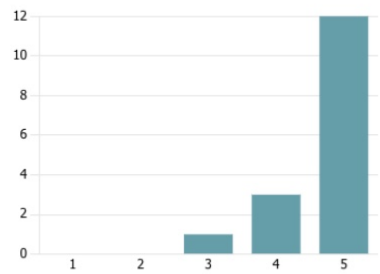
29. I think that I would need the support of a technical person to be able to use this web-app

1.44
Average Rating



30. I found the various functions in this web-app were well integrated

4.69
Average Rating



31. I found the various functions in this web-app were well centralised to direct attention to vital information

4.25
Average Rating

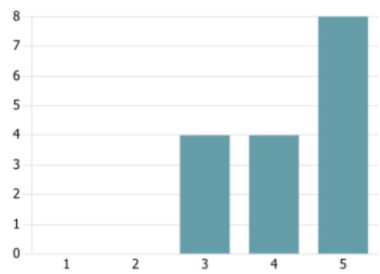
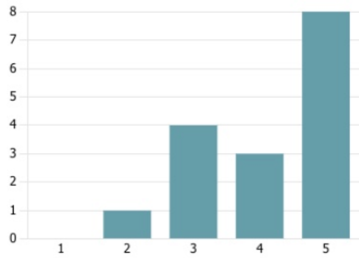


Figure A.17: *This screenshots for the responses to the survey outlined in Section 7.*

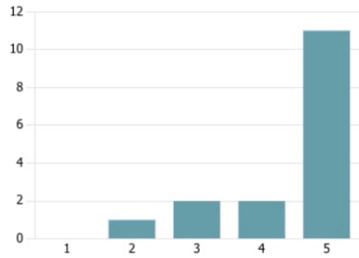
32. I thought the colours of the web-app contributes to a comfortable interaction

4.13
Average Rating



33. I thought the varying font sizes for text boxes enhanced readability, especially for potential users from the elderly demographic

4.44
Average Rating



34. I found the error communications and user guide of the web-app to be an effective user-system communication

4.50
Average Rating

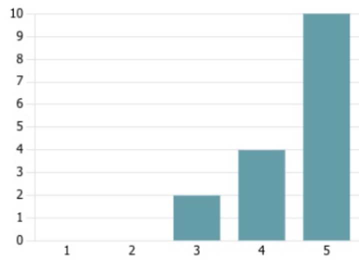
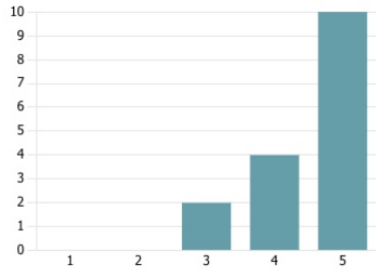


Figure A.18: *This screenshots for the responses to the survey outlined in Section 7.*

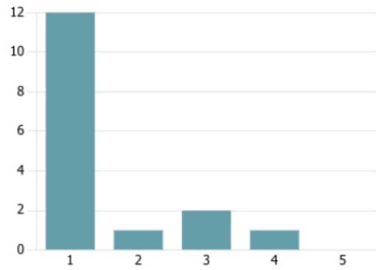
35. I found the default settings helpful in streamlining your initial interaction with the web-app

4.50
Average Rating



36. I thought there was too much inconsistency in this web-app

1.50
Average Rating



37. I would imagine that most people would learn to use this web-app very quickly

4.63
Average Rating

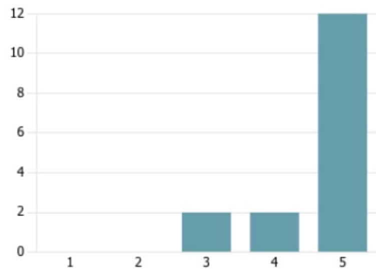
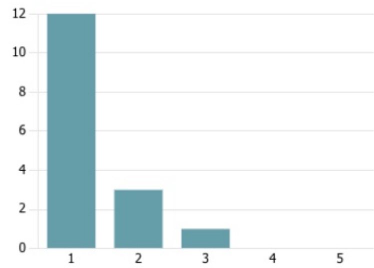


Figure A.19: *This screenshots for the responses to the survey outlined in Section 7.*

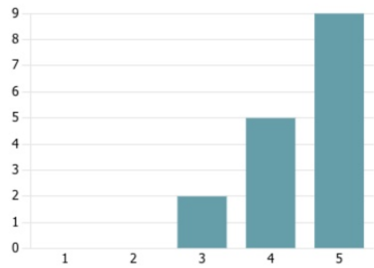
38. I found the web-app very cumbersome to use

1.31
Average Rating



39. I felt very confident using the web-app

4.44
Average Rating



40. I needed to learn a lot of things before I could get going with the web-app

1.75
Average Rating

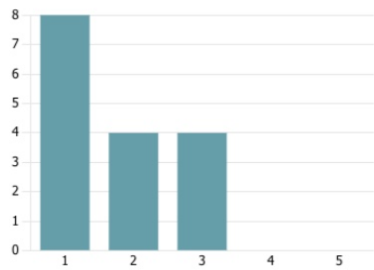


Figure A.20: *This screenshots for the responses to the survey outlined in Section 7.*

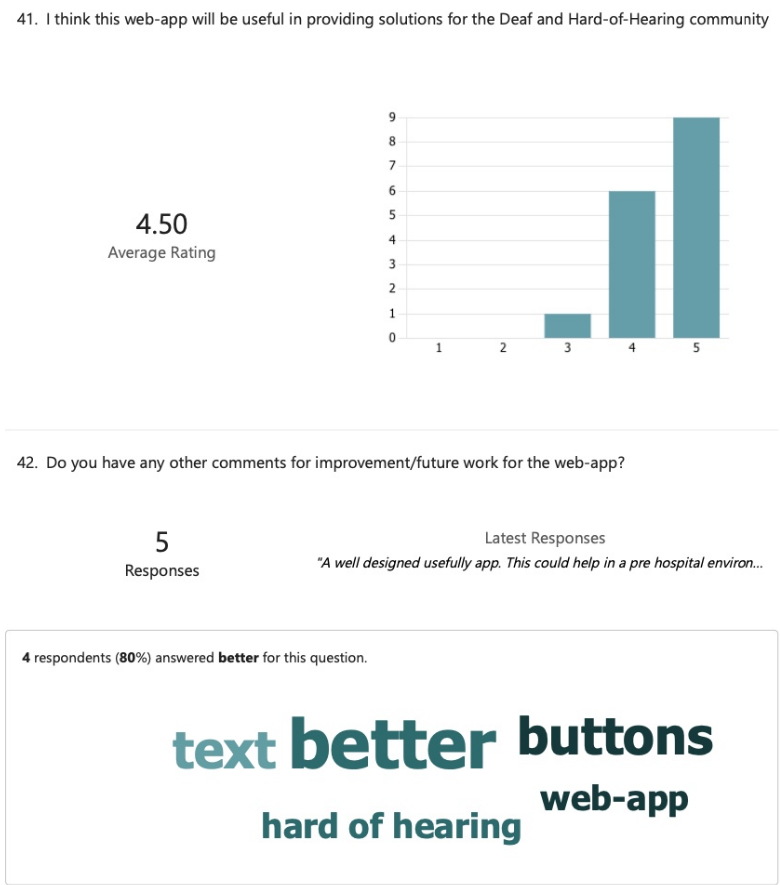


Figure A.21: *This screenshots for the responses to the survey outlined in Section 7.*

8 Bibliography

A. Adeel, J. Ahmad, H. Larijani, and A. Hussain. A novel real-time, lightweight chaotic-encryption scheme for next-generation audio-visual hearing aids. *Cognitive Computation*, 12, 05 2020. doi: 10.1007/s12559-019-09653-z. Last accessed: 2024-02-02.

Atlassian. What is a code review how it can save time. URL `https://www.atlassian.com/agile/software-development/code-reviews`. Last accessed: 2024-03-18.

J. Borg, A. Lindström, and S. Larsson. Assistive technology in developing countries: a review from the perspective of the convention on the rights of persons with disabilities. *Prosthetics and Orthotics International*, 35(1), 2011. ISSN 0309-3646. URL `https://journals.lww.com/poijournal/fulltext/2011/35010/assistive_technology_in_developing_countries__a.3.aspx`. Last accessed: 2023-12-10.

E. D. Borre, K. Kaalund, N. Frisco, G. Zhang, A. Ayer, M. Kelly-Hedrick, S. D. Reed, S. D. Emmett, H. Francis, D. L. Tucci, et al. The impact of hearing loss and its treatment on health-related quality of life utility: a systematic review with meta-analysis. *Journal of General Internal Medicine*, 38(2):456–479, 2023. Last accessed: 2024-01-11.

E. Britannica. Deafness, January 2024. URL `https://www.britannica.com/science/deafness`. Last accessed: 2024-01-08.

P. Broadbent, R. Thomson, D. Kopasker, G. McCartney, P. Meier, M. Richiardi, M. McKee, and S. V. Katikireddi. The public health implications of the cost-of-living crisis: outlining mechanisms and modelling consequences. *The Lancet Regional Health–Europe*, 27, 2023. Last accessed: 2024-01-31.

J. Brooke. Sus: A quick and dirty usability scale. *Usability Eval. Ind.*, 189, 11 1995. Last accessed: 2024-02-06.

K. Burrows. A not-so-short history of deaf technology. 2022. Last accessed: 2024-01-06.

CalHearing. A brief history of hearing loss, Jan 2023. URL `https://calhearing.com/a-brief-history-of-hearing-loss/`. Last accessed: 2024-01-08.

J. S. Choi, M. E. Adams, E. M. Crimmins, F. R. Lin, and J. A. Ailshire. Association between hearing aid use and mortality in adults with hearing loss in the usa: a mortality follow-up study of a cross-sectional cohort. *The Lancet Healthy Longevity*, 5(1):e66–e75, 2024. Last accessed: 2023-01-11.

L. F. Chu and B. K. Chan. Evolution of web site design: implications for medical education on the internet. *Computers in biology and medicine*, 28(5):459–472, 1998. Last accessed: 2024-01-25.

Clerk. URL `https://clerk.com/docs/quickstarts/react`. Last accessed: 2024-01-31.

A. B. Consortium. Chapter 10: Moscow prioritisation. URL `https://www.agilebusiness.org/dsdm-project-framework/moscow-prioririsation.html`. Last accessed: 2024-01-14.

E. David and O. Selfridge. Eyes and ears for computers. *Proceedings of the IRE*, 50(5):1093–1101, 1962. Last accessed: 2024-01-16.

A. Davis, P. Smith, M. Ferguson, D. Stephens, and I. Gianopoulos. Acceptability, benefit and costs of early screening for hearing disability: A study of potential screening tests and models. *Health technology assessment (Winchester, England)*, 11:1–294, 10 2007. doi: 10.3310/hta11420. Last accessed: 2023-12-14.

M. W. Docs. Web speech api - web apis: Mdn, Feb 2023. URL `https://developer.mozilla.org/en-US/docs/Web/API/Web_Speech_API`. Last accessed: 2024-01-26.

M. A. Ferguson, P. T. Kitterick, L. Y. Chong, M. Edmondson-Jones, F. Barker, and D. J. Hoare. Hearing aids for mild to moderate hearing loss in adults. *Cochrane Database of Systematic Reviews*, (9), 2017. Last accessed: 2023-12-14.

M. R. Firm. Hearing aids market worth $10.2 billion by 2026. URL `https://www.marketsandmarkets.com/PressReleases/hearing-aids.asp`. Last accessed: 2023-12-08.

Google. URL `https://support.google.com/accessibility/android/#topic=9078842`. Last accessed: 2024-02-02.

D. Grant. Full-featured user authentication made easy with clerk - this dot labs, Jun 2023. URL `https://www.thisdot.co/blog/full-featured-user-authentication-made-easy-with-clerk`. Last accessed: 2024-01-31.

N. E. Groce. deafness on martha's vineyard, November 2016. URL `https://www.britannica.com/science/deafness-on-Marthas-Vineyard`. Last accessed: 2024-01-08.

L. M. Haile, K. Kamenov, P. S. Briant, A. U. Orji, J. D. Steinmetz, A. Abdoli, M. Abdollahi, E. Abu-Gharbieh, A. Afshin, H. Ahmed, et al. Hearing loss prevalence and years lived with disability, 1990–2019: findings from the global burden of disease study 2019. *The Lancet*, 397 (10278):996–1009, 2021. Last accessed: 2024-01-05.

S. E. Harpe. How to analyze likert and other rating scale data. *Currents in pharmacy teaching and learning*, 7(6):836–850, 2015. Last accessed: 2024-03-18.

P. Harrison. Hearing aid prices in the uk 2024, Jan 2024. URL `https://www.hearingaid.org.uk/hearing-aids/hearing-aid-prices-uk`. Last accessed: 2024-01-11.

E. Jiménez-Arberas and E. Díez. Psychosocial impact of assistive devices and other technologies on deaf and hard of hearing people. *International Journal of Environmental Research and Public Health*, 18(14), 2021. ISSN 1660-4601. doi: 10.3390/ijerph18147259. URL `https://www.mdpi.com/1660-4601/18/14/7259`. Last accessed: 2023-11-24.

D. Jurafsky and J. H. Martin. Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition. Last accessed: 2024-01-16.

D. H. Klatt. Software for a cascade/parallel formant synthesizer. *the Journal of the Acoustical Society of America*, 67(3):971–995, 1980. Last accessed: 2024-02-23.

A. Koenecke, A. Nam, E. Lake, J. Nudell, M. Quartey, Z. Mengesha, C. Toups, J. R. Rickford, D. Jurafsky, and S. Goel. Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences*, 117(14):7684–7689, 2020. Last accessed: 2024-01-16.

J. L. Kröger and P. Raschke. Is my phone listening in? on the feasibility and detectability of mobile eavesdropping. In *Data and Applications Security and Privacy XXXIII: 33rd Annual IFIP WG 11.3 Conference, DBSec 2019, Charleston, SC, USA, July 15–17, 2019, Proceedings 33*, pages 102–120. Springer, 2019. Last accessed: 2024-03-04.

B. J. Lawrence, D. M. Jayakody, R. J. Bennett, R. H. Eikelboom, N. Gasson, and P. L. Friedland. Hearing loss and depression in older adults: a systematic review and meta-analysis. *The Gerontologist*, 60(3):e137–e154, 2020. Last accessed: 2024-01-11.

C.-C. Liao and Y.-H. Hsu. Colour-emotion association and colour preference for elder care robot appearance. *Coloration Technology*, 139(4):369–384, 2023. doi: https://doi.org/10.1111/cote.12653. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/cote.12653. Last accessed: 2023-10-11.

F. R. Lin, E. J. Metter, R. J. O'Brien, S. M. Resnick, A. B. Zonderman, and L. Ferrucci. Hearing Loss and Incident Dementia. *Archives of Neurology*, 68(2):214–220, 02 2011. ISSN 0003-9942. doi: 10.1001/archneurol.2010.362. URL https://doi.org/10.1001/archneurol.2010.362. Last accessed: 2023-12-14.

G. Livingston, J. Huntley, A. Sommerlad, D. Ames, C. Ballard, S. Banerjee, C. Brayne, A. Burns, J. Cohen-Mansfield, C. Cooper, et al. Dementia prevention, intervention, and care: 2020 report of the lancet commission. *The Lancet*, 396(10248):413–446, 2020. Last accessed: 2024-01-05.

C. Lopez-Lloreda. How speech-recognition software discriminates against minority voices, Oct 2020. URL https://www.scientificamerican.com/article/how-speech-recognition-software-discriminates-against-minority-voices/. Last accessed: 2024-01-16.

D. G. Loughrey, M. E. Kelly, G. A. Kelley, S. Brennan, and B. A. Lawlor. Association of age-related hearing loss with cognitive function, cognitive impairment, and dementia: a systematic review and meta-analysis. *JAMA otolaryngology–head & neck surgery*, 144(2):115–126, 2018. Last accessed: 2024-01-11.

E. Mahmoudi, P. Zazove, M. Meade, and M. M. McKee. Association Between Hearing Aid Use and Health Care Use and Cost Among Older Adults With Hearing Loss. *JAMA Otolaryngology–Head Neck Surgery*, 144(6):498–505, 06 2018. ISSN 2168-6181. doi: 10.1001/jamaoto.2018.0273. URL https://doi.org/10.1001/jamaoto.2018.0273. Last accessed: 2024-01-08.

C. Mathers, A. Smith, and M. Concha. Global burden of hearing loss in the year 2000. *Global burden of Disease*, 18(4):1–30, 2000. Last accessed: 2023-12-14.

P. Morville. User experience design, Jun 2004. URL https://semanticstudios.com/user_experience_design/. Last accessed: 2023-10-11.

MozDevNet. Web audio api - web apis: Mdn. URL https://developer.mozilla.org/en-US/docs/Web/API/Web_Audio_API. Last accessed: 2024-02-04.

MozDevNet. Biquadfilternode - web apis: Mdn, Apr 2023. URL https://developer.mozilla.org/en-US/docs/Web/API/BiquadFilterNode. Last accessed: 2024-03-03.

U. Nations. Convention on the rights of persons with disabilities (crpd). URL https://social.desa.un.org/issues/disability/crpd/convention-on-the-rights-of-persons-with-disabilities-crpd. Last accessed: 2023-12-10.

NIDCD, Oct 2022. URL `https://www.nidcd.nih.gov/health/hearing-aids`. Last accessed: 2024-01-11.

J. Pornel and G. Saldaña. Four common misuses of the likert scale. *Philippine Journal of Social Sciences and Humanities*, 18:12–19, 12 2013. Last accessed: 2024-03-18.

D. Powell, E. Oh, F. Lin, and J. Deal. Hearing impairment and cognition in an aging world. *Journal of the Association for Research in Otolaryngology : JARO*, 22, 05 2021. doi: 10.1007/s10162-021-00799-y. Last accessed: 2023-10-04.

PyPI. vosk. URL `https://pypi.org/project/vosk/`. Last accessed: 2024-03-14.

Red-Hat, May 2020. URL `https://www.redhat.com/en/topics/api/what-is-a-rest-api`. Last accessed: 2024-01-08.

Samsung, Sep 2023. URL `https://www.samsung.com/uk/sustainability/accessibility/hearing/`. Last accessed: 2024-02-02.

M. J. Scherer. Outcomes of assistive technology use on quality of life. *Disability and rehabilitation*, 18(9):439–448, 1996. Last accessed: 2023-11-24.

H. Semary, K. A. Al-Karawi, and M. M. Abdelwahab. Using voice technologies to support disabled people. *Journal of Disability Research*, 3(1), 2024. Last accessed: 2024-02-23.

A. A. Shah, A. Adeel, J. Ahmad, A. Al-Dubai, M. Gogate, A. Bishnu, M. Diyan, T. Hussain, K. Dashtipour, T. Ratnarajah, and A. Hussain. A novel chaos-based light-weight image encryption scheme for multi-modal hearing aids, 2022. Last accessed: 2023-12-10.

A. M. Sheffield and R. J. Smith. The epidemiology of deafness. *Cold Spring Harbor perspectives in medicine*, 9(9):a033258, 2019. Last accessed: 2024-03-17.

Y. Song, Y. Yang, and P. Cheng. The investigation of adoption of voice-user interface (vui) in smart home systems among chinese older adults. *Sensors*, 22(4):1614, 2022. Last accessed: 2024-01-28.

A. Support. URL `https://support.apple.com/en-gb/guide/iphone/iphb210be03a/ios`. Last accessed: 2024-02-02.

X. Tan, J. Chen, H. Liu, J. Cong, C. Zhang, Y. Liu, X. Wang, Y. Leng, Y. Yi, L. He, et al. Naturalspeech: End-to-end text-to-speech synthesis with human-level quality. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024. Last accessed: 2024-02-23.

V. Tangcharoensathien, W. Witthayapipopsakul, S. Viriyathorn, and W. Patcharanarumol. Improving access to assistive technologies: challenges and solutions in low- and middle-income countries. *WHO South East Asia J Public Health*, 7(2):84–89, Sept. 2018. Last accessed: 2023-12-10.

G. Tao, G. Charm, K. Kabacińska, W. C. Miller, and J. M. Robillard. Evaluation tools for assistive technologies: a scoping review. *Archives of Physical Medicine and Rehabilitation*, 101(6): 1025–1040, 2020. Last accessed: 2024-01-14.

Usability.gov, May 2014a. URL `https://www.usability.gov/what-and-why/user-interface-design.html#:~:text=Create%20consistency%20and%20use%20common,site%20to%20help%20facilitate%20efficiency`. Last accessed: 2023-11-24.

Usability.gov, Feb 2014b. URL `https://www.usability.gov/what-and-why/user-experience.html#:~:text=User%20experience%20(UX)%20focuses%20on,the%20group%20managing%20the%20project`. Last accessed: 2023-11-24.

M. E. Valentinuzzi. Hearing aid history: From ear trumpets to digital technology. *IEEE Pulse*, 11(5):33–36, 2020. doi: 10.1109/MPULS.2020.3023833. Last accessed: 2024-01-05.

WHO. Deafness and hearing loss, Feb 2023. URL `https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss`. Last accessed: 2023-12-10.

P. Wordmark. Eagle speaker recognition and identification. URL `https://picovoice.ai/platform/eagle/`. Last accessed: 2024-03-13.