Sounds Accessible: Envisioning Accessible Audio-Media Futures with People with Aphasia

Filip Bircanin
Department of Informatics
King's College London
London, United Kingdom
filip.bircanin@kcl.ac.uk

Alexandre Nevsky
Department of Informatics
King's College London
London, United Kingdom
alexandre.nevsky@kcl.ac.uk

Himaya Perera
Department of Informatics
King's College London
London, United Kingdom
himaya.perera@kcl.ac.uk

Vaasvi Agarwal
Department of Informatics
King's College London
London, United Kingdom
vaasvi.agarwal@kcl.ac.uk

Eunyeol Song
Department of Informatics
King's College London
London, United Kingdom
eunyeol.song@kcl.ac.uk

Madeline Cruice
Department of Language and Communication Science
City St George's, University of London
London, United Kingdom
m.cruice@city.ac.uk

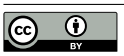
Timothy Neate
Department of Informatics
King's College London
London, United Kingdom
timothy.neate@kcl.ac.uk

Abstract

Audio-media, such as radio and podcasts, are a vital means to engage with global events, access education, or offer entertainment. However, for people with complex communication needs, such as aphasia, there can be accessibility challenges. While accessibility research has largely focused on audiovisual media, little work has considered audio-media, particularly for users with complex communication needs. To address this gap, we undertook six co-design workshops with 10 people with aphasia to re-imagine access to audio-media. We uncover how our co-designers perceive audio-media as more than a tool, but a part of daily intimacies; shaping social relationships and contributing to therapeutic recovery. Through a Research-through-Design process culminating in one low-fidelity and three high-fidelity technology probes that embody novel accessibility interventions, our findings further challenge conventional approaches to audio-media accessibility and signal new directions for future design.

CCS Concepts

• **Human-centered computing** → **Accessibility technologies**; *Empirical studies in accessibility*; *Accessibility theory, concepts and paradigms.*



CHI '25, Yokohama, Japan

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3714000

Keywords

Accessibility, audiovisual, media, aphasia, complex communication needs, envisioning, probes, prototype

ACM Reference Format:
Filip Bircanin, Alexandre Nevsky, Himaya Perera, Vaasvi Agarwal, Eunyeol Song, Madeline Cruice, and Timothy Neate. 2025. Sounds Accessible: Envisioning Accessible Audio-Media Futures with People with Aphasia. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 22 pages. https://doi.org/10.1145/3706598.3714000

1 Introduction

Audio-media is integral to the lives of many. A 2024 OFCOM (UK Office for Communication) report [71] highlights its importance – 92% of people in the UK listen to some form of audio content (e.g., radio, podcasts or audiobooks) at least once a week. Despite its ubiquity, accessibility in audio-media remains under-explored, leaving a significant gap in understanding how individuals with complex communication needs (CCNs) engage with auditory experience. Recognising that audio-media services serve public interests and benefit all audiences highlights the need to include people with CCNs in accessibility research and design of accessible interventions. This is crucial not only to support their engagement, but also to challenge the ableist norms of speech fluency, articulation, and comprehension, particularly for people with communication impairments such as aphasia [70] – the community of focus in this work.

Present research on accessibility to digital media focuses largely on the audiovisual – e.g., video media such as TV and streamed video [68]. Generally, these audiovisual accessibility interventions are designed for users with visual and auditory disabilities, such

as captions and audio description respectively. Some recent work, however, has considered the needs of users with CCNs for audiovisual media access [66, 67]. In contrast, research on accessibility in audio-only media remains sparse, focusing primarily on auditory impairments (e.g., [95, 96]). This focus often overlooks a crucial challenges: while accessibility to audio content is often framed in terms of *listening*, it must also address the cognitive and linguistic demands of *understanding*. Ongoing developments – such as internet-delivered content with high potential for individualisation [2, 3] – offer opportunities to bridge this gap and extend accessibility to diverse audiences. However, despite the promise of these advancements, we still lack a clear understanding of how such interventions can support users with CCNs and address their unique challenges.

This paper explores accessible futures for people with aphasia and communication difficulties. Existing technology, as Jones et al. [44] notes, fail to provide efficient access to podcasts, with interventions driven by market goals like shareability and listener expansion, sidelining accessibility. This market focus paradoxically limits user choice, shifting audio-media from a sanctuary of critical engagement [34] to a tool for audience profiling and customisation. Audio media streaming platforms like BBC Sounds are shifting toward personalization and immersive content [12], yet remain overwhelming for some users, such as those with dyslexia [20], emphasising the need for accessibility that prioritizes inclusion over marketability.

In this paper, we shift the focus to interventions that improve access to audio-media for people living with the language impairment aphasia, offering a novel perspective on how people with CCNs engage with and experience audio-media. Focusing on the experiential dimensions of audio-media, we aim to uncover how people with aphasia listen to and interact with these technologies in everyday life, highlighting the unique challenges and opportunities that arise in the context of audio-only formats. We address the following key research questions: (1) *How can interventions be envisioned to enhance the engagement and experience of audio-media for individuals living with aphasia, considering the community's unique communication needs and the experiential dimensions of listening?* and (2) *What form should these interventions take to effectively integrate with the everyday lives of people with aphasia?*. Additionally, we explore the social dynamics of audio-media consumption and how these envisioned interventions might influence and shape this experience.

Our work contributes to a broader understanding of how people with aphasia experience audio-media, providing insights into the barriers and opportunities that arise in these auditory interactions. Through this lens, we envision new roles and functions for audio-media, proposing design directions that extend beyond traditional accessibility interventions to embrace the full spectrum of auditory experiences. As the first study of its kind, it addresses the notable lack of research on audio-media use among people with aphasia – and more broadly, those with CCNs – broadening the scope of media accessibility beyond narrow interventions like captions. These interventions, which consider audio-media in general, and in particular on a long-spoken dialogue, aim to shape how the accessibility community's focus on audio-media in relation to diverse communication needs. To this end, we offer the following contributions:

(1) A detailed exploration of the needs, experiences, and barriers for people living with aphasia in the context of audio-media consumption, providing new understandings of how these individuals engage with and perceive conventional and emerging audio technologies.
(2) An exploration of new roles and functions for audio-media, highlighting the potential for our proposed tailored accessibility interventions that align with the unique communication needs of people with aphasia.
(3) Implications for design, offering insights and recommendations on how to effectively create and implement accessibility interventions for audio-media that are both meaningful and accessible to individuals with aphasia, ensuring their full participation in auditory experiences.

Through this research, we not only address the practicalities of designing for accessibility, but also describe three key directions for transforming audio-media: multi-representational enhancement (e.g., adding visual or textual aids), reformatting content to offload long-spoken formats (e.g., summarising or segmenting), and creating entirely new content tailored to specific needs, such as aphasia-friendly formats. Our exploratory approach emphasises the breadth of possibilities, envisioning new research directions, and encouraging deeper exploration of scenarios where these interventions can enrich and redefine the auditory landscape.

2 Foregrounding Accessible Audiovisual Media Research in HCI

Screen-based media, audio-only media and social media have emerged as vital means of participation in contemporary life; offering platforms that extend far beyond mere communication. They form social threads that weave individuals into the fabric of the community; fostering a sense of belonging and shared experience [99]. In the civic sphere, these media forms are not just informative but also *transformative* – equipping citizens with the tools to engage with the public discourse and cultivate a more active and informed citizenry [59, 101]. Culturally, audiovisual media serves a dual role; reflecting the diverse societal values while simultaneously shaping and preserving the collective ideas, traditions and creative expressions [69].

Over the last decade, there has been significant growth in accessibility research for audiovisual content resulting in new broadcasting guidelines designed to enhance access for all viewers [82]. Historically, the discipline of Audiovisual Translation has dedicated considerable effort to addressing accessibility, with a substantial body of work focused on subtitling, dubbing, adaptations, voice-overs, audio descriptions, sign interpreting, and respeaking [22]. This focus on accessibility has naturally extended into the field of Human-Computer Interaction (HCI), where research and design efforts have primarily concentrated on technological solutions that include caption accessibility [49, 73, 80] and audio descriptions [46, 72]. While Mack et al. [56] reported a growing trend of accessibility research in the HCI community, the trend within audiovisual HCI research has remained mostly focused on users with hearing and visual disabilities [68].

Although much digital content consumption has shifted to more technologically capable platforms – e.g., TikTok [33, 81] or live-streaming [40] – much of the existing research on audiovisual accessibility interventions remain centred around 'conventional' screen viewing [68]; 'TV' and its standard access supports (i.e., captions and audio description). The recent rise of content distributed through the internet – and the high customisability this implies – might enable novel accessibility interventions which meet any user's needs. Further, advances in deep learning approaches and large language models (LLMs) might enable accessibility interventions which support the needs of users with CCNs, such as through simplified summarisation of video content [35], or through commercial-level summarisation capabilities [1].

Finally, audio-media, even in the context of significant growth in podcasting and audiobooks, lags behind audiovisual media in terms of accessibility advancements [32]. Currently, audio access is often treated as an isolated component within audiovisual interventions, or treated as access to non-digital contexts and environments (i.e., [43]), and with only limited efforts made to address the unique needs of disabled audiences in audio-only formats.

2.1 Audio-media Accessibility

While radio continues to maintain its relevance in an era dominated by audiovisual media [97], its enduring presence is increasingly tangled with the resurgence of "listening" through its audiovisual components and the rise of new internet formats. Traditionally considered distinct from audiovisual media research and design, sound media – encompassing radio, podcasts, audiobooks and similar, fast-growing formats [71] – have undergone a transformation, acquiring a new materiality that blurs the lines between sound and vision [13]. What was once a purely auditory experience now often includes visual elements. This evolution is exemplified by BBC 1 Radio, which in 2014 reached one million subscribers on YouTube, with cameras becoming a regular feature in radio studios. Common streaming platforms now offer some visual element – e.g., Spotify's Canvas[2]; short rolling videos which offer a visual complement to music. This shift has initiated interest among researchers who are beginning to explore the accessibility of these transformed media landscapes [44].

To improve access to audio-media researchers have introduced various options for neurotypical audiences [44]; from podcast properties such as reliable and source-rich metadata [23, 32], acoustic features [100] and novel trajectory-based recommendations [11]. For example, the growing use of LLMs has further placed emphasis on summarisation, yet for non-accessibility purposes, such as to compose informative and catchy descriptions in text – i.e., [75], audio [44], or video format [94]. Despite this growing attention, much of the research remains focused on neurotypical users [21, 97], focusing largely on brand strategies and consumer interests. Little engagement is directed toward understanding and improving accessibility for diverse audiences.

Conventional radio, radio-on-demand, and podcasting – collectively termed 'soundwork' [12] – have primarily been considered by accessibility research through the lens of hard-of-hearing audiences. Within this stream of research, there are two main considerations. First, the topic of audio enhancement, often explored under the umbrella of audio personalisation, has predominantly focused on improving audio quality and accessibility for hard-of-hearing audiences [86, 95]. This, relatively small, area of research has predominantly concentrated on the possibilities of remixing sounds, with key efforts aimed at allowing listeners to control noise disruptions, enable spatial sound separation, and enhance the clarity of audio. The second stream focuses on information access, improving metadata attributes and textual representation, specifically examining their availability and exploring ways to improve their visual presentation for hard-of-hearing listeners [23, 32]. While this research offers valuable insights, expanding inclusion to a broader spectrum of users with disabilities – particularly those with CCNs – would open up a wealth of unexplored possibilities.

2.2 Designing with People with Aphasia

Aphasia is a language impairment that can affect a person's reading, writing, speaking, and listening abilities, often resulting from a stroke or other brain injury [41, 63]. The nature and severity of aphasia can vary greatly from person to person, which means that individuals living with aphasia may experience the same piece of audio-media in vastly different ways. However, it is important to recognise that aphasia does not impact a person's intelligence, ability to form opinions, or problem-solving abilities, leaving certain cognitive capacities intact [91]. This distinction highlights the importance of developing assistive technologies that cater to the specific communication challenges posed by individuals' aphasia while respecting the cognitive and intellectual capabilities of these individuals.

Designing and developing assistive technologies that genuinely address the accessibility needs of communities living with disabilities requires the direct involvement of those who live with these challenges daily [7, 56, 62]. This is particularly crucial in areas with limited empirical research, such as auditory comprehension in people with aphasia -encompassing acoustic and phonemic processing difficulties [47]. It seems that these challenges may arise from various factors, including for example phoneme misperception, grammatical processing difficulties, and working memory difficulties [92]. As audio media continues to grow as a dominant mode of communication, ensuring its accessibility for individuals with aphasia is essential for equitable access to information and cultural participation.

Participatory design (PD) has been successfully employed across various domains within HCI to involve diverse groups, including people who are deaf-blind [4], older adults [54], and individuals with dementia [53]. Similarly, in the realm of aphasia, co-design has informed the development of high-tech augmentative and alternative communication aids [29, 45], assistive technologies for daily tasks [15, 62], and tools for engaging in artistic digital content creation [65, 89]. Aphasia often co-occurs with other post-stroke impairments, such as motor or cognitive challenges, further complicating recruitment and active engagement in research activities. These complexities necessitate tailored approaches to PD, ensuring that the methods are accessible and meaningful for co-designers

[1] https://summarize.ing for YouTube video summaries
[2] Spotify Canvas: https://artists.spotify.com/en/canvas

with aphasia. Tasks should be concise and straightforward, with verbal instructions supplemented by accessible materials, including text, images, and verbal explanations [39, 83]. Probing for feedback rather than relying on think-aloud methods can reduce cognitive load and prevent fatigue, making the process more inclusive and effective. Using alternatives to standard language-based tasks common in PD can support co-designers in interacting with abstract ideas, such as through the use of tangible design language [98] or representative video clips [54].

3 Methodology

In this section, we outline the overarching Research-through-Design (RtD) methodological approach that guided our project, comprising six workshop sessions. We provide an overview of the workshop structure, illustrating how each session naturally evolved from the previous one, shaped by the group dynamics, co-designer presence, and their evolving ideas.

3.1 Research-through-Design

This paper presents an exploratory study rooted in RtD [87], an approach for envisioning futures and exploring what the world could and should be. RtD combines design-led exploration of materials, scenarios, and concepts with reflective practices that emerge through creating artefacts, services, and systems [36]. Unlike more artefact-centric methods, our use of RtD was inherently flexible and reflexive [10], allowing for the continual re-framing of goals and situations throughout the co-design process [8]. Our methodology was not confined to a single artefact; instead, it evolved in response to the dynamics of the sessions, adapting to the co-designers' insights and needs [102]. This flexibility was key in enabling a truly co-designer driven process – 'drifting by intention' [48] – where the exploration of accessible audio-media interventions was continually informed by the lived experiences of those directly impacted. Leveraging RtD enabled us to generate a wide range of design possibilities, aligning with Zimmerman and Forlizzi [102]'s framing of RtD as producing knowledge that functions as '*proposals* rather than *predictions*'.

3.2 Overview of the Workshop Sessions Guided by Research-through-Design

Co-designers were recruited through Aphasia Re-Connect, an aphasia charity which provides activities and support. Our team was given a regular slot at the charity's weekly drop-in sessions, allowing us to integrate into co-designers' routines. This relaxed environment let co-designers control the pace, taking breaks or pausing until the following week if needed. Of the ten co-designers involved, each session typically had around four, with attendance rotating. To ensure continuity, we updated absent co-designers on developments and began each session with a recap of key topics, presented on a flipchart to support memory retention and understanding.

This section, supported by Table 2 (see Appendix), highlights the dynamic and evolving nature of the six workshop sessions. The first three workshops embraced a broad, exploratory approach, aiming to uncover key challenges and user needs. Rather than adhering to a rigid plan, the sessions co-evolved with the group's dynamics, informed by weekly developments and real-time reflections on emerging insights. Decisions were often made in hindsight, adapting fluidly to the co-designers' priorities and the organic flow of discussions. Given the unexplored nature of this area, the rapid pace of the workshops, and the use of emerging technologies in our workshops (e.g., GenAI – for more detail see Section 4), we simultaneously explored and assessed the value of the interventions as they were being designed. This iterative, reflexive process generated actionable insights – generative knowledge – that addressed immediate accessibility needs while laying the groundwork for future design directions.

Sessions were facilitated by two researchers experienced working with CCNs, with input from a speech and language therapist (SLT) at key points. As shown in Figure 1, the first three sessions, following the double diamond model [31], focused on divergent thinking to generate a wide range of intervention ideas, including imaginative and 'magic' solutions [27, 30], with a shift toward convergence starting in Session 3. In the fourth session, we introduced three high-fidelity technological probes, refining the discussion to more realistic and actionable design opportunities. By the fifth session, we presented a tangible low-fidelity radio mock-up (see Figure 2b) to instigate conversations about collective content creation and curation.

The first two workshops established a shared understanding of auditory needs and listening contexts, laying the groundwork for deeper inquiry and RtD. These exploratory discussions captured a wide range of experiences, setting the agenda for future sessions. However, it quickly became clear that abstract discussions, while insightful, lacked the grounding needed for deeper critique and reflection. This realisation shaped Workshop 3, shifting the focus toward envisioning future possibilities. The creative momentum highlighted the need to move beyond conceptual dialogue to tangible, accessible interventions. co-designers expressed a desire for concrete tools – something they could interact with and evaluate directly. This feedback led to the creation of high-fidelity probes, not as final solutions but as starting points to inspire ideation and meaningful discussions. This marked a pivotal moment, steering the workshops toward actionable, co-designer driven design pathways.

The probes acted as catalysts for engagement, enabling co-designers to evaluate features like multimodality and pacing in practical terms. Designed to gather insights and spark inspiration, they embodied the idea of boundary objects, bridging imagination and reality to provoke reflection and ideation [14, 98]. By presenting examples of automated transcripts, summarisation, and personalisation using generative AI tools, we showcased the potential of these technologies, expanding co-designers' understanding of what could be achieved – concepts largely unfamiliar to our demographic. This approach minimised abstraction, giving co-designers concrete experiences to envision accessible interventions with clarity – important for co-design with users with aphasia [98]. While probes were introduced in Session 4, co-designers strongly advocated dedicating the next session to exploring a community-led initiative – a subject of considerable importance to them. This ensured that their vision of a collaborative, inclusive radio platform was not overshadowed by the focus on refining probes. By using a low-fidelity radio mockup, we created a symbolic framework to embed co-designer driven

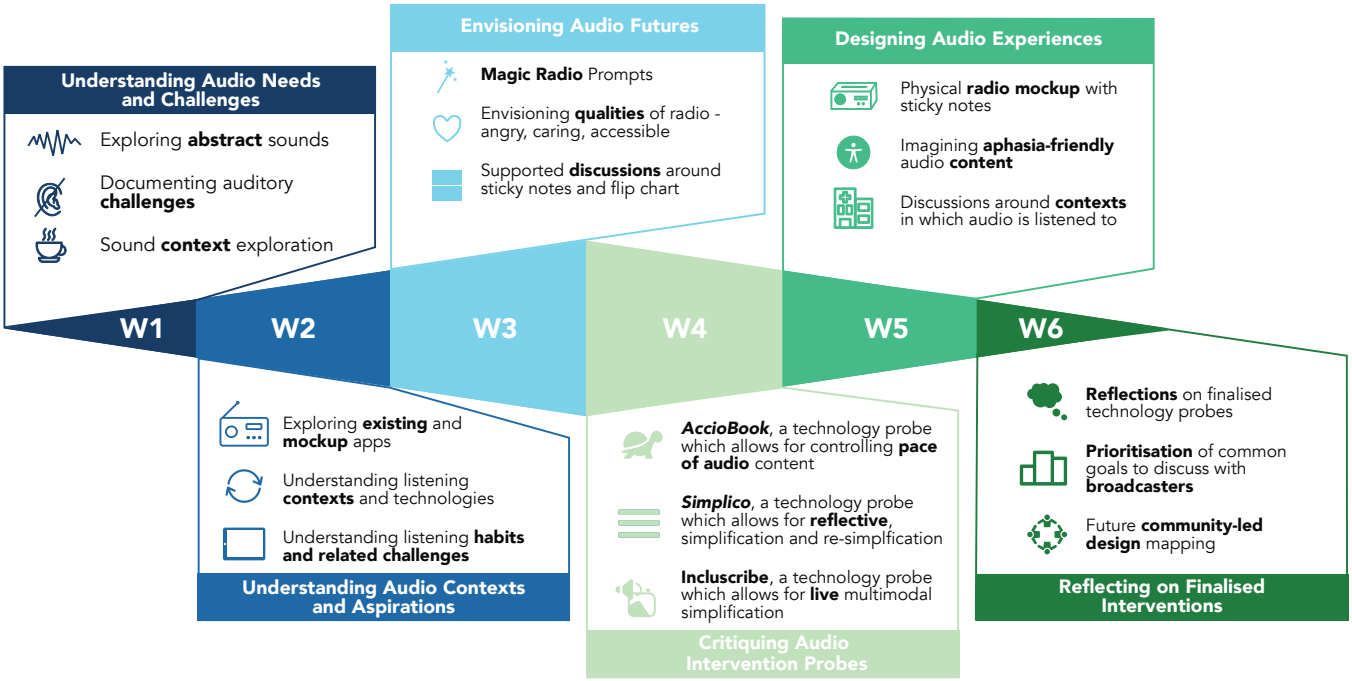


Figure 1: Process of co-design workshops, through the divergent thinking workshops (W1–W3), where the possibilities are explored, to the convergent workshops (W4–W6), where we begin to converge on envisioned solutions and reflect on their implications.

ideas, allowing for dynamic input and creative exploration. Workshop 5 became a bridge, honouring the group's priorities while maintaining the study's trajectory toward the deeper critique and synthesis planned for Workshop 6.

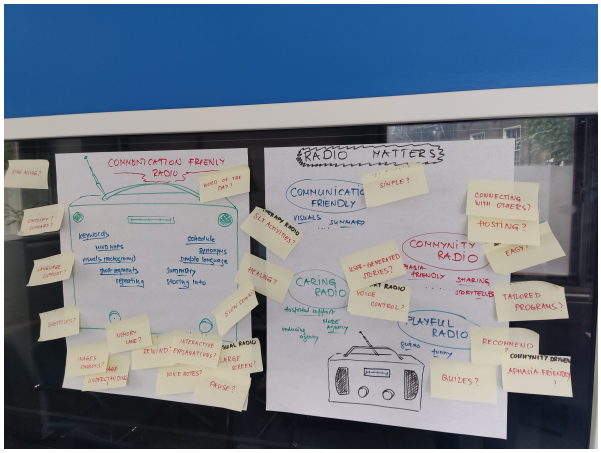
Finally, Workshop 6 synthesised the insights and decisions from earlier sessions, revisiting the probes to solidify priorities for real-world application. This iterative process emphasised the dynamic and responsive nature of the study, allowing its focus to co-evolve with the group's needs.

3.3 Co-Designers

Building on the trust established through previous research, we fully informed them about this new project. They were allowed to ask questions and discuss their involvement with family and friends before providing informed consent, ensuring they understood the scope of their participation. This was all supported by an SLT.

3.3.1 Recruitment. This project was conducted as part of a broader initiative on accessible media interventions for users with aphasia. The focus on audio-only formats emerged organically from informal conversations and interactions with individuals attending Aphasia Re-Connect outside the workshop sessions. These exchanges highlighted a significant unmet need, prompting the creation of an initial flyer to recruit co-designers for exploratory workshops. Recruitment was intentionally inclusive, avoiding rigid classifications like expressive or receptive aphasia (e.g., Broca's or Wernicke's – receptive vs. expressive difficulties). With guidance from a seasoned SLT at the charity, bringing over 30 years of experience, we recognise auditory processing challenges transcend traditional clinical categories. Clinicians acknowledge that disruptions in Broca's areas can impact Wernicke's functions and vice versa [79]. Such classifications, while useful diagnostically, impose artificial distinctions that fail to reflect the dynamic realities of aphasia, particularly in auditory contexts. By focusing on co-designers' lived experiences rather than rigid labels, we ensured the design of audio interventions addressed the broad, interconnected spectrum of challenges people face.

Following best practices [57], we conducted the sessions in a familiar and comfortable setting. Each session, lasting up to 1 hour and 45 minutes, took place in a dedicated room within an Aphasia Re-Connect charity, where co-designers were comfortably seated around a table, engaging with the material via a screen or whiteboard – see Figure 7. In total, we recruited 10 co-designers – see Table 1. Each session was supported by an SLT involved at the centre, providing communication support, occasionally helping us to devise materials and deciding on the best approach. All co-designers were fluent in English prior to their stroke, and none of the co-designers used augmentative and alternative communication in the workshops. The mean age of the co-designers was 67.7 years (SD = 10.71 years), reflecting a moderate dispersion of ages within the group. The mean duration of aphasia among co-designers was 8.9 years (SD = 4.13 years).



(a) Important design themes and intervention ideas from session 3

(b) Portable radio mockup capturing community-led audio media initiative from session 5

3.4 Data Analysis

We adopted a qualitative descriptive approach, integrating elements of narrative inquiry [85], such as co-designers' stories detailing their post-stroke media experience and identifying relevant design requirements. Co-design methods were employed to look into the nuanced aspects of post-stroke media experiences, specifically to gain a comprehensive understanding of who the individuals are, their challenges and needs they encounter, and where these experiences take place within their everyday lives.

We conducted reflexive thematic analysis [16] to gain insights into our co-designers experiences and consider how these insights could inform design of audio-media experiences. This approach was framed within a constructivist–interpretivist paradigm [25] exploring the experiences, interactions, and interpretations of individuals. The researchers were not objective observers but were actively engaged in interpreting and understanding perspectives in a collaborative manner. The research team held regular meetings to discuss and exchange notes, reflecting on key topics and pain points while planning the next workshops. Data from W1–W3 was thematically analyzed, shaping the design and implementation of the probes introduced in W4. Following the completion of W4-W6, all findings were synthesized through a final round of thematic analysis, which informed the future directions presented to co-designers. This analysis was conducted manually using Miro, where data was iteratively coded, clustered into thematic categories, and visually mapped to identify accessibility challenges and co-designer driven insights.

The data analysis involved identifying patterns and themes that emerge from co-designers' narratives, always considering the context of the co-designers' own meanings and understandings. We looked for specific instances of how co-designers made audio-media technology part of their everyday post-stroke life, to determine audio-media experience use in generationally and culturally specific ways, and as part of the shifting ecology of media presence.

The probes emerged as a cohesive synthesis of findings from W1-W3, with each workshop building on insights from the previous one (see Table 2). For instance, the use of generative AI to create images and keywords highlighted the potential of automated accessibility features – possibilities previously unfamiliar to co-designers. This approach aimed to demonstrate broader possibilities for accessibility in audio-media, focusing on understated yet impactful considerations. The synthesis was structured around three key principles: (a) representing the multifaceted nature of audio listening by assigning diverse roles to media (e.g., enhancing communication potential); (b) addressing the lack of structured research in this domain, often driven by anecdotal insights (e.g., [44]); and (c) emphasising simple yet vital accessibility tweaks identified by co-designers as crucial. For example, temporal accessibility interventions were explored to reduce cognitive effort, helping users overcome barriers to meaningful engagement with audio-media.

4 Resulting Technology Probes

Based on prior workshop findings (see Table 2, we deployed three technology probes in both Sessions 4 and 6. For ease of reading, we present the final versions of these prototypes (e.g., those in Session 6) here in isolation, before discussing the findings, from which they were borne and shaped.

AccioBook, presented in Figure 3, was developed as a result of our co-designers' desire to manage the pace at which audio content was presented. AccioBook allows users to control the pace at which spoken content is delivered, complemented with AI-generated visuals of ongoing events. AccioBook takes the form of an iOS audiobook app, deployed on iPhone. The front-end developed was developed as a native Swift app, with a Django back-end. The Librosa Python library was used for segmenting the audio; segmenting by bins of < 40 dB. Synchronisation of the story's text (from Project Gutenberg) was done via Google's Cloud's Speech-to-Text, and images were generated with DALL-E 3. The content used in AccioBook was an audiobook of Oscar Wilde's 'The Canterville Ghost'.

Simplico, presented in Figure 4, was developed as a result of our co-designers' wish for comprehension support at points where they experienced challenges understanding the content. A 15-second summary is provided throughout the audio playing, but the key features allow for the simplification and the re-simplification of

Table 1: List of co-designers in the workshops, along with demographic data and communication challenges. ASR provides an index of the severity of the aphasia from 0 (speech, writing and/or auditory comprehension are not functional) to 4 (Although the individual feels that they have a problem with language, this is barely apparent to the listener who may not detect any problem)

Name (attendance)	Age	Years w/ aphasia	Aphasia Severity	Communication challenges
Nancy (1–6)	85	6	ASR: 4 – High level, mild aphasia	When tired, finding and expressing words can be tricky. Reading and writing similarly.
Chris (1–6)	52	12	ASR: 4 – High level, mild aphasia	Word finding and greater difficulty expressing longer or more complex ideas. Reading and writing limited. Following conversation in very noisy, fast environments can be more challenging.
Nathan (1–6)	57	14	ASR: 3 – Receptive and expressive aphasia	Word finding difficulty. Understanding is more difficult in crowded or fast paced circumstances.
Nick (5, 6)	65	14	ASR: 4 – High level, mild aphasia	Fatigue can hamper expression reading and writing.
Sava (5)	78	1	ASR: 4 – High level, mild aphasia	Occasionally slower than before her stroke, occasional word finding difficulty.
Joely (1–3, 6)	62	13	ASR: 2 – Receptive and expressive aphasia	Marked word finding difficulty with problems formulating sentences. Following conversation in very noisy, fast environments can be more challenging.
Eliot (1, 2)	74	6	ASR: 2 – Moderate receptive and expressive aphasia	Difficulty reading and finding the right words and expressing more complex ideas. Following conversation in very noisy, fast environments can be more challenging.
Vanessa (3)	77	7	ASR: 2 – Receptive and expressive aphasia	Severe aphasia reduced to limited one-word utterances. Uses written word to communicate her ideas. Comprehension requires support through rephrasing slowed speech, pausing and written word.
Ethan (1)	54	10	ASR: 2 – Receptive and expressive aphasia	Marked word finding and problems putting words together. Understanding is more difficult in crowded or fast paced environments.
Adam (6)	73	6	ASR: 3 – Receptive and expressive aphasia	Can communicate verbally but it can be difficult to find the right words. It can be difficult to follow what people are saying. Can feel confused and have memory issues.

content to allow the user to pause and support their comprehension. Simplico allows for multiple iterations of simplification, from a 'Grade 10' to less than a 'Grade 4' reading age. Simplico also allows for alternative representations such as keywords, emoji and AI-generated images. Simplico was implemented as a web application, with a Python back-end and ran on a laptop. Speech recognition of the audio content was accomplished with the SpeechRecognition Python library. Simplification requests were accomplished with OpenAI's API for ChatGPT-4 for text, images and emoji – e.g., with the prompt "*summarise this description with 4 emojis*". The content used in Simplico was an audiobook of 'Little Women' by Louisa May Alcott.

IncluScribe, presented in Figure 5, was developed as a result of our co-designers' wish to have additional modalities to complement audio content, with an aim of enhancing comprehension without interfering with the audio experience. Similar to Simplico, a live updating summary is provided, but IncluScribe also afford 'live' visuals which summarise the content and update at a regular interval. The user is able to choose the time threshold for the update (e.g., the last 10% of audio). IncluScribe was implemented as a web application and run on a laptop, using the React framework. Similarly, the OpenAI API was used for image generation and text summarisation. Prompts for summarisation/generation included "*Summarise the given text to make understandable for someone with aphasia. Maximum number of sentences should be 3 and maximum words in each*

Figure 3: 'AccioBook' technology probe. The AccioBook technology probe supports users in controlling the pace of the audio content. The interface allows for its text to be displayed (a) and enables an 'add pause' feature, which segments the audio into chunks, denoted by red parts of the audio waveform in (b), and enables the user to choose a time of pause between speech (between 1 and 5 seconds), using the 'add pause' button (c), finally an AI-generated visual representation of the scene is also provided (d). The key for interface controls is shown in (e), including a 'rewind' feature which returns to the beginning of the current segment or goes back through previous segments.

sentence should be 8". In all cases, prompts were based on aphasia literature on 'aphasia-friendly' content (e.g., [77, 84]) and on trial and error with the ChatGPT chatbot interface. The content used in IncluScribe was David Attenborough's 'Frozen Planet' and the audiobook of 'The Magic Shop', by H.G. Wells.

5 Findings

The findings presented here are a synthesis of insights from all six co-design sessions. Instead of focusing solely on the technology probes and radio mockup, we compiled overarching themes that emerged throughout the process (see Figure 6). While we aimed to capture an equal distribution of co-designers' views, this was not always possible due to varying attendance. Certain individuals feature more prominently – i.e., three co-designers attending all six sessions. This approach nonetheless provides a broad perspective on the challenges, opportunities, and emotional connections with audio-media, highlighting peoples' unique needs and visions for the future. While certain themes overlap – such as temporal access, cognitive load, and multimodal enhancement – we find it essential to consider each separately. Each offers distinct elements significant enough to warrant individual attention.

5.1 Temporal Audio Access

Simplico and IncluScribe were designed to embody key ideas from the early co-design sessions, focusing on two central concepts: asynchronous support, which encouraged reflective interaction with the audio-media content, and synchronous interventions, designed to appear in real-time alongside the content. While certain features, such as image generation, overlapped between the two probes, each maintained distinct contextual meaning based on its specific interaction model. This careful balancing of temporal accessibility

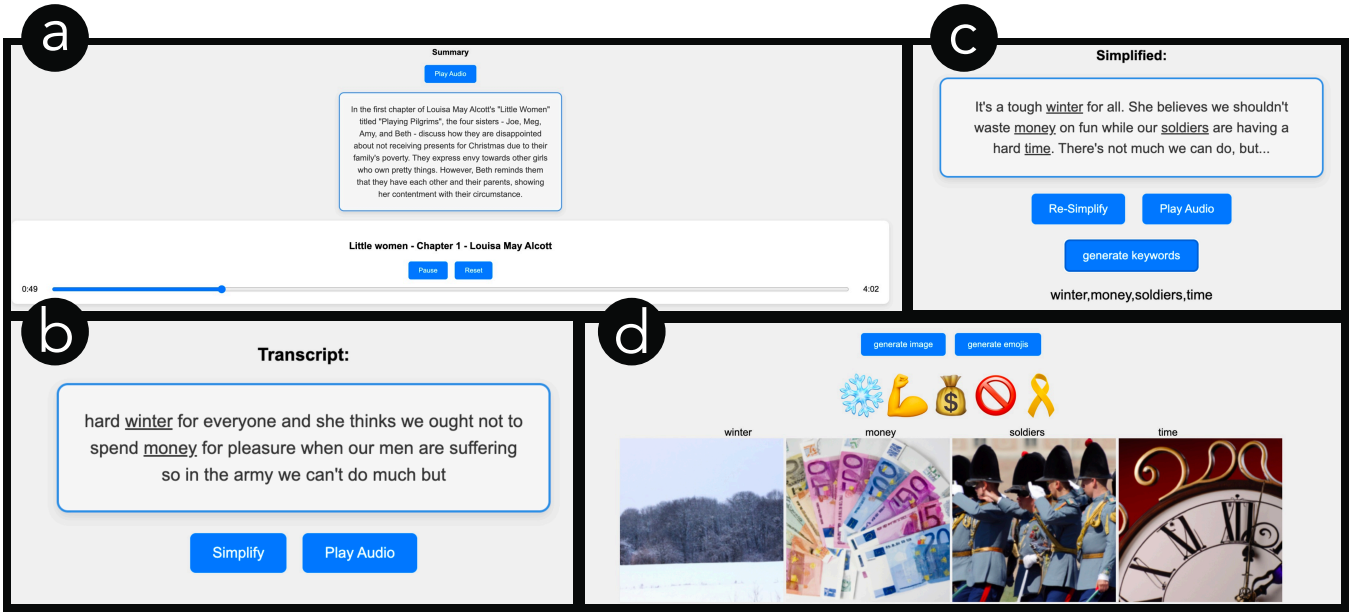


Figure 4: 'Simplico' technology probe. Simplico enables users to simplify moments in audio they find challenging. Users are provided with an updating summary (a) to orient themselves, but may also pause and simplify: (b) shows the direct transcript, allowing them to simplify further keywords, which are underlined; (c) allows for further simplification and keyword highlighting; then (d) finally allows for emoji or visual representations of the keywords highlighted by the simplification.

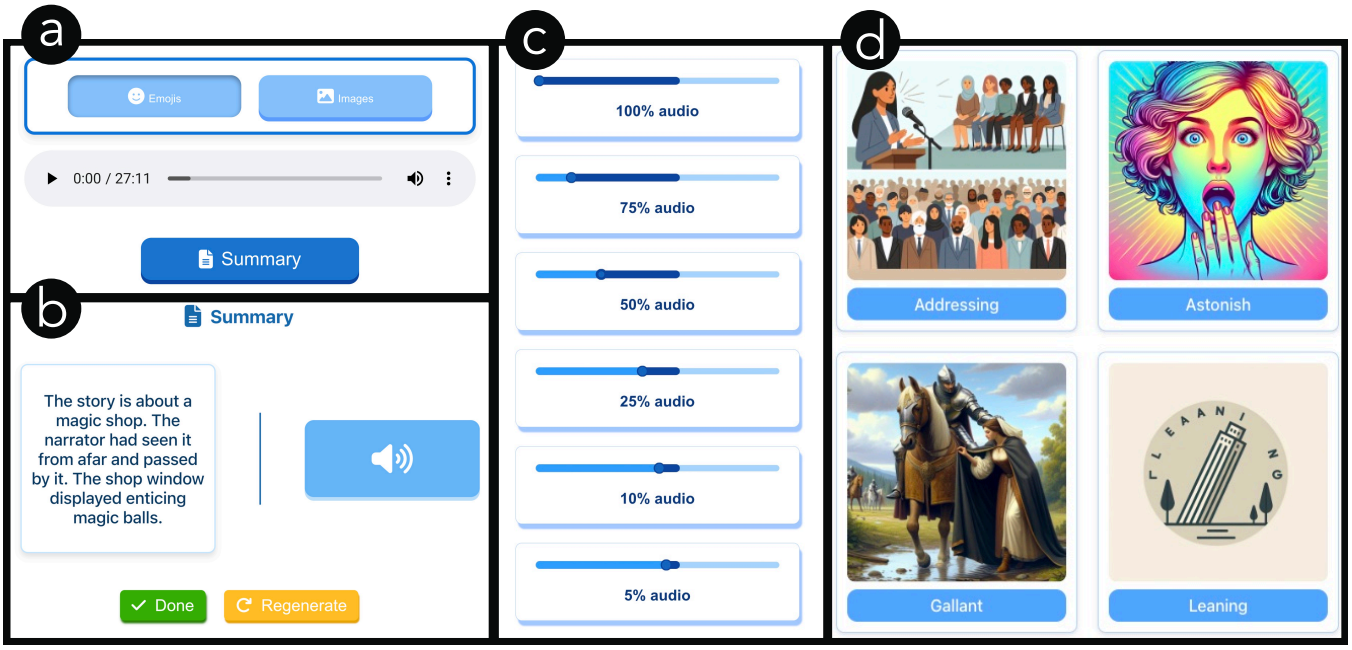


Figure 5: 'IncluScribe' technology probe. IncluScribe focuses on access to live audio. The main UI is shown in (a). A live, updating summary is shown below (b). The user is able to select how much of the previously listened content is summarised by pressing buttons ranging from 5% to 100% (c), finally users can access visual representations of keywords for further support (d).

allowed us to explore the nuances of how co-designers would like to engaged with audio-media.

5.1.1 Sequential Processing and Reorientation. Our co-designers frequently surfaced the challenge of *centring* themselves within

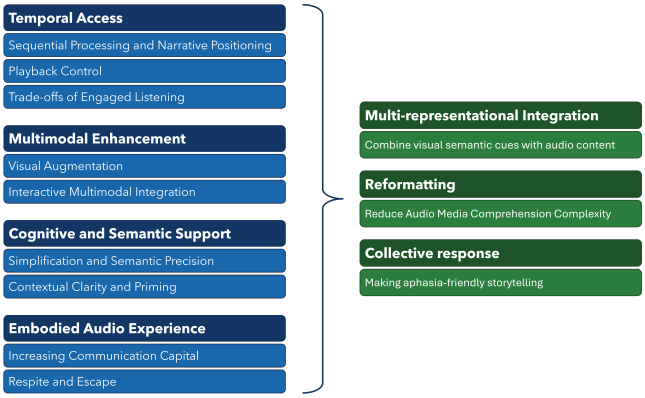


Figure 6: The main themes and sub-themes that emerged from the data analysis, along with the main takeaways from the six workshops.

stories in complex auditory environments, particularly when faced with long streams of continuous speech, such as in audio dramas or any long-form auditory content (i.e., audiobooks). The difficulty in processing connected speech often led to a disorienting sense of being lost within the narrative. When discussing Simplico's simplification and IncluScribe summarisation feature, Nancy proposed an idea centred on the sequencing of audio narratives, expressing, "*I like positioning myself*". After her stroke, she encountered a specific problem with discourse recovery – struggling to determine the sequence of events. She would often lose her way in the narrative, unable to clearly demarcate the beginning, middle, or end of the radio story. This challenge was echoed by Chris and Nathan, highlighting a need to maintain a narrative continuity and reorientation of complex auditory content – i.e., pause, simplify, and reflect features offered by all three probes.

> "*I listened to a radio program about the democratic, the demo...crat [Nancy encounters a blockage, struggling to recall the word but continues despite the difficulty], where Harris was endorsed. 'You can see me having a difficulty now'. And if I tell the story now, I have the difficulty in putting them in sequence. Sometimes I do mindmaps as branches of a tree and then various pieces coming off [referring to the method she uses to visually organise information]. (Kamala was endorsed, what is Biden's role in the speech). It is all sequenced. I presented people with mind maps and my personal problem and they say we quite like that. It is quite simple.*"

Nancy's insight resonated with others, sparking reflection on the difficulties of navigating speech and the associated challenges – i.e., short-term memory retrieval. As discussions unfolded in Workshop 3, co-designers began to articulate how these struggles are intricately tied to the experience of listening. Chris, in particular, spoke of the once vivid memories that now seemed to have lost their potency, no longer anchored by the words that used to evoke them. Chris expressed this by saying, "*You know there is something happened, something very important, but you forget. You do not know where to go, because it is a short-term memory loss*".

This recognition of feeling adrift in the auditory landscape suggests the need for tools that could help listeners regain their bearings on their own time. The 'smart rewind' feature in AccioBook was inspired by this. It became particularly significant, as Chris emphasised the value of being able to "*re-listen, go back to it for whatever reason*," highlighting the importance of features that allow individuals to reorient themselves within the flow of conversation and information. Specifically, co-designers discussed the potential of the IncluScribe that uses sliders instead of buttons in Workshop 4 to navigate and segment the storyline by choice, offering a more intuitive way to manage and revisit complex auditory content. However, in Workshop 6 buttons were preferred over sliders due to easy operation – some people had a problem with fine motor movement. The value of this feature was also seen as significant; with an intent to shorten to content and split up the narrative such that it can be listened to in segments allowing for rest and respite – "*Some people don't have an ability to remember*" (*Eliot*) or "*Refresh, refresh, refresh-ment* " (*Chris*).

The first time Nancy introduced the idea of mindmaps in Workshop 2 and 3, the others took it with a grain of salt. However, by Workshop 6 the conversation shifted when Nancy brought in a tangible map she created herself compiling information from different radio programs (see Figure 8). This visual tool, used regularly by her sparked interest. For the first time, other group members began to see its potential. Nick and Chris, in particular, highlighted its relevance for more complex topics like governance, politics and voting. Chris shared an example from the news about an upcoming podcast "*There's going to be a podcast about easy, easy, easy governance – it'll be idiot-proof*" – a concept he felt could benefit from a visual aid. Nick agreed, adding – "*It's like a tree, that's what we need*" – emphasising the value of this approach for clarifying intricate subjects.

5.1.2 Pacing and Pausing for Comprehension. Many of our co-designers exhibited signs of apraxia – difficulty in speech articulation – and anomia – challenges in naming objects – but their need for time in conversation extended well beyond the act of speaking. It also encompassed the crucial time needed to fully understand and process the communication being directed at them. This dual need for time, both to articulate and to comprehend, inspired a compelling accessibility vision to slow the pace of speech, incorporating deliberate pauses between words or sentences. These pauses, particularly after more complex phrases and "*long words*" (*Chris*), would offer individuals the necessary space to grasp the content more effectively. Reacting to AccioBook 3 Nancy noted, "*this is very good...,*" highlighting the impact this feature could have on her listening experience. Chris echoed this appreciation, emphasising that these intentional pauses would greatly enhance their ability to engage with and truly understand audio content. Most people emphasised the importance of pacing, where the rhythm of speech can be tailored to their cognitive needs. This was already present in their daily audio-media routines as Nancy states "*I listen to BBC Sounds, because I can relisten*". Nick too would repeatedly indulge in 'The Hitchhiker's Guide to the Galaxy' available to him in an audio format.

However, presenting AccioBook probe led to an interesting discussion. Nathan warned about a risk of long prolongation in the

Figure 7: Picture from a session, with co-designers listening to audio from a technology probe from the laptop.

face of forgetting "*It can be confusing, trying to remember last thing said. But then you do and you lost where you are*". Nick, yet, said "*this is better than going back in time, if you can, to go back. to where you lost the plot*" while Chris insisted that this can be "*patronising*" if the pause is lengthy. Adam added that "*I found have found to be very useful when I first started my reading and understanding. This has brought up something in me, I haven't read a book since I've had stroke which was six years ago*". This implied that when interacting with AccioBook a need for better control mechanisms is needed. For instance, two co-designers favoured control over the use of pauses and timetag annotations. In particular, Chris suggested to annotate the audio content over the course of an episode or radio program. Also an automated system was suggested by Nick suggesting a goal-directed system to tag the content.

5.1.3 A Mindful Act of Auditory Navigation – The Effort of Listening. For people with aphasia, the effort required to communicate and understand is not something to be shied away from; it is an integral part of their daily experience – i.e., communication cards that provide basic personal information as seen with Nathan, or to use written and drawn lexical support, as Vanessa did, who could only verbalise two words. These strategies are not just practical but essential in daily interactions. For instance, when visiting a GP, people bring pre-prepared speech cards to articulate their symptoms – an adaptation that has become a routine part of their lives [28].

To facilitate this, co-designers emphasised the need for structured audio-media support, such as having a schedule or introductory materials that prepare them for what they are about to listen to. Nancy evoked the value of curation "*It is a schedule that would help you,*" highlighting the value of preparation to make the audio experience more accessible and meaningful.

Chris explained why it is important to actively engage "*If somebody is telling me something and I don't understand, I'm not going to just say good and smile. No, I don't understand. Point. And they could be using big words, but I'm not going to say, OK, OK, that's good. No, I don't understand. I need it straight. I don't understand*". This insistence on clarity over complacency illuminates the importance of reducing confusion and ensuring that communication is effective and transparent.

Some co-designers were willing to make trade-offs – between immersing themselves in the atmosphere of the story and a complete understanding of the plot. For example, Joely expressed this by leaving out the time to listen to the story of 'Little Women' for a bit longer before Simplico's output and IncluScribe's summary, stating "*I like the summary as you go more into it*" referencing to extend the listening even at the cost of not understanding everything.

> Researcher: "*There are two types of summary here, one is beforehand, before even listening, to get to know the characters and the plot. And the other is getting the summary as you go along without giving you the information prior to listening. So which one do you prefer more, the one where you have the summary before you start listening, or as you go you decide on the length of the summary and the level of information that you want to have? [alluding to the 'Little Women']*". Adam: *"For me, it would be the second."* Joely: *"Yes"*. Nathan: *"Yes, the second one"*. Nathan: *"Joely, did you get it from the summary?"* Joely: *"Yep, yep!"*.

However, while the novelty of certain content and stories sparked excitement, co-designers still valued having prior contextual understanding, especially when dealing with more challenging content. This feature needed to be individually tailored, considering factors such as previous familiarity – "*Little Women is a widely regarded book, most young girls in our time have read this book.*" The complexity of narratives with numerous characters, events spanning different eras, and varied contexts highlighted the importance of this preparation.

5.2 Multimodal Enhancements

We discuss the transformative potential of integrating visual and interactive elements into audio-media. These multimodal enhancements can profoundly enrich the listening experience for people

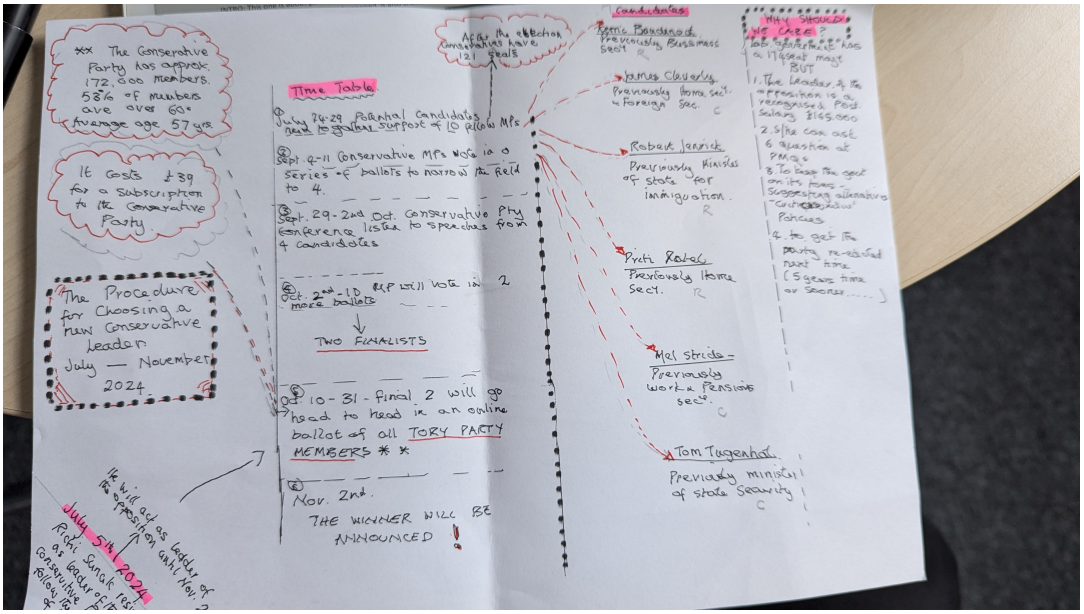


Figure 8: A handwritten mindmap created by Nancy about the Conservative Party in the UK

with aphasia. Through the seamless blending of visual augmentation with interactive technologies like text-to-speech or voice activation, we will illustrate how these interventions reinforce comprehension but also empower listeners to engage independently with the audio content. While our findings highlight a variety of accessible interventions, we intentionally place emphasis on the visual elements, recognising the significant role these elements play in bridging the gap between hearing and understanding.

5.2.1 Visual Augmentation. While audio is central, co-designers also recognised the value of visual aids in reinforcing spoken content. co-designers with aphasia expressed a clear desire for compensatory strategies that would enhance their understanding of spoken content. While our live captioning feature, which includes live transcripts and captions for radio (see IncluScribe, Figure 5), showed considerable promise, the co-designers emphasised the need for further refinement.

As Chris thoughtfully suggested, "*Could you underline the words, can you highlight it? Even with the keywords that the app offers, it would be good to highlight the words as they speak along.*" This visual reinforcement was seen as particularly beneficial within a caption system for radio, where, as Chris noted, "*if you are stuck on the word, it would help*". This feature captured by IncluScribe was then offered in our next iteration and presented in Workshop 6 confirming again its value for the community. "*We need a Kindle for radio,*" as Nancy, Chris, and Nathan envisioned it, a tool that will merge the auditory experience with visual support, much like the Kindle text-to-speech feature.

Most of our co-designers were older adults who had encountered discussions about generative AI through mainstream media, yet had little understanding of its functions. The revelation that IncluScribe can generate images, synchronously, capable of integrating visual elements into everyday radio content was met with surprise. This sparked a discussion about how such images should be presented and what their implications might be for people with aphasia. Nathan and Chris dismissed the use of emojis, viewing them as largely ineffective beyond the most basic expressions "*With these new emojis you need a dictionary*" (Chris). However, this was not universally shared, and the conversation quickly expanded into a broader debate about the appropriateness of different types of images for people with aphasia. Nancy was particularly adamant that cartoon images could be of significant value, as they clearly depict meaning in a simple, less overwhelming way. She acknowledged that some might view this approach as patronising or childish, even potentially offensive, but she offered a poignant perspective: "*And I understand that, but on the other hand, you are a child learning to speak again. It is very difficult, and I have no problem with that.*"

The image accuracy emerged as crucial for our co-designers, underscoring the potential usefulness of this feature. As they listened together to Simplico's audio content of 'Little Women,' three images were generated alongside respective keywords. Reflecting on the results, Nick muses with a burst of laughter, "*When I think of Little Women, I don't think of that image!*". Adam and Nancy chimed in with a suggestion, proposing that the images could be improved by drawing from the vast array of existing book illustrations, particularly for such a well-known title, leveraging already established datasets *"You could use the images from the movies, a lots of films were made about it" (Nancy).* More so, Adam was puzzled by why the images simply capture the atmosphere of olden times "*what is been, beeeen [prolonged], you've told me that book, was written in 1800s and the pictures in my head do not*". They took it even further, adopting a meticulous approach to the importance of image details; "*A picture is worth 1,000 words, isn't it,*" Nick remarked. Others were quick to point out specific aspects of images that were critical for understanding, noting, how attention to such details how, for them,

every element of an image carried meaning. Commenting on the images generated by Simplico Nancy added "*Even the shoes, it is not concentrating on the heels*".

This reliance on visuals was not unusual for our co-designers, as many of them had already integrated images into their daily routines. In Workshop 1, the importance of visuals over sound quickly emerged. co-designers emphasised how crucial visual cues were in helping them understand and orient themselves, particularly in unfamiliar settings. "*I will always look at the screen and the signs,*" Ethan noted when discussing the role of visuals in public transport. Ethan, in particular, was insistent about their significance, cutting in to say, "*I always look for visuals. Otherwise, I won't understand. I sit and look around. If I went to Paris, I would be lost.*". For some, the visuals were even more central to communication. Two co-designers relied on written notes and sketches to express themselves. Vanessa, for instance, often communicated by writing or drawing, using these methods to ask questions or seek confirmation. Visuals, for these individuals, were not just supplementary but vital for making meaning.

5.2.2 Interactive Multimodal Integration. While our accessible content – such as summaries, captioning, and transcripts – has undoubtedly expanded access, co-designers expressed a clear need for voice read-outs (text-to-speech) to further enhance their experience. This was particularly emphasised by Eliot, who, in Workshop 2, shared his struggle with co-occurring post-stroke visual impairment. Battling oscillopsia – a condition resulting from an aneurysm– Eliot described his difficulty reading, as everything appears doubled and distorted, with images, words, and everyday objects jiggling in front of him. As we discussed the BBC Sounds app and its existing program summaries, Eliot pointed out that while these summaries and transcripts augmented audio content, they were not necessarily accessible to him. This observation became a pivotal moment in the discussion, as Eliot's idea – despite his absence from the following four workshops – resonated with the group. His request for a text-to-speech feature reappeared throughout the sessions, with others frequently recalling his need and incorporating it into the conversation around technology probes. In discussion of these Simplico's and IncluScribe's features three different alternatives and directions emerged.

First, the introduction of voice read-outs in Workshop 3, later reinforced during the presentation of Simplico in Workshop 6, highlighted the crucial importance of speech quality and the variety of presenters' voices. This feedback highlighted the nuanced preferences of our co-designers, with three of them emphasising that the choice of voice was important. They expressed a strong preference for familiar presenters whose voices they found to be not only alluring and accessible but also consistent with their existing listening habits. This preference highlights a gendered and branded response to the feature, where the voice becomes part of the trusted radio experience. For instance, when the Simplico in Workshop 3 introduced a synthetic voice to read out summaries or captions, caused confusion at first. They found the synthetic voice to be undesirable and out of sync "*disattaching" (Nick)* them from the established tempo and quality of the radio presenters they listened to "*Because if you are listening to a play, you don't want to interrupt the flow of the play" (Nancy).* While being aware that it might not be easy to mimic the same voice they were ready to make trade-offs Adam: "*If the voice was similar, it would make more sense, it would join the two together, as opposed of having a big split*". However, others did recognise the importance of voice quality as they struggled with accents and what they called mispronunciations, an important for their everyday understanding of the speech content *"It annoys me, the pronunciation, it is difficult. Even the American English can be a problem. I wouldn't understand".*

Secondly, an interesting suggestion emerged from Workshop 1 and 2. Despite most co-designers reporting no significant auditory impairments, many highlighted difficulties in recognising certain speech sounds or localising them. For instance, in Workshop 1, Ethan shared his struggle with understanding speech sounds: *"(Researcher): Radio can be hard for you? (Ethan): Yes, hard to understand words sometimes."* In Workshop 2, Eliot initiated a conversation about the challenge of localising sounds: *"(Researcher): So, you find it hard to localise, to figure out where it's coming from? (Eliot): Yes, I can hear but don't know where."* These discussions led to reflections on the potential benefits of offering alternative voices – reinforcing the importance of audio quality and providing options for a plurality of voices. While the synthetic voice was mainly criticised for its separation from the actual voice it was still interesting as an alternative as long as it met certain requirements *"Punctuation, there were not full stops, it was too fast and too artificial".*

Finally, an intriguing alternative emerged during Workshop 6, while discussing Simplico and IncluScribe, both of which featured options for simplification and summarisation. The idea of training their own voice to read out summaries, aligning with their internal voice. This concept was sparked by a conversation about the potential of generative AI to enable such personalisation – offering a more tailored and intimate listening experience that co-designers had not previously considered. The possibility of hearing familiar voices, particularly their own, was viewed as a way to bridge the gap between the auditory experience and personal identity, creating a more resonant and meaningful interaction with the content.

5.3 Cognitive and Semantic Support

In this section, we will focus on the cognitive and linguistic strategies that make audio-media more accessible for people with aphasia. We will examine how techniques like language simplification, keyword generation, and semantic precision reduce cognitive load and enhance understanding. By offering clear, relatable language and allowing for active engagement, these supports enabled our co-designers to grasp complex concepts and interact with content in a way that is both meaningful and tailored to their individual needs.

5.3.1 Simplification and Semantic Precision. Simplico's feature of audio simplification and keyword generation invoked a thoughtful discussion about the significance of providing effective cues. Among our group, those with more pronounced anomia (difficulty with word finding) emphasised the importance of carefully reconsidering the type of keywords, the level of associative language, and the structure and degree of simplification.

In Workshop 4, the initial discussion around summarisation and keyword selection sparked excitement among co-designers who saw great potential in Simplico. However, the enthusiasm

shifted in Workshop 6 with co-designer's being drawn to the layered simplification (i.e., a summary of 100 words, a summary of 20 words, a 1-sentence summary and finally the list of keywords) overshadowing keyword selection. Most people have not seen the advantage beyond a reminder due to short-memory retrieval issues. Chris saw this more as a reminder of *"going back yourself and you want to remember"*. While simplification earned more praise, the keyword feature's functional value was seen only in terms of associative and relational language. In everyday media language, many co-designers found that standardised speech metrics and abstract concepts, such as those used in informative programs like the news, often felt disconnected from their lived experiences and language use. Chris highlighted the challenge of grasping abstract notions, using examples like '100 yards' and suggested that relational and associative language, paired with well-chosen keywords, could be more helpful – *"Just tell me how many bus stops"*.

While Simplico probe included a simplification of recently heard content by shortening the sentences and removing the transcript complexity, co-designers expressed the need for precision and thoughtful word choice. Nancy cautioned against oversimplification, highlighting the potential for added confusion *"It's a book. It's a pen. It's a table and the brain thinks, oh, that's really good. We've got the bread, but it makes an enormous amount of difference if it's over or under or before or after. This is what I find difficult, and a lot of people with aphasia find difficult. The little words. The just the keyword, just the noun wouldn't be that helpful, I don't think. What others think?"*. This then begun a deeper discussion about the balance between simplifying content and maintaining clarity. For Chris and Nathan keywords on their own, without more context, and with no syntactic linkage were difficult to understand prior to listening to any type of radio program *"Everything for me would just be a line of words" (Chris).*

5.3.2 Contextual Clarity and Priming. People with aphasia in our group often experienced a profound gap between hearing and understanding, where the words they heard did not easily translate into comprehension. This challenge was particularly acute in the early months after a stroke, when sounds that once made sense became elusive and fragmented. For our co-designers, the act of listening transformed into an exercise in deciphering – a process that demanded not just the reception of sound but a careful reconstruction of meaning. This reconstruction was often piecemeal, relying on fragments of speech to piece together a coherent understanding of the radio program. For instance, when listening to airport announcements as an example from Workshop 1, Joely described how catching just a single word – often one that might be difficult to discern even for typical listeners – became a critical starting point for understanding. This isolated word served as an anchor, allowing Joely to connect it to the broader context and gradually piece together the intended message.

> Researcher: "*Do you understand what they said and what is happening here? [The researcher played the sound of an airport announcement]*" Joely: "*Boarding.*" Researcher: "*Yes. Boarding ... [nodded encouragingly and widened their eyes, signalling Joely to continue]*" Joely: "*Airport.* Researcher: "*Yes. This is an airport announcement? Is this typical of how you sometimes understand sound announcements? [...]*"

This approach to managing communication has naturally extended to how people with aphasia engage with radio and other audio-media. The need to reduce cognitive load during listening initiated discussions about the importance of contextual priming before interacting with content. Nancy, for example, expressed a strong preference for understanding the context of a book or program before engaging with it: *"I would prefer...like in the back of the book, somebody could tell you what it's about."*. Chris had a similar observation in Workshop 2. This desire for a preliminary overview reflects a broader need for support in processing auditory information, ensuring that listening is not overwhelming but rather a more manageable and enriching experience.

The ideas for interventions varied among co-designers. Nathan, for instance, emphasised the importance of content simplification, while also valuing contextual priming: *"So I would prefer that instead of telling me which actors in there, because I'm not seeing them, I'm just hearing voices."* On the other hand, Nancy expressed a strong preference for knowing more about the host, their guests, and even a brief biography, suggesting that this additional context would further enrich the listening experience. These differing preferences highlight the nuanced and individualised needs of people with aphasia, emphasising the importance of tailored approaches to audio-media accessibility. While most of the podcast episodes they listened to on BBC Sounds app were already showing program summaries and host names they felt this was not truly accessible in terms of its visual appearance and detail of information (either simplified or providing more contextual elements like specific themes, a linear sequencing of events etc.).

5.4 Embodied Audio Experience: Emotional and Social Integration

We now delineate the profound emotional and social dimensions of audio-media for people with aphasia, revealing its role as more than just a medium of consumption. We will explore how radio and podcasts serve as sanctuaries, offering comfort, connection, and a sense of belonging. Through these reflections, we will illustrate how audio-media weaves into the very fabric of their daily lives, supporting not only their cognitive engagement but also their emotional well-being and social identity.

5.4.1 Increasing Communicative Capital. For people with aphasia, the recovery journey was deeply intertwined with speech and language therapy, a cornerstone in the development of expressive and receptive language skills. It was not uncommon for them to view everyday occurrences through the lens of SLT, where even the act of listening to the radio or podcasts became part of a broader therapeutic regime, aligned with the language discipline they navigated daily. However, this connection extended beyond mere therapy. Most of our co-designers changed their approach to communication deepening their appreciation for the nuances of language. *"I became more sensitive to the conversation that actors had with each other since I had my stroke. I think I became more conscious of the way people speak, and the way people use words now. The way people put words together. I didn't take it for granted anymore."*. To this end,

co-designers channelled their language discipline into envisioning a variety of radio features that would extend their engagement with the language. They imagined a radio that not only entertained but also fostered continued language development, incorporating ideas such as language quizzes, voice notes, a 'word of the day' segments, and read-along formats. These features were envisioned as a way to seamlessly integrate the rigour of their language journey into their daily media experience, making the radio not just a source of respite, but a tool for ongoing linguistic growth.

However, this embrace of language discipline was not without its tensions. Some co-designers expressed a clear desire to escape the relentless pressure of language use, revealing an internal conflict between the drive to improve and the need to take a break from the weights of everyday language demands. Two co-designers, in particular, display a reluctance to fully embrace the expectation of perfect language use. One co-designer reflected, *"Before I was living with aphasia, I didn't understand every word being said to me. Now I think we have a sort of cultural perfection. It is very important that we understand every word and any word has the significance because we've lost our speech."*.

Co-designers in our study began to envision new formats, looks and tactile experiences that could better their needs. One idea was the concept of radio that could print a receipt – essentially a punch card that captures the key points of a radio episode. As some of our co-designers carry around such cards/notes they envisioned this printed summary to serve as a conversation starter, a tangible artefact that bridges the gap between the auditory experience and everyday interactions. Nathan expressed enthusiasm for this *"It could be in. Yeah, you could say, 'I was listening to a good radio program, but here is the summary.'"*. The researcher then suggested *"Almost like a card deck that you sometimes use?"* to which Nathan responded with clear excitement, *"That would be good."*.

This internal struggle was further illustrated by Chris, who asserted, "*I cheat at everything. No, it's like if I can't say a word, I go to my phone and Google. And I'll say to Google, I'll write the word and it would say, oh, OK, that's how you say it.*" In response, Nancy reassured them, *"That is not cheating, that is sensible."* Yet, Nancy also acknowledged the exhausting burden of constantly navigating these linguistic challenges, adding, "*When I come up to a word, and I have difficulty pronouncing or finding it, in a split second, I will substitute, but this is exhausting, especially for people with aphasia, all the time.*"

These reflections bring to light the dual nature of language recovery for people with aphasia – where the desire to design language-related audio-media features coexists with the very real need to ease the strain of constant linguistic effort and suggest accessibility (section on accessibility) when communication breakdowns ensue.

5.4.2 The Role of Audio-media as Respite and Escape. A theme that emerged from four co-designers was the use of media as a form of respite. For instance, Ethan's relationship with media was comforting, a form of a respite from the burden of everyday communication. The pressure to use language, even to converse with family members, children, and his speech therapist became overwhelming and expected – "*I can't remember much anyway, but what I do remember is kind of needed the TV. Hmm, because I couldn't, my walk wasn't great. And so being around TV was kind of bond. And the children always want to talk"*. Chris and Nathan shared a similar sentiment. Under the pressure of rehabilitation, time spent in a hospital, and hard recovery work ahead, they sought respite wanting to connect to the outer world and familiar things *"Hearing familiar voices was important"* (*Chris*). For Nancy her portable radio was so important that she let her family know "*I drew it, drew it. I couldn't speak, I drew my radio. I said please bring this. Radio four was an enormous comfort for me*".

Beside respite people also felt a need to design for a place of sanctuary. Just after stroke people faced a deep sense of disorientation, loss of self, and a time of profound adjustment, as echoed by all "*I am 14 years old. It has been 14 years since my stroke, and I had to learn to read and write like in school"* (*Chris*). As the stroke had altered their physical disabilities people emphasised the importance of situating audio-media within a context where they felt most vulnerable. The hospital, surfaced the conversation the most, as a place that could play a crucial role in offering connection and sense of belongingness, especially in Workshop 5 *"Hospital is very isolating and very frightening"* (*Nancy*).

'Truth' was an important overarching theme and a concern, envisioning a community-led radio that forms a discourse around accurate, reliable and comforting information *"Telling as it is, should be the name of the radio initiative"* (*Sava*). Further, agency and identity was discussed as a need to create content that will help early stroke patients to accept who they are and their new self *"A strange voice coming out of your mouth. You do not sound the same"* (*Nathan*) and "*That strange voice*" (*Joely*), as Chris confirms *"it simply does not come out"*. People considered other design arrangements – reluctance to create audio content that gives advice and provide a more positive recovery lens. Sava suggested a *"a collection of voices and stories"* to foreground diversity and *"tips of living well"*. Many ideas were suggested but not entirely polished acknowledging the effort and complexity of the matter *"You are lying in your bed, someone comes and gives this little box and says, you know, turn it on. How do you process that, where do you take that. Where do I take that, you are never going to get cured, where?"*. Nick continues *"You do an upbeat. As Chris was saying, you have to be honest. I guarantee you will get better, you won't come to where you were"*. Other design themes and frictions emerged; Nick placed a value on a physical access to audio-media – *"You can listen to radio from any position"* while Sava *"would rather take a flyer"*. This led to thinking about simplification and tangible elements *"You can just have preset buttons, 1,2,3,4"* and importance of audio content over flyers *"I would just like to hear a comforting voice"*. It was important to acknowledge peoples' fears of the possibility of having another stroke. While Sava saw little value of a hospital radio she changed her sentiment throughout the Workshop 5 making a parallel with a reason to come to the charity confirming a sense of belongingness that radio initiative should embody *"That's one of the things that made me sign up to come here was the fear that I might have another stroke"* and as Chris added *"It is just like coming to this group, knowing other people who's got aphasia"*. This exploration highlighted how audio-media, when thoughtfully integrated into these various contexts, could serve as a sanctuary, providing both solace and a sense of community in spaces where it is most needed.

6 Discussion

In envisioning their audio-media futures, people with aphasia saw audio-media as a profound, multifaceted entity intertwined with their lives. Beyond enabling comprehension, co-designers recognised audio-media's deeper roles in shaping social connections, intellectual engagement, and emotional well-being. For them, audio-media was less about passive consumption and more about an active, layered process requiring attention, interpretation, and memory. Their vision included transformative possibilities: radio as a sanctuary, content creation as empowerment, and a medium bridging fluent and disfluent speech. They imagined an audio landscape enriched by personal experiences, embracing the complexity of their interaction with sound and redefining accessibility as inclusion and connection. This is significant for HCI, where audio interventions have often been limited to context-aware systems or recommender tools promoting specific content. Our findings offer a broader, human-centered perspective, reimagining audio-media as an inclusive, transformative space.

6.1 Multi-representational Integration

Hearing perception is a unitary and holistic phenomenon [58], where listening engages more than just the auditory sense. Our findings reveal that people with aphasia frequently mobilise multiple senses to interpret and make sense of the world – i.e., using visual semantic cues such as changing the seating position (e.g., Ethan). These adaptive strategies highlight the interplay of sensory modalities in navigating their everyday experiences. Our study reveals that small, thoughtful accessibility interventions can greatly enhance audio consumption, demonstrating the power of subtle modality adjustments to create meaningful impact, particularly through multi-representational integration [1]. We call for increased attention to other modalities that can enhance audio-only formats offering interesting proposals for future research.

Our findings show that people with aphasia value visual augmentation, yet much of the existing research on visual-graphic support for people with aphasia has been conducted in clinical settings (i.e., speech-language therapy) [17]. These settings often prioritise rapid comprehension and conversation continuity, sometimes requiring individuals to extract key information under time constraints. In contrast, the listening contexts explored in our study prioritise comprehension and engagement, free from the pressure to articulate and/or identify objects.

Discussing IncluScribe's features has revealed that people with aphasia are sensitive to language discrepancies, particularly regarding the precision and clarity of visual cues accompanying the audio content. For example, discussions about image selection sparked a debate among the group members, emphasising the importance of maximising visual accuracy and the option to choose between images and/or icons for people with aphasia [55]. While co-designers recognised the value of image production and representation, it remains unclear how these factors influence auditory comprehension over extended periods in solitary settings (e.g., at home rather than in a group). Furthermore, questions persist regarding the timing of visual interventions, their sustained effectiveness, and the impact of image quality. A key limitation of this study is the unexplored risk of inaccurate AI-generated content, presenting an opportunity for future research to assess its impact on comprehension and strategies for mitigating misrepresentation.

Enhancing audio can also offer significant potential. Research on auditory comprehension, particularly the processing of connected speech in aphasia, is sparse [92]. While treatments targeting language stimulation or cognitive components like attention and memory show promise [92], many individuals report challenges navigating poor acoustic environments [74], a finding echoed in our study. For example, Eliot, a co-designer with severe aphasia and vision impairment, highlighted the transformative potential of personalised audio enhancements as an additional layer for audio comprehension. Enhancing acoustic features in radio and podcasts, however, is no simple task. Segregating music, speech, and non-verbal sounds remains technically challenging [9]. Object-based audio (OBA), which attaches metadata to audio for adaptive playback, offers a promising solution [93], enabling personalisation through adjustments like speech-to-noise ratios and spatial separation [95]. While OBA has improved accessibility for hard-of-hearing audiences [86], its integration within the audio-only platforms in combination with visual accessibility tools remains unexplored. These innovations introduce tensions between personalisation and artistic vision. Dynamic modifications, such as appending visual or audio enhancements, risk distorting the original intent of the media. Misrepresentations or 'hallucinations' could emerge, reshaping how listeners interpret the content, potentially undermining the creative message. Balancing accessibility with fidelity to the media's artistic integrity remains a significant challenge.

Additionally, care must be taken with multiple representations; more is not always better [61]. Our study shows that while these visual representations can take many forms, their value lies in their contextual relevance. For instance, on a train, a simple visual representation might suffice (e.g. Chris - highlighting), while in other settings, mind maps might prove useful (e.g., Nancy). However, the circumstances under which these tools are most effective remain unclear. Future work should explore how best to align visual representations with specific audio contexts and genres. A table might offer precise readability, a phase plot could enhance perceptual inference [1] or mind maps might provide a more intuitive grasp of sequential storytelling (i.e., Nancy's personal experience).

When offering different formats of engagement with audio-only media, it is also crucial to question whether multi-representational formats and visual radio/podcasting are meaningful or risk disrupting deep-rooted cultural habits tied to traditional radio listening. The culture of listening, shaped by loyalty to specific presenters and familiar auditory routines, plays a pivotal role in tailoring future use. Our co-designers exemplified the importance of these habits, emphasising their attachment to long-established radio traditions [12, 13], and yet they have also hinted on the rising trend to consume audio-only formats differently.

Conventional radio, alongside its extensions like podcasts, radio-on-demand, and long formats (e.g., audio dramas and audiobooks), is reshaping audio-media interaction [19]. While audio-media has historically been tied to a passive listening culture-offering sensory relief and ambient engagement [38]-a shift toward integrating visual elements is emerging, aimed at boosting brand awareness and creating immersive experiences [13, 50]. Innovative features

like the German radio station *SWR3*'s smart rewinds and vertical timeshift scales illustrate how younger audiences are embracing interactive controls, mirroring second screen trends in audiovisual media consumption [64, 78]. This evolution aligns with practices like temporal dilation, where platforms like Netflix manipulate the perception of time through binge-watching, challenging traditional linear media norms [18, 26]. For people with aphasia, these visual integrations add cognitive effort but also foster new opportunities for engagement. co-designers demonstrated a readiness to embrace such changes, reflecting a broader shift in listening habits, even among older audiences [5]. Balancing these innovations with the accessibility and cultural heritage of audio-media remains a key design challenge.

6.2 Reformatting and Reducing Cognitive Load

The co-designers frequently emphasised how cognitive load and auditory fatigue complicate comprehension. While the challenge of understanding words and phonemes is well-documented in aphasia, it is far from the only factor at play. While we can highlight a distinction between language and reasoning [91], the disrupted language faculties can still increase the effort required from other cognitive functions, such as attention and memory. This heightened demand often leads to overwhelm and fatigue, particularly affecting short-term memory, as co-designers repeatedly noted. Many reported the need to re-engage with the same content multiple times, reflecting the cumulative toll of processing disruptions on comprehension. For designers, this underscores the critical need to create accessible interventions that are sensitive to the cognitive load of auditory comprehension. While visual cues can make sounds and audio listening more visible, transparent, and graspable, they can also impose cognitive strain [61]. Equally important is ensuring that the audio content and sounds from such media remain predictable, reducing uncertainty and supporting ease of comprehension. In this context, temporal accessibility emerges as a promising alternative. For example, segmenting content into smaller, manageable units provides a practical solution, reducing cognitive strain and enhancing the user's ability to process information effectively.

Exploring AccioBook with the charity's group revealed promising avenues for future research. While our segmentation approach allowed co-designers to isolate speech sounds and reorient their attention in long, spoken dialogues, feedback indicated a desire for more advanced control beyond sentence-by-sentence segmentation. co-designers frequently used repetition to revisit audio, suggesting a need for annotated language complexity to enhance usability. Emerging techniques, such as intuitive audio segmentation interfaces [90], could streamline this process, reducing cognitive effort. Future work should prioritise designing user-friendly, efficient tools that enable people with aphasia to navigate long spoken dialogues effortlessly. The trade-offs between information retention and segmentation require attention. As co-designers noted, working memory challenges [60] often complicate their ability to balance pauses and continuity in extended audio content.

The use of Simplico highlighted another effective strategy. co-designers valued its hierarchical summarisation feature, which distilled long dialogues into simplified content paired with concise keywords. Future research could explore alternative strategies, such as starting with high-level summaries that progressively expand into more detailed layers [51], offering greater flexibility in engaging with complex content. Moreover, the precision and representation of keywords remain a critical area for investigation, particularly in contexts like podcast browsing and keyword searches [76]. Understanding how relational language and keyword selection influence accessibility could unlock new opportunities for intuitive content navigation tailored to the needs of people with aphasia. A significant effort by our co-designers has also been put to emphasise pre-understanding, familiarity and pre-training to consume the audio-content through content summarisation or any technique that can help to understand content prior to listening to reduce the cognitive load. However, the nature of familiarisation and pre-training remains to be explored-could it involve central topical concepts as focal objects, a simplified temporal plot to anchor understanding, or merely naming and briefly explaining key elements as they appear. Additionally, future work should assess generative AI's effectiveness not just in refining image precision but in ensuring linguistic clarity, particularly where precise language is key to comprehension.

Our findings show that gauging language precision, cognitive effort and post-audio reflection for people with complex needs warrants more research. Current audiovisual studies show a tendency to invest little effort via *select and settle* consumer behaviour [37]. However, our study shows that people with aphasia do not shy away from the laborious work of audio comprehension. Liang et al. [52] show high investment in setting specific goals to recommend audio content. However, recommending audio content is different to explicit goal-setting systems and future research should explore the thresholds of the cognitive effort investments as many people in our study reported fatigue and tiredness after exposure to long-stream of audio content to understand articulate or understand speech (i.e., the weight of substitution reported by our co-designers). Furthermore, this opens an avenue for a in-depth exploration of a more systematic execution and realistic scenario of home (in-the-wild) long-term deployment.

Distinguishing between interventions integrated in synchrony with the content and those applied before or after listening emerged as a key consideration for our group. However, achieving seamless interaction requires deeper exploration. While early indications suggest the value of these interventions, their precise implementation and future potential remain unclear. The co-designers showed a preference for post-listening engagement, and investing effort to process and understand content. However, this raises a trade-off: temporal interventions, like synchronised summaries, can help offload cognitive effort but may also introduce a split-attention effect [88], increasing cognitive demands by creating multiple access points. Striking a balance is crucial, as the effort required to enhance understanding could inadvertently impose additional strain. This delicate interplay underscores the need for thoughtful, user-centred design.

6.3 Collective Response to Audio-media Accessibility

Our study highlights the importance of broadening the scope of design interventions. Enhancing and reformatting content are intrinsically tied to the quality and representation of audio-media. While our probes leveraged automated speech recognition (ASR), this technology presents challenges, including noise and inaccuracies. Integrating metadata can improve ASR performance [44], offering a viable pathway to accessibility. Future strategies should focus on equipping creators with guidelines to enrich audio-media (e.g., title, keywords, accessible transcripts) , making it more accessible and inclusive.

Our study illustrates the importance of designing for early stroke experiences and the role that audio-media could play in shaping the identity, fostering a sense of belongingness, managing emotions, false expectations, lack of information, and need for personal expression and adaptation of people with aphasia, an important considerations brought up by people with aphasia. For instance, Nancy's early stroke experiences illustrate this, as she expressed a deep desire for her portable radio while bedridden. She emphasised the challenge of making radio accessible for those confined to bed, in hospitals – often overwhelmed by audiovisual and medical information, or being affected by physical disabilities – yet still yearning for connection to the outside world. Baker et al. [6] emphasise the need for interventions that target sensory communication difficulties, participation in everyday activities, and in particular, a social support during recovery. Co-designers in our study echoed this by proposing a community-led radio initiative [24] to address these needs, much like the digital companions designed for cardiac patients to provide tailored information [42]. Future research should explore the potential for people with aphasia to create or curate aphasia-friendly content that addresses language loss, identity reconstruction, and the challenge of overwhelming stroke-related information.

7 Conclusion

Our study with people living with aphasia advocates for interventions that not only reimagine new accessible audio-media experiences but also establish standardised layers of accessibility – such as captioning, transcription, and other essential visual aids. While we take a broad approach to envision audio-media in its totality, our study opens up several exciting research directions. As new media formats emerge,longstanding accessibility challenges persist, requiring thoughtful solutions tailored to diverse listening contexts -whether live, on-demand, or genre-specific. The four technology probes in this study served as a testbed for assessing new accessibility features, contributing to a deeper understanding of accessibility needs, experiences, and barriers, exploring new roles for audio media, and generating design implications that reimagine audio-media technologies to foster more meaningful and inclusive auditory experiences.

Acknowledgments

We would like to thank our co-designers and the staff at Aphasia ReConnect. This work was funded through an EPSRC DTP partnership and through the CA11y Project via an EPSRC New Investigator Award (EP/X012395/1 – PI: Neate).

References

[1] Shaaron Ainsworth. 2018. *Multiple representations and multimedia learning.* Routledge, London, UK. 96–105 pages.

[2] M. Armstrong, A. Churnside, M.E.F. Melchior, M. Shotton, and M. Brooks. 2014. Object-Based Broadcasting - Curation, Responsiveness and User Experience. In *International Broadcasting Convention (IBC) 2014 Conference.* Institution of Engineering and Technology, Salford, United Kingdom, 8 pages. https://doi.org/10.1049/ib.2014.0038

[3] Michael Armstrong and Maxine Glancy. 2023. *Frameworks for understanding personalisation.* Technical Report WHP 404. BBC, UK. 51 pages.

[4] Shiri Azenkot and Emily Fortuna. 2010. Improving Public Transit Usability for Blind and Deaf-Blind People by Connecting a Braille Display to a Smartphone. In *Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility* (Orlando, Florida, USA) *(ASSETS '10).* Association for Computing Machinery, New York, NY, USA, 317–318. https://doi.org/10.1145/1878803.1878890

[5] Christina Baade. 2023. *Hearing age in music streaming well-being, marketing and older listeners.* Bloomsbury, London, UK, Chapter 3, 44–61.

[6] Caroline Baker, Abby M Foster, Sarah D'Souza, Erin Godecke, Ciara Shiggins, Edwina Lamborn, Lucette Lanyon, Ian Kneebone, and Miranda L Rose. 2022. Management of communication disability in the first 90 days after stroke: a scoping review. *Disability and rehabilitation* 44, 26 (2022), 8524–8538.

[7] Liam Bannon, Jeffrey Bardzell, and Susanne Bødker. 2018. Introduction: Reimagining Participatory Design—Emerging Voices. *ACM Transactions on Computer-Human Interaction* 25, 1 (Feb. 2018), 1–8. https://doi.org/10.1145/3177794

[8] Jeffrey Bardzell, Shaowen Bardzell, and Lone Koefoed Hansen. 2015. Immodest Proposals: Research Through Design and Knowledge. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15).* Association for Computing Machinery, New York, NY, USA, 2093–2102. https://doi.org/10.1145/2702123.2702400

[9] Mathieu Barthet, Steven Hargreaves, and Mark Sandler. 2011. *Speech/Music Discrimination in Audio Podcast Using Structural Segmentation and Timbre Recognition.* Springer Berlin Heidelberg, Berlin, Germany, 138–162. https://doi.org/10.1007/978-3-642-23126-1_10

[10] Jesse Josua Benjamin, Joseph Lindley, Elizabeth Edwards, Elisa Rubegni, Tim Korjakow, David Grist, and Rhiannon Sharkey. 2024. Responding to Generative AI Technologies with Research-through-Design: The Ryelands AI Lab as an Exploratory Study. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference* (Copenhagen, Denmark) *(DIS '24).* Association for Computing Machinery, New York, NY, USA, 1823–1841. https://doi.org/10.1145/3643834.3660677

[11] Greg Benton, Ghazal Fazelnia, Alice Wang, and Ben Carterette. 2020. Trajectory Based Podcast Recommendation. https://doi.org/10.48550/ARXIV.2009.03859

[12] Richard Berry. 2020. Radio, music, podcasts–BBC Sounds: Public service radio and podcasts in a platform world. *Radio Journal: International Studies in Broadcast & Audio Media* 18, 1 (2020), 63–78.

[13] Richard Berry. 2023. *Radio in the Round Reflections on the Future of Sound Media.* Bloomsbury, London, UK, Chapter 30, 504–519.

[14] Kirsten Boehner, Janet Vertesi, Phoebe Sengers, and Paul Dourish. 2007. How HCI interprets the probes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '07).* Association for Computing Machinery, New York, NY, USA, 1077–1086. https://doi.org/10.1145/1240624.1240789

[15] Jordan L. Boyd-Graber, Sonya S. Nikolova, Karyn A. Moffatt, Kenrick C. Kin, Joshua Y. Lee, Lester W. Mackey, Marilyn M. Tremaine, and Maria M. Klawe. 2006. Participatory Design with Proxies: Developing a Desktop-PDA System to Support People with Aphasia. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, New York, NY, USA, 151–160. https://doi.org/10.1145/1124772.1124797

[16] Virginia Braun and Victoria Clarke. 2012. Thematic analysis. In *APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological.* Vol. 2. American Psychological Association, USA, 57–71. https://doi.org/10.1037/13620-004

[17] Jessica Brown and Amber Thiessen. 2018. Using images with individuals with aphasia: Current research and clinical trends. *American Journal of Speech-Language Pathology* 27, 1S (2018), 504–515.

[18] Deborah Chambers. 2019. Emerging temporalities in the multiscreen home. *Media, Culture and Society* 43, 7 (Aug. 2019), 1180–1196. https://doi.org/10.1177/0163443719867851

[19] Sylvia Chan-Olmsted and Rang Wang. 2020. Understanding podcast users: Consumption motives and behaviors. *New Media and Society* 24, 3 (Oct. 2020), 684–704. https://doi.org/10.1177/1461444820963776

[20] Rachel-Ann Charles-Hatt and Thomas Sayers. 2021. Reframing public service radio: The case of BBC Sounds. *Radio Journal: International Studies in Broadcast & Audio Media* 19, 2 (2021), 291–309.

[21] Rachel-Ann Charles-Hatt and Thomas Sayers. 2021. Reframing public service radio: The case of BBC Sounds. *Radio Journal: International Studies in Broadcast and Audio Media* 19, 2 (Oct. 2021), 291–309. https://doi.org/10.1386/rjao_00044_1

[22] Frederic Chaume. 2019. Localizing Media Contents: Technological Shifts, Global and Social Differences and Activism in Audiovisual Translation. In *The Routledge companion to global television*. Routledge, London, UK, 320–331.

[23] Amelia Chelsey. 2021. Is There a Transcript? Mapping Access in the Multimodal Designs of Popular Podcasts. In *Proceedings of the 39th ACM International Conference on Design of Communication* (Virtual Event, USA) *(SIGDOC '21)*. Association for Computing Machinery, New York, NY, USA, 46–53. https://doi.org/10.1145/3472714.3473622

[24] Roberto Cibin, Maurizio Teli, and Sarah Robinson. 2019. Institutioning and Community Radio. A comparative perspective. In *Proceedings of the 9th International Conference on Communities & Technologies - Transforming Communities* (Vienna, Austria) *(C&T '19)*. Association for Computing Machinery, New York, NY, USA, 143–154. https://doi.org/10.1145/3328320.3328392

[25] D. J. Cohen and B. F. Crabtree. 2008. Evaluative Criteria for Qualitative Research in Health Care: Controversies and Recommendations. *The Annals of Family Medicine* 6, 4 (July 2008), 331–339. https://doi.org/10.1370/afm.818

[26] Rebecca Coleman. 2017. Theorizing the present: digital media, pre-emergence and infra-structures of feeling. *Cultural Studies* 32, 4 (Dec. 2017), 600–622. https://doi.org/10.1080/09502386.2017.1413121

[27] Arthur Cropley. 2006. In Praise of Convergent Thinking. *Creativity Research Journal* 18, 3 (2006), 391–404. https://doi.org/10.1207/s15326934crj1803_13

[28] Humphrey Curtis, Ying Hei Lau, and Timothy Neate. 2024. Breaking Badge: Augmenting Communication with Wearable AAC Smartbadges and Displays. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 914, 25 pages. https://doi.org/10.1145/3613904.3642327

[29] Humphrey Curtis, Zihao You, William Deary, Miruna-Ioana Tudoreanu, and Timothy Neate. 2023. Envisioning the (In)Visibility of Discreet and Wearable AAC Devices. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–19. https://doi.org/10.1145/3544548.3580936

[30] Herie B. de Vries and Todd I. Lubart. 2017. Scientific Creativity: Divergent and Convergent Thinking and the Impact of Culture. *The Journal of Creative Behavior* 53, 2 (Dec. 2017), 145–155. https://doi.org/10.1002/jocb.184

[31] Design Council. n.d.. The Double Diamond: A universally accepted depiction of the design process. https://www.designcouncil.org.uk/our-resources/the-double-diamond/ Accessed: 2025-01-30.

[32] Becca Dingman, Garreth W. Tigwell, and Kristen Shinohara. 2021. Designing a Podcast Platform for Deaf and Hard of Hearing Users. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility* (Virtual Event, USA) *(ASSETS '21)*. Association for Computing Machinery, New York, NY, USA, Article 59, 4 pages. https://doi.org/10.1145/3441852.3476523

[33] Patricio Domingues, Ruben Nogueira, José Carlos Francisco, and Miguel Frade. 2020. Post-mortem digital forensic artifacts of TikTok Android App. In *Proceedings of the 15th International Conference on Availability, Reliability and Security*. ACM, New York, NY, USA, 1–8. https://doi.org/10.1145/3407023.3409203

[34] Christina Dunbar-Hester. 2014. *Low power to the people: Pirates, protest, and politics in FM radio activism*. MIT Press, Cambridge, Massachusetts, USA.

[35] Chaoyou Fu, Yuhan Dai, Yongdong Luo, Lei Li, Shuhuai Ren, Renrui Zhang, Zihan Wang, Chenyu Zhou, Yunhang Shen, Mengdan Zhang, Peixian Chen, Yanwei Li, Shaohui Lin, Sirui Zhao, Ke Li, Tong Xu, Xiawu Zheng, Enhong Chen, Rongrong Ji, and Xing Sun. 2024. Video-MME: The First-Ever Comprehensive Evaluation Benchmark of Multi-modal LLMs in Video Analysis. arXiv:2405.21075 [cs.CV] https://arxiv.org/abs/2405.21075

[36] William Gaver. 2012. What should we expect from research through design?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 937–946. https://doi.org/10.1145/2207676.2208538

[37] Maxine Glancy, Lauren Ward, Nick Hanson, Andy Brown, and Michael Armstrong. 2020. Object-Based Media: An Overview of the User Experience. , 16 pages.

[38] Megan E. Graham. 2022. Ambient ageism: Exploring ageism in acoustic representations of older adults in AgeTech advertisements. *Frontiers in Sociology* 7 (Oct. 2022), 14. https://doi.org/10.3389/fsoc.2022.1007836

[39] Brian Grellmann, Timothy Neate, Abi Roper, Stephanie Wilson, and Jane Marshall. 2018. Investigating Mobile Accessibility Guidance for People with Aphasia. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility* (Galway, Ireland) *(ASSETS '18)*. Association for Computing Machinery, New York, NY, USA, 410–413. https://doi.org/10.1145/3234695.3241011

[40] Noor Hammad, Erik Harpstead, and Jessica Hammer. 2023. GameAware Streaming Interfaces. In *Companion Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Stratford, ON, Canada) *(CHI PLAY Companion '23)*. Association for Computing Machinery, New York, NY, USA, 248–253. https://doi.org/10.1145/3573382.3616041

[41] Katerina Hilari and Sarah Northcott. 2016. "Struggling to stay connected": comparing the social relationships of healthy older people and people with stroke and aphasia. *Aphasiology* 31, 6 (Aug. 2016), 674–687. https://doi.org/10.1080/02687038.2016.1218436

[42] Isabel Höppchen, Stefan Tino Kulnik, Alexander Meschtscherjakov, Josef Niebauer, Franziska Pfannerstill, Jan David Smeddinck, Eva-Maria Strumegger, Faith Young, and Daniela Wurhofer. 2024. "Be with me and stay with me": Insights from Co-Designing a Digital Companion to Support Patients Transitioning from Hospital to Cardiac Rehabilitation. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference* (Copenhagen, Denmark) *(DIS '24)*. Association for Computing Machinery, New York, NY, USA, 890–904. https://doi.org/10.1145/3643834.3661633

[43] Maarten Houben, Rens Brankaert, Saskia Bakker, Gail Kenning, Inge Bongers, and Berry Eggen. 2019. Foregrounding Everyday Sounds in Dementia. In *Proceedings of the 2019 on Designing Interactive Systems Conference* (San Diego, CA, USA) *(DIS '19)*. Association for Computing Machinery, New York, NY, USA, 71–83. https://doi.org/10.1145/3322276.3322287

[44] Rosie Jones, Hamed Zamani, Markus Schedl, Ching-Wei Chen, Sravana Reddy, Ann Clifton, Jussi Karlgren, Helia Hashemi, Aasish Pappu, Zahra Nazari, Longqi Yang, Oguz Semerci, Hugues Bouchard, and Ben Carterette. 2021. Current Challenges and Future Directions in Podcast Information Access. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval* (Virtual Event, Canada) *(SIGIR '21)*. Association for Computing Machinery, New York, NY, USA, 1554–1565. https://doi.org/10.1145/3404835.3462805

[45] Shaun K. Kane, Barbara Linam-Church, Kyle Althoff, and Denise McCall. 2012. What We Talk about: Designing a Context-Aware Communication Tool for People with Aphasia. In *Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility* (Boulder, Colorado, USA) *(ASSETS '12)*. Association for Computing Machinery, New York, NY, USA, 49–56. https://doi.org/10.1145/2384916.2384926

[46] Masatomo Kobayashi, Kentarou Fukuda, Hironobu Takagi, and Chieko Asakawa. 2009. Providing synthesized audio description for online videos. In *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, Pennsylvania, USA) *(Assets '09)*. Association for Computing Machinery, New York, NY, USA, 249–250. https://doi.org/10.1145/1639642.1639699

[47] Jill Kries, Pieter De Clercq, Robin Lemmens, Tom Francart, and Maaike Vandermosten. 2023. Acoustic and phonemic processing are impaired in individuals with aphasia. *Scientific Reports* 13, 1 (2023), 11208.

[48] Peter Gall Krogh and Ilpo Koskinen. 2020. *Drifting by Intention: Four Epistemic Traditions from within Constructive Design Research*. Springer International Publishing, New York, NY, USA. https://doi.org/10.1007/978-3-030-37896-7

[49] Raja S. Kushalnagar, Gary W. Behm, Joseph S. Stanislow, and Vasu Gupta. 2014. Enhancing caption accessibility through simultaneous multimodal information: visual-tactile captions. In *Proceedings of the 16th International ACM SIGACCESS Conference on Computers and Accessibility* (Rochester, New York, USA) *(ASSETS '14)*. Association for Computing Machinery, New York, NY, USA, 185–192. https://doi.org/10.1145/2661334.2661381

[50] Tal Laor. 2020. How does it 'sound'? Audiences, broadcasters, and managers on visual radio in Israel. *Convergence: The International Journal of Research into New Media Technologies* 27, 4 (Aug. 2020), 1038–1054. https://doi.org/10.1177/1354856520942404

[51] Daniel Li, Thomas Chen, Albert Tung, and Lydia B Chilton. 2021. Hierarchical Summarization for Longform Spoken Dialog. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '21)*. Association for Computing Machinery, New York, NY, USA, 582–597. https://doi.org/10.1145/3472749.3474771

[52] Yu Liang, Aditya Ponnada, Paul Lamere, and Nediyana Daskalova. 2023. *Enabling Goal-Focused Exploration of Podcasts in Interactive Recommender Systems*. Association for Computing Machinery, New York, NY, USA, 142–155. https://doi.org/10.1145/3581641.3584032

[53] Stephen Lindsay, Katie Brittain, Daniel Jackson, Cassim Ladha, Karim Ladha, and Patrick Olivier. 2012. Empathy, Participatory Design and People with Dementia. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 521–530. https://doi.org/10.1145/2207676.2207749

[54] Stephen Lindsay, Daniel Jackson, Guy Schofield, and Patrick Olivier. 2012. Engaging Older People Using Participatory Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 1199–1208. https://doi.org/10.1145/2207676.2208570

[55] Xiaojuan Ma, Jordan Boyd-Graber, Sonya Nikolova, and Perry R. Cook. 2009. Speaking through pictures: images vs. icons. In *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, Pennsylvania, USA) *(Assets '09)*. Association for Computing Machinery, New York, NY, USA, 163–170. https://doi.org/10.1145/1639642.1639672

[56] Kelly Mack, Emma McDonnell, Dhruv Jain, Lucy Lu Wang, Jon E. Froehlich, and Leah Findlater. 2021. What Do We Mean by "Accessibility Research"? A Literature Survey of Accessibility Papers in CHI and ASSETS from 1994 to 2019. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 371, 18 pages. https://doi.org/10.1145/3411764.3445412

[57] Kelly Mack, Emma McDonnell, Venkatesh Potluri, Maggie Xu, Jailyn Zabala, Jeffrey Bigham, Jennifer Mankoff, and Cynthia Bennett. 2022. Anticipate and Adjust: Cultivating Access in Human-Centered Methods. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 603, 18 pages. https://doi.org/10.1145/3491102.3501882

[58] Lawrence E Marks. 2014. *The unity of the senses: Interrelations among the modalities*. Elsevier, Netherlands.

[59] Hans Martens and Renee Hobbs. 2015. How Media Literacy Supports Civic Engagement in a Digital Age. *Atlantic Journal of Communication* 23, 2 (March 2015), 120–137. https://doi.org/10.1080/15456870.2014.961636

[60] Jamie F Mayer and Laura L Murray. 2002. Approaches to the treatment of alexia in chronic aphasia. *Aphasiology* 16, 7 (2002), 727–743.

[61] Richard E Mayer and Roxana Moreno. 2003. Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist* 38, 1 (2003), 43–52.

[62] Karyn Moffatt, Joanna McGrenere, Barbara Purves, and Maria Klawe. 2004. The Participatory Design of a Sound and Image Enhanced Daily Planner for People with Aphasia. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 407–414. https://doi.org/10.1145/985692.985744

[63] Reg Morris, Alicia Eccles, Brooke Ryan, and Ian I. Kneebone. 2017. Prevalence of anxiety in people with aphasia after stroke. *Aphasiology* 31, 12 (March 2017), 1410–1415. https://doi.org/10.1080/02687038.2017.1304633

[64] Timothy Neate, Michael Evans, and Matt Jones. 2016. Designing Visual Complexity for Dual-screen Media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 475–486. https://doi.org/10.1145/2858036.2858112

[65] Timothy Neate, Abi Roper, Stephanie Wilson, Jane Marshall, and Madeline Cruice. 2020. CreaTable Content and Tangible Interaction in Aphasia. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376490

[66] Alexandre Nevsky, Filip Bircanin, Madeline Cruice, Stephanie Wilson, Elena Simperl, and Timothy Neate. 2024. "I Wish You Could Make the Camera Stand Still": Envisioning Media Accessibility Interventions with People with Aphasia. In *ACM SIGACCESS Conference on Computers and Accessibility* (St. John's, NL, Canada) *(ASSETS '24)*. Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3663548.3675598

[67] Alexandre Nevsky, Timothy Neate, Elena Simperl, and Madeline Cruice. 2024. Lights, Camera, Access: A Closeup on Audiovisual Media Accessibility and Aphasia. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3613904.3641893

[68] Alexandre Nevsky, Timothy Neate, Radu-Daniel Vatavu, and Elena Simperl. 2023. Accessibility Research in Digital Audiovisual Media: What Has Been Achieved and What Should Be Done Next?. In *ACM International Conference on Interactive Media Experiences* (Nantes, France) *(IMX '23)*. Association for Computing Machinery, New York, NY, USA, 1–21. https://doi.org/10.1145/3573381.3596159

[69] Horace M. Newcomb and Paul M. Hirsch. 1983. Television as a Cultural Forum: Implications for Research. *Quarterly Review of Film Studies* 8, 3 (jun 1983), 45–55. https://doi.org/10.1080/10509208309361170

[70] Kristen Nunn, Victoria Tilton-Bolowsky, and Ayelet M Kershenbaum. 2024. Moving Toward Anti-Ableist Practices in Aphasia Rehabilitation and Research: A Viewpoint. *American Journal of Speech-Language Pathology* na, na (2024), 1–11.

[71] OFCOM. 2024. Audio listening in the UK.

[72] Rita Oliveira, Jorge Ferraz de Abreu, and Ana Margarida Almeida. 2011. An approach to identify requirements for an iTV audio description service. In *Proceedings of the 9th European Conference on Interactive TV and Video* (Lisbon, Portugal) *(EuroITV '11)*. Association for Computing Machinery, New York, NY, USA, 227–230. https://doi.org/10.1145/2000119.2000166

[73] Pilar Orero, Marta Brescia-Zapata, and Chris Hughes. 2021. Evaluating subtitle readability in media immersive environments. In *Proceedings of the 9th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-Exclusion* (Online, Portugal) *(DSAI '20)*. Association for Computing Machinery, New York, NY, USA, 51–54. https://doi.org/10.1145/3439231.3440602

[74] Robyn O'Halloran, Brooke Grohn, and Linda Worrall. 2012. Environmental Factors That Influence Communication for Patients With a Communication Disability in Acute Hospital Stroke Units: A Qualitative Metasynthesis. *Archives of Physical Medicine and Rehabilitation* 93, 1 (jan 2012), S77–S85. https://doi.org/10.1016/j.apmr.2011.06.039

[75] Jimin Park, Chaerin Lee, Eunbin Cho, and Uran Oh. 2024. Enhancing the Podcast Browsing Experience through Topic Segmentation and Visualization with Generative AI. In *Proceedings of the 2024 ACM International Conference on Interactive Media Experiences* (Stockholm, Sweden) *(IMX '24)*. Association for Computing Machinery, New York, NY, USA, 117–128. https://doi.org/10.1145/3639701.3656324

[76] Jimin Park, Chaerin Lee, Eunbin Cho, and Uran Oh. 2024. Enhancing the Podcast Browsing Experience through Topic Segmentation and Visualization with Generative AI. In *Proceedings of the 2024 ACM International Conference on Interactive Media Experiences* (Stockholm, Sweden) *(IMX '24)*. Association for Computing Machinery, New York, NY, USA, 117–128. https://doi.org/10.1145/3639701.3656324

[77] Gill Pearl and Madeline Cruice. 2017. Facilitating the Involvement of People with Aphasia in Stroke Research by Developing Communicatively Accessible Research Resources. *Topics in Language Disorders* 37, 1 (Jan. 2017), 67–84. https://doi.org/10.1097/tld.0000000000000112

[78] Marta Perrotta. 2023. *The Listener of the Future: Exploring Public Service Broadcasters' Strategies About Radio Apps*. Bloomsbury, London, UK, Chapter 8, 134–149.

[79] Siddharth Ramanan, Ajay D Halai, Lorna Garcia-Penton, Alistair G Perry, Nikil Patel, Katie A Peterson, Ruth U Ingram, Ian Storey, Stefano F Cappa, Eleonora Catricala, et al. 2023. The neural substrates of transdiagnostic cognitive-linguistic heterogeneity in primary progressive aphasia. *Alzheimer's Research & Therapy* 15, 1 (2023), 219. https://doi.org/10.1186/s13195-023-01350-2

[80] Anni Rander and Peter Olaf Looms. 2010. The accessibility of television news with live subtitling on digital television. In *Proceedings of the 8th European Conference on Interactive TV and Video* (Tampere, Finland) *(EuroITV '10)*. Association for Computing Machinery, New York, NY, USA, 155–160. https://doi.org/10.1145/1809777.1809809

[81] Jacob M. Rigby, Duncan P. Brumby, Anna L. Cox, and Sandy J. J. Gould. 2016. Watching movies on netflix. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. ACM, New York, NY, USA, 714–721. https://doi.org/10.1145/2957265.2961843

[82] Pablo Romero-Fresco. 2019. *Accessible Filmmaking: Integrating translation and accessibility into the filmmaking process*. Routledge, London, UK. https://doi.org/10.4324/9780429053771

[83] Abi Roper, Ian Davey, Stephanie Wilson, Timothy Neate, Jane Marshall, and Brian Grellmann. 2018. Usability Testing - An Aphasia Perspective. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility* (Galway, Ireland) *(ASSETS '18)*. Association for Computing Machinery, New York, NY, USA, 102–106. https://doi.org/10.1145/3234695.3241481

[84] T. A. Rose, L.E. Worrall, L.M. Hickson, Hoffmann, and T.C. 2011. Aphasia friendly written health information: Content and design characteristics. *International Journal of Speech-Language Pathology* 13 (2011), 335–347. Issue 4. https://doi.org/10.3109/17549507.2011.560396

[85] Margarete Sandelowski. 2000. Whatever happened to qualitative description? *Research in Nursing and Health* 23, 4 (2000), 334–340. https://doi.org/10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g

[86] Ben Shirley and Lauren Ward. 2020. Intelligibility versus comprehension: understanding quality of accessible next-generation audio broadcast. *Universal Access in the Information Society* 20, 4 (July 2020), 691–699. https://doi.org/10.1007/s10209-020-00741-8

[87] Pieter Jan Stappers and Elisa Giaccardi. 2017. *Research through Design* (2nd ed.). The Interaction Design Foundation, US, 1–94.

[88] John Sweller. 2020. Cognitive load theory and educational technology. *Educational technology research and development* 68, 1 (2020), 1–16.

[89] Carla Tamburro, Timothy Neate, Abi Roper, and Stephanie Wilson. 2022. Comic Spin: A Comic Creation Tool Enabling Self-expression for People with Aphasia. *ACM Transactions on Accessible Computing* 15, 2 (May 2022), 1–27. https://doi.org/10.1145/3508500

[90] Yi Tang, Chia-Ming Chang, Xi Yang, and Takeo Igarashi. 2023. SyncLabeling: A Synchronized Audio Segmentation Interface for Mobile Devices. *Proc. ACM Hum.-Comput. Interact.* 7, MHCI, Article 226 (sep 2023), 19 pages. https://doi.org/10.1145/3604273

[91] Rosemary A. Varley, Nicolai J. C. Klessinger, Charles A. J. Romanowski, and Michael Siegal. 2005. Agrammatic but Numerate. *Proceedings of the National Academy of Sciences* 102, 9 (Feb. 2005), 3519–3524. https://doi.org/10.1073/pnas.0407470102

[92] Sarah E. Wallace, Janet Patterson, Mary Purdy, Kelly Knollman-Porter, and Patrick Coppens. 2022. Auditory Comprehension Interventions for People With Aphasia: A Scoping Review. *American Journal of Speech-Language Pathology* 31, 5S (oct 2022), 2404–2420. https://doi.org/10.1044/2022_ajslp-21-00297

[93] Tim Walton, Michael Evans, David Kirk, and Frank Melchior. 2018. Exploring object-based content adaptation for mobile audio. *Personal and Ubiquitous Computing* 22, 4 (feb 2018), 707–720. https://doi.org/10.1007/s00779-018-1125-6

[94] Sitong Wang, Zheng Ning, Anh Truong, Mira Dontcheva, Dingzeyu Li, and Lydia B Chilton. 2024. PodReels: Human-AI Co-Creation of Video Podcast Teasers. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference* (Copenhagen, Denmark) *(DIS '24)*. Association for Computing Machinery, New York, NY, USA, 958–974. https://doi.org/10.1145/3643834.3661591

[95] Lauren Ward and Ben Shirley. 2019. Personalization in Object-based Audio for Accessibility: A Review of Advancements for Hearing Impaired Listeners. *Journal of the Audio Engineering Society* 67, 7/8 (Aug. 2019), 584–597. https://doi.org/10.17743/jaes.2019.0021

[96] Lauren Ward, Ben Shirley, and Jon Francombe. 2018. *Accessible Object-Based Audio Using Hierarchical Narrative Importance Metadata.* Technical Report. BBC. https://downloads.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP395.pdf

[97] Maarten Wijnants, Eva Geurts, Hendrik Lievens, Peter Quax, and Wim Lamotte. 2021. Stay Tuned! An Investigation of Content Substitution, the Listener as Curator and Other Innovations in Broadcast Radio. In *Proceedings of the 2021 ACM International Conference on Interactive Media Experiences* (Virtual Event, USA) *(IMX '21)*. Association for Computing Machinery, New York, NY, USA, 120–133. https://doi.org/10.1145/3452918.3458793

[98] Stephanie Wilson, Abi Roper, Jane Marshall, Julia Galliers, Niamh Devane, Tracey Booth, and Celia Woolf. 2015. Codesign for people with aphasia through tangible design languages. *CoDesign* 11, 1 (Jan. 2015), 21–34. https://doi.org/10.1080/15710882.2014.997744

[99] D. Yvette Wohn and Eun-Kyung Na. 2011. Tweeting about TV: Sharing Television Viewing Experiences via Social Media Message Streams. *First Monday* 6, 3 (Feb. 2011), 13 pages. https://doi.org/10.5210/fm.v16i3.3368

[100] Longqi Yang, Yu Wang, Drew Dunne, Michael Sobolev, Mor Naaman, and Deborah Estrin. 2019. More Than Just Words: Modeling Non-Textual Characteristics of Podcasts. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining* (Melbourne VIC, Australia) *(WSDM '19)*. Association for Computing Machinery, New York, NY, USA, 276–284. https://doi.org/10.1145/3289600.3290993

[101] Xinyu Andy Zhao and Crystal Abidin. 2021. TikTok 'Fox Eye' Trend and Everyday Activism: Gen-Z Agency in an Audiovisual Narrative Case Study. In *Proceedings of the 22nd Annual Conference of the Association of Internet Researchers. (AoIR '21')*. University of Illinois Libraries, IL, USA, 5 pages. https://doi.org/10.5210/spir.v2021i0.12267

[102] John Zimmerman and Jodi Forlizzi. 2014. *Research Through Design in HCI.* Springer New York, New York, NY, USA, 167–189. https://doi.org/10.1007/978-1-4939-0378-8_8

Appendix

Table 2: Table that shows the main goals and takeaways of the six workshops.

Workshop Agenda	Main Goal	Session Takeaways
1: Understanding Needs	Exploring auditory comprehension and processing challenges through experiential inquiry to uncover diverse needs and strategies for navigating accessible audio content.	**Major Themes/Takeaways**: *Diversity in auditory needs*: Highlighted varied challenges in processing sounds, including auditory fatigue and difficulty localizing or interpreting sounds. *Compensatory strategies*: co-designers use alternative methods like repositioning or combining sensory inputs to navigate auditory content. **Agenda for Workshop 2**: *Exploring listening contexts and preferences*: Investigate daily listening habits (e.g., devices, environments, formats like podcasts or radio). *Assessing technological barriers*: Understand how current platforms support or hinder accessibility in audio-media. *Evaluating audio format relevance*: Discuss how different formats (live radio, on-demand, or hybrid) meet user needs.
2: Understanding Context	Exploring the relevance of radio in co-designers' lives, understanding typologies of devices and platforms used for audio consumption, and assessing the perceived value of accessible audio-media experiences.	*Contextual Listening*: Examined how auditory comprehension unfolds in real-world interactions with audio-media, including strategies co-designers with aphasia use to correct or adapt their listening practices. *Real-Life Barriers*: First instances of challenges tied to accessibility, such as difficulties for visually impaired co-designers navigating graphics within audio apps. *Initial Future Directions*: Began identifying potential amendments for group sessions to address these challenges more effectively and incorporate early ideas for improving accessibility.
3: Envisioning Audio Futures	Co-envisioning accessible audio-media futures by exploring 'what if' scenarios to creatively address barriers, foster imagination, and identify design opportunities for inclusive interventions.	*Community Radio*: Emphasis on fostering community connection through shared, user-generated stories, and tailored, aphasia-friendly programming. *Playful Radio*: Incorporating engaging and entertaining features such as quizzes and interactive content to encourage participation and enjoyment. *Caring Radio*: Creating content that supports emotional well-being, with elements like soothing voice control and calming or 'healing' radio experiences. *Tailored Accessibility Features*: Exploring recommendations for slow speech, rewinds, and pauses to accommodate diverse user needs. *Communication-Friendly Design*: Prioritising simplicity, summaries, and hosting features to make radio content accessible for people with aphasia.
4: Critiquing Audio Intervention Probes and envisioning audible accessibility	To use exemplary probes as concrete, in-situ prompts for exploring specific accessibility requirements in greater depth, refining abstract ideas into tangible interventions, and setting a collaborative agenda for advancing accessible radio content	*Personalisation is Key*: co-designers valued interventions that allow for customisation, such as selecting image types (symbols, cartoons, emojis) and tailoring playback features (e.g., time-tagging, 'where you left off'). *Multimodal Access*: Visual representations alongside audio content are critical, especially during moments of linguistic struggle, to aid comprehension and reduce cognitive load. *Memory Support*: Tools for repetition, playback, and transcription play a significant role in bridging short-term and long-term memory, facilitating better retention. *Engagement and Motivation*: Affective aids, like quizzes or playful interactions, were seen as highly motivating and aligned with group preferences. *Pacing Content*: Slowing down audio content or introducing tools to control its pace was universally appreciated.
5: Designing Audio Experiences	To design community-driven audio experiences by engaging co-designers with a low-fidelity radio mockup, capturing their ideas for content curation, and envisioning new formats and features for inclusive and accessible radio.	*Affordable Accessibility*: Emphasised the need for low-cost, practical interventions to make accessible media widely available. *Bridging Gaps with Media*: Highlighted the role of media in addressing information gaps and supporting individuals in navigating aphasia-related challenges. *Emotional Comfort and Identity Support*: Recognised the value of familiar voices, personalised content, and stigma-reducing narratives to rebuild identity and provide emotional reassurance. *Aphasia-Friendly Content*: Identified the importance of curating programs tailored for people with aphasia, including motivational tools, coping strategies, and actionable tips.
6: Reflecting on Finalised Accessible Recommendation	To reflect on the insights and interventions developed throughout the workshops, refine key design principles for accessible audio-media, and collaboratively prioritise actionable goals for future co-design efforts and implementation	*Analysis and Synthesis of Findings*: Consolidate insights and feedback gathered throughout the workshops into actionable themes and design principles. *Presentation to the Group*: Share synthesised findings with co-designers to collaboratively decide on future directions. *Focus on "In the Wild" Studies*: Plan for next steps involving real-world testing and application of the developed interventions to validate their practicality and impact.