Tuning in to Personalised Accessible Video: A Home Deployment Study with People with Aphasia

ANONYMOUS AUTHOR(S)

Audiovisual media is increasingly central to everyday information access, entertainment, and social participation. However, this media remains largely inaccessible for people with complex comprehension needs (e.g., aphasia) due to this media's heavy reliance on language-based comprehension. Traditional accessibility features (e.g., captions, audio description) inadequately address these barriers because of their dependence on language processing. In this paper, we conducted the first longitudinal, in-home deployment of audiovisual accessibility features, developing four accessibility features and deploying them for two weeks in the home of nine people with aphasia, allowing us to investigate how these features were appropriated in everyday viewing contexts. Our findings reveal that successful adoption of personalisation requires reducing adoption frictions through careful mapping between interaction and adaptation, and supporting feature appropriation. Further, we detail the nuanced negotiations of accessibility needs between co-viewers. We contribute design implications for accessible content personalisation that considers both individual needs and social viewing dynamics.

CCS Concepts: • **Human-centered computing → Accessibility theory, concepts and paradigms**.

Additional Key Words and Phrases: Accessibility, aphasia, video, deployment, complex communication, probe

ACM Reference Format:

Anonymous Author(s). 20XX. Tuning in to Personalised Accessible Video: A Home Deployment Study with People with Aphasia. In *Title (XYZ 'XX), Month A–Month B, 20XX, City, Country.* ACM, New York, NY, USA, 24 pages. https://doi.org/xxx

1 Introduction

Ensuring that everyone has access to audiovisual media (e.g., TV, video) is crucial. This is particularly true as these media have become pervasive in our everyday lives, being central to the ways we interact with each other and the world around us. We now rely on video to stay informed [75, 103], for entertainment [100], for cultural enrichment [66], social connections [29], and more. While central to our modern lives, video content is not always accessible to all people – those living with certain disabilities, such as those who are blind or low-vision (BLV) or d/Deaf or hard of hearing (DHH), can find video to be inaccessible [56, 93]. Researchers, technologists, and disability activists have developed assistive technologies to address the accessibility needs of these communities, notably captions[1] and audio description [47, 50, 54, 94]. These standard accessibility features, however, rely heavily on language, either written or auditory, and are therefore inaccessible to many people living with complex communication needs (CCNs). Limited prior work has explored ways in which accessibility features for video should be designed for people with CCNs, such as Nevsky et al. [62] co-designing novel accessibility features with people living with aphasia – an acquired impairment of language that can affect or all of the language modalities [6] (understanding, speaking, reading, and/or writing), with the nature and severity varying between individuals, yet does not impact intelligence or reasoning. These features

[1]The term captions will be used to refer to subtitles, captions, and closed captions.

Manuscript submitted to ACM

for viewers with CCNs, however, affect the content in ways more complex than traditional accessibility features – with unknown impacts on long-term and shared viewing experiences. Addressing these gaps requires moving beyond lab-based evaluations to understand how accessibility features function in real-world settings and social contexts.

Therefore, in this paper we present an at-home deployment of a novel content delivery platform with integrated accessibility features specifically designed for people with aphasia. We investigate the impact of these features in situ over 2 weeks with nine participants ($N = 9$) with a range of aphasia profiles. This approach responds to growing calls within the accessibility and HCI communities to conduct in-situ deployments that capture the lived experiences of people with disabilities in their everyday environments [36, 101]. We deployed our system in participants' homes to gain situated insights into how people with aphasia engage with accessibility features that leverage content personalisation, and how these impact the social aspects of in-home viewing. Our findings show that participants valued the ability to adapt content to their accessibility needs, with multi-track volume controls being the most used feature (present during over 90% of watch time), while speed controls were seen as a 'nice-to-have', and spotlight and simplified captions received mixed responses. We also identify interaction latency (i.e., having to leave the content to adjust settings) as an accessibility barrier in its own right, and show how content personalisation becomes socially negotiated during shared viewing. Our main contributions are:

(1) The first in-home deployment of audiovisual accessibility features, capturing how these features are appropriated in everyday contexts
(2) Design implications for personalisation features that support not only individual accessibility, but also shared viewing experiences for people with complex communication needs
(3) A methodological account of how in-home deployments with people with aphasia can surface nuanced, ecologically valid insights into accessibility needs and usage patterns

2 Related Work

2.1 Accessibility in Audiovisual Media

Audiovisual media plays a vital role in cultural participation and expression, serving as a key conduit for social and informational engagement [66, 100]. However, significant accessibility barriers can prevent people from interacting with these media, rendering these media inaccessible to many, including for people who are DHH [50], people who are BLV [37], people with cognitive impairments [38, 88], dementia [28], or language impairments [14, 64]. In response, HCI research has focused on developing and refining technologies to address these challenges, with particular emphasis on improving access for people with hearing and visual impairments through captions and audio descriptions [65].

While initially developed as the standard accessibility features for television viewing [78], they have been adapted for different viewing contexts, devices, and media types. For instance, captions have been adapted to function in 360° immersive video, including when watched using a VR headset [69, 102]. Captions have also been enhanced to provide additional visual cues, such as emotional or tonal information [22, 23]. Additional features have been added, such as dynamically following the speaker [19, 47] or automatically chunking through semantic analysis [48]. Prior research has also explored how captions should be presented and used in novel media formats, such as short form social media video content [44, 56].

Audio descriptions have similarly received significant interest, with research focusing particularly on the preference of BLV viewers for the style and content of the descriptions [37, 68]. Research has also focused on allowing viewers to customise the audio descriptions to better meet their preferences [61], as well as better integrating audio descriptions in new video formats, such as immersive video [40]. Additionally, prior work has explored ways of improving audio description production, such as by involving BLV describers [39] or through LLM-enabled automation [91]. Despite these advances, many other communities are still under-represented in accessibility research, such as individuals with unique communication barriers, for whom these features are not suitable due to their over reliance to language [64].

2.2 Personalisation as Accessibility

Accessibility features often take a "one-size-fits-all" approach which often fail to meet the needs of heterogeneous populations with complex and diverse needs, such as individuals with cognitive and sensory impairments [34], and those living with communication and language difficulties, such as people with aphasia. Aphasia is a language impairment that can affect understanding, speaking, reading, and/or writing, with the nature and severity varying between individuals [6, 59]. These pan-modal challenges can affect comprehension and expression of language across modalities, yet do not impact intelligence, reasoning, problem-solving, or memory [92]. As stroke is the primary cause of aphasia [24], many people with aphasia also experience co-occurring impairments, such as hemiplegia [83] or fatigue [98], which can further influence their interactions with media and viewing experience. Personalisation offers a promising avenue to address the needs of people with aphasia by tailoring audiovisual content to their language, cognitive, and physical contexts. Flexible media – a production method in which audiovisual elements are labelled with metadata and assembled at play time – enables highly personalised viewing experiences that can adapt the complexity to match the viewer's unique needs [2, 18, 70]. This capacity for accessible personalisation allows media to be configured from the outset to meet diverse accessibility needs [35], shifting viewers from passive consumers to active participants in their own experience [43]. For example, allowing the viewer to adapt the volume level of different audio streams for better comprehension [96, 97], or integrating second screens to interact with branching narratives and exercise 'temporal agency' [15].

Personalisation has also been understood as a form of translation – not only between languages, but across semiotic, sensory, and contextual layers of media [16]. Concepts such as localising media to the viewer's context [17], or transcreation, which reshapes the semiotic layers to augment the original [79], are particularly relevant for accessibility. In the context of aphasia, this could involve simplifying or restructuring the language, pairing text with reinforcing imagery such as visual cues, or adjusting the temporal delivery of content to support comprehension. In doing so, the viewer is creating a new piece of media that better serves their own needs without diminishing the original intent. Personalisation can also extend to the viewing environment. For example, relaxed shared viewing setups have supported individuals with autism spectrum disorder (ASD) by adapting the social and physical context of the viewing experience [1]. With technological advances in metadata and content generation, including the use of machine-learning approaches to generatively adapt or create content [53], there is growing potential for targeted and scalable personalisation. While flexible media and other personalisation technologies have been explored for sensory and cognitive accessibility, their systematic application to address the diverse, multimodal needs of people with aphasia remains under-explored, representing a significant gap and opportunity for innovation in accessible audiovisual media design.

Table 1. List and description of the four accessibility features available to the participants.

Features	Description
Simplified Captions	The user can turn captions on or off. When the captions are on, the user can choose to see a version of the captions that replaces verbatim speech with shortened and simplified text.
Spotlight	The user can toggle on spotlighting, which highlights the current speaker on the screen.
Speed Controls	The user can manually adjust the playback rate of the video by 5% increments, or allow the system to automatically slow down the content based on pace and complexity of speech.
Multi-Track Volume Controls	The user can control the volume of the speaker, music, and background audio tracks.

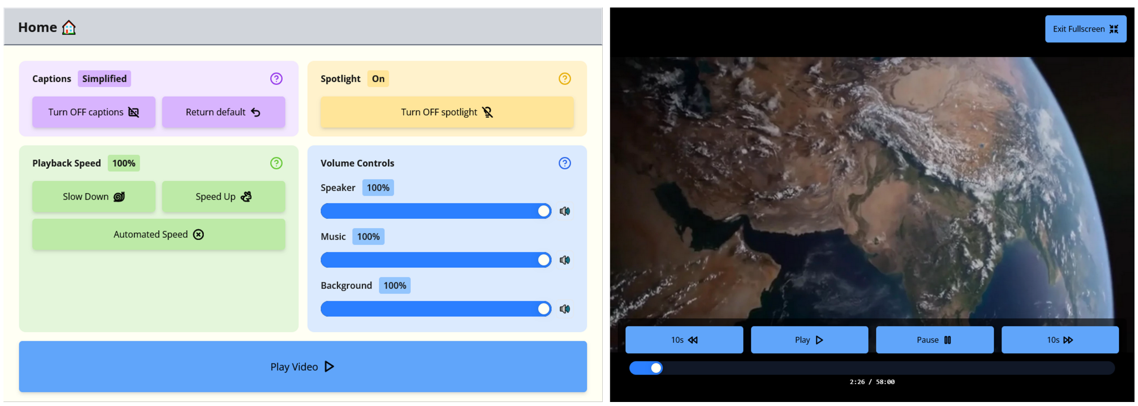

Fig. 1. Screenshots of the system. On the left is the feature controls page and on the right is the video player.

3 Methodology

3.1 System Overview

The system that we co-designed for the deployment took the form of a web application which consisted of a small video library and a custom video player that allowed the user to interact with four accessibility features – see Figure 1 for a screenshot of the controls and Table 1 for the list of features and their description. The video library contained five pieces of video content that we received from the BBC for this study, including a nature documentary, an episode of a travel cooking show, an episode of a talk show, an episode of a comedy TV show, and an episode of a quiz show. The average video duration was 39 minutes (22 minutes to 58 minutes). All participants' actions inside the web application were logged to a database. Additionally, the system had event-trigged ecological momentary assessments (EMAs) [84], asking the participants to assess their experience with our system on a 5-point Likert scale. The participants were given a Samsung Galaxy Tab A9+ to watch the content, as well as a Google Chromecast, which we offered to connect to their TV and allowed them to mirror the tablet display onto that TV. This was done so that all our participants had the same interaction with our system, while also allowing them to watch the video content in a familiar context. We also gave

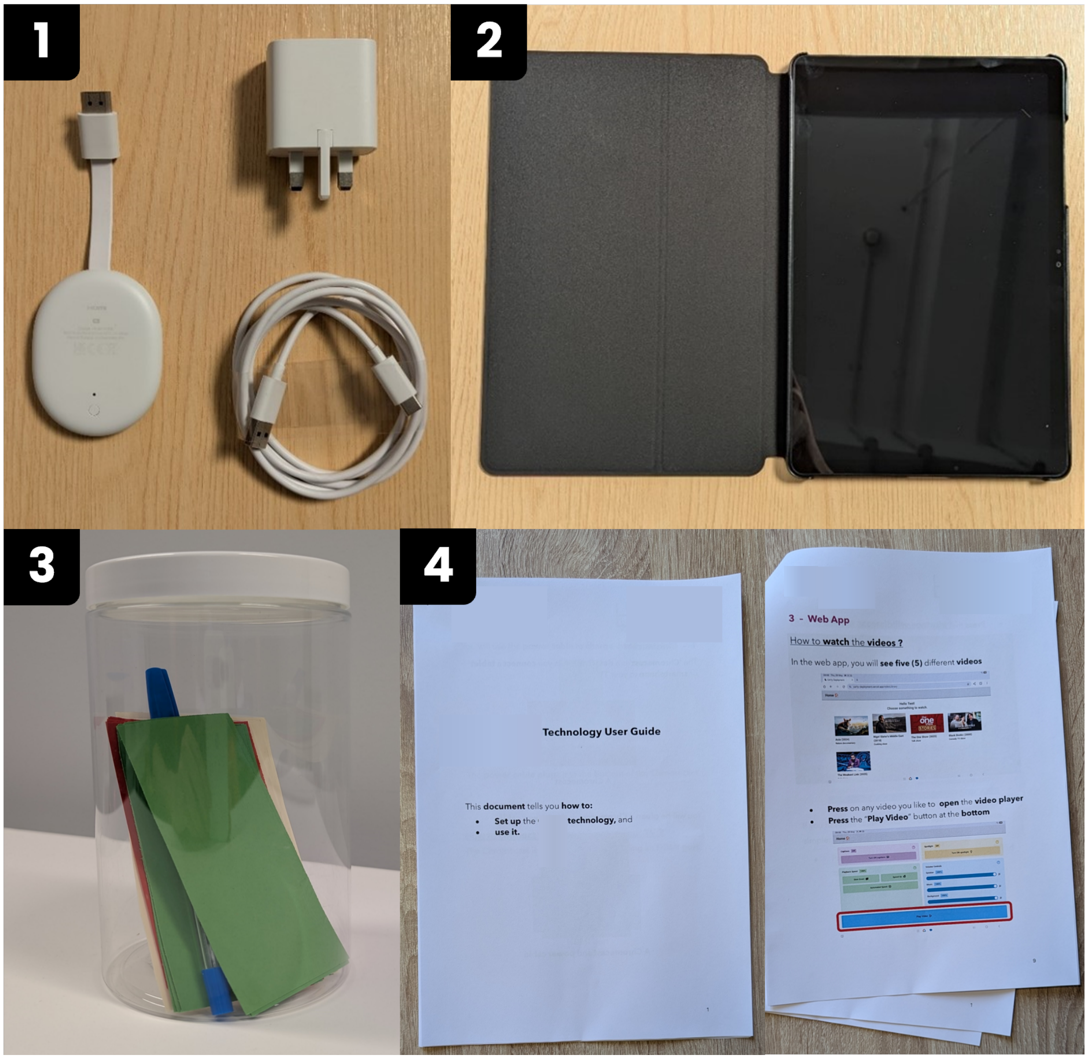

Fig. 2. Equipment given to participants for the deployment: (1) the Chromecast, (2) the tablet, (3) the notes jar, and (4) the user guide.

the participants an aphasia-friendly 'user guide' which had step-by-step instructions on how to use the tablet, how to mirror the tablet screen onto the TV, as well as an explanation of the four accessibility features and how their controls worked. The guide also had our contact information on it, in case the participants experienced any problems during the deployment. See Figure 2 for all materials given to participants.

3.2 Study Design

Prior works successfully co-designed novel accessibility features that adapt audiovisual media through personalisation with people with aphasia [7, 62, 63]. These, however, lacked key insights from participants' real-life viewing experiences as they were done outside the home setting and at distinct time points. Real-world deployments can foster deeper engagement and reflection by allowing participants to interact with systems in the context of their everyday routines [45, 51]. This is particularly important when working with people with aphasia, who often rely on contextual cues and support from familiar environments and people to aid comprehension [42]. These affordances are often absent in lab-based evaluations, limiting the ecological validity of findings. Moreover, studies have demonstrated that the contexts in which people use technology significantly shape how they interact with and make sense of the interactive systems [4, 55]. Embedding technology in real-world contexts can also foster a sense of ownership, reflection, and advocacy among participants – outcomes that are particularly valuable when working with marginalised or under-represented groups [49]. Collecting data through a context-aware and event-triggered method in the home can reveal nuanced, situated experiences that are often inaccessible through traditional lab-based methods [33]. In doing so, our system acted as a knowledge probe [30], enabling us to surface underlying tensions in adapting video content for accessibility and to observe how personalisation practices unfolded in both individual and shared viewing contexts. This approach produced knowledge through the 'annotation' of the system [30], grounded in participants' lived experiences and oriented toward envisioning futures of video personalisation that empower rather than constrain. By conducting an in-home deployment, we enabled participants to familiarise themselves with the technology, explore features at their own pace, and engage with the content in a setting that supports their natural viewing practices, while ensuring ecological validity and capturing the subtle ways in which our participants adapt and respond to the technology.

3.3 Participant Recruitment and Demographics

Table 2. Participant demographics.

Name	Gender	Age	Years w/ Aphasia
P1	Male	71	14
P2	Male	53	12.5
P3	Female	58	50+
P4	Female	73	1.5
P5	Male	44	11
P6	Male	77	3
P7	Male	60	5.5
P8	Male	39	2
P9	Male	70	12

We recruited a total of 9 participants (2F, 7M) to participate in the in-home deployment through <anonymous charity>, a charity for people with aphasia that offers group support sessions, and the <anonymous> mailing list ran through <anonymous university>. We received ethical approval for this research from the <anonymous university> Ethics Board. The participants were recruited through <anonymous charity> during their weekly drop-in sessions in which two of the authors volunteer. We introduced and explained the research to the people present at the drop-in session, providing opportunities for questions and answers. Those who expressed interest in participating in the study were given a flyer and participant information sheet (PIS) prepared in an aphasia accessible format [80], so that they had time to go through it with their support network at their own time. The PIS and all other research materials were designed to be aphasia friendly with the help of a speech and language therapist with 18 years of experience working with people with aphasia, and following appropriate guidelines [3]. When communicating with people with aphasia, supported communication techniques were used [42]. At least a week after the initial contact, those who were still interested in participating met with a member of the research team, accompanied by another person of their choice if they wished. The researcher went through the aphasia-accessible PIS with the potential participant, using supported communication techniques, and answered questions, provided clarification and checked understanding, as needed. Only after that process was a potential participant asked if they wanted to participate. If they indicated that they did, the researcher went through the aphasia-accessible consent form with the person, providing written consent. After the informed consent process was completed, we collected demographic data, as well as information to enable and support the in-home deployment – the participants home address, their contact details, contact details for their next of kin, presence of animals in the home, and any pre-existing medical conditions that the research team should be aware of. This information was kept separate from the research data and the forms were shredded once the deployment was completed. The process was mirrored for those recruited through the <anonymous> mailing list, except the meetings were held online or by phone, with research documents sent by email prior and collected at the first home deployment meeting.

Our participants had an average age of 60.6 (39 – 77) and have had aphasia for an average of 12.4 years (1.5 – 50+) prior to the deployment – see Table 2. They had varied levels of tech proficiency and diverse video content consumption patterns, ranging from only occasionally watching films to spending most of their day watching TV and online video content. All participants were fluent in English prior to their aphasia – with **P3** and **P8** having non-English mother tongues, Luganda and Kirundi respectively. All participants were diagnosed with aphasia following a stroke, except for **P4** who had an intracranial subdural hematoma. Apart from aphasia, **P6** also had dysarthria and dyspraxia, **P7** had cognitive communication disorder, and **P8** had apraxia. Five participants (**P3**, **P4**, **P5**, **P7**, and **P8**) also had right-sided hemiparesis. None of the participants used any high-tech augmentative and alternative communication devices to communicate during our interactions with them. We compensated the participants with 100 GBP in vouchers for a store of their choosing, to thank them for their time commitment.

3.4 Deployment Procedure

Once we recruited our participants, we scheduled a convenient time for them to come to their home and set up the technology. We conducted a pre-deployment semi-structured interview, which we audio recorded only to avoid needlessly invading participants' privacy – this is not best practice when working with people with aphasia, since many rely on multi-model communication. At most two researchers were present for these pre-deployment interview, which were divided into two sections. We first assessed the participants' self-esteem through the 10-item Visual Analogue Self-Esteem Scale (VASES) [11]. We then conducted a semi-structured interview supported by Likert scales, which explored the participants' viewing experiences, video content understanding, how they communicate about their viewing experiences, how they feel while watching video content, and questions on the technologies they use to watch video content, including any accessibility challenges they face. Following the interview, we set up the technology. This involved setting up the Chromecast to the participants' television and walking through the web application. To maximise the likelihood of our participants feeling comfortable using the technology and to reduce potential frustrations, once we walked through the web application and the casting process, we asked them to go through the steps on their own. If they had any usage barriers or questions, these were problem-solved in the moment, and the researchers did not leave until the participants could use technology independently. To support the participants during the post-deployment interview, we gave each participant a plastic 'memory jar' with coloured cards and a pen, which the participants could use to write down their thoughts as they used the system. The cards were in three colours – green for good points, red for bad points, and white for neutral points. To make this process accessible to those who had difficulties with writing, they could also take pictures with their phones that captures moments they wanted to discuss in their post-deployment interviews.

After one week, a researcher contacted each participant to check in and schedule the post-deployment interview and technology pick up. If the participant had any difficulties at the check-in, these were problem-solved with the research team. The post-deployment interview followed the same structure as the pre-deployment one – we first assessed the participants' overall self-esteem using the 10-item VASES, followed by a semi-structured interview. This interview started by asking the same questions as the pre-deployment interview, but focusing on the time when the participants had the video content, followed by exploring their viewing experience when using the system. We then asked questions about the web application, focusing on their overall experience, their thoughts on the accessibility features, which features they found most useful to support their understanding, their experience watching the video content with others, and the ways in which they would have wanted to adapt the features to their own needs. To support the interview, participants also used the notes they took on the coloured papers in the 'memory jar', with some participants' taking more extensive notes in their own notepads.

3.5 Data Analysis

As part of the deployment, we collected both quantitative and qualitative data. The quantitative data included the participants' usage logs, answers to the Likert scale questions, and the self-reported EMA ratings. These were analysed by looking at the participants' feature usage to build an understanding of their viewing behaviour, along with their self-reported experience with the system. To analyse the qualitative data, all interview audio recordings were initially transcribed using the Nvidia Parakeet automatic speech recognition (ASR) model, which is specifically tailored for English transcription and supports various language features [2]. The transcripts were then manually fixed by the first author using NVivo 15 [3], correcting any mistakes generated by the automatic transcription process. These transcripts were then thematically analysed by the first author, following recommendations by Braun and Clarke [10]. The themes identified focused on how participants approached adopting the new accessibility features, how they interacted with the system, the benefits and drawbacks they experienced, and how shared viewing affected their experience. The coding was done inductively, allowing themes to emerge from participants' accounts of how they engaged with the accessibility features.

4 Results

This section presents our finding in two parts. We first report quantitative results, including feature usage patterns and self-reported EMA ratings, to provide an overview of how participants engaged with the system. We then turn to qualitative insights, organised into themes that explain how participants adopted and negotiated accessibility features, the challenges they faced, and the social dynamics that shaped their viewing experiences.

4.1 Quantitative Results

Participants generally found the accessibility features helpful for understanding video content, with all participants finding the features at least somewhat helpful – see Figure 3. The *multi-track volume control* was the most valued

[2]We tested multiple ASR models which failed to capture many nuances of language, such as pauses and filler words common for people with aphasia.
[3]https://lumivero.com/products/nvivo/

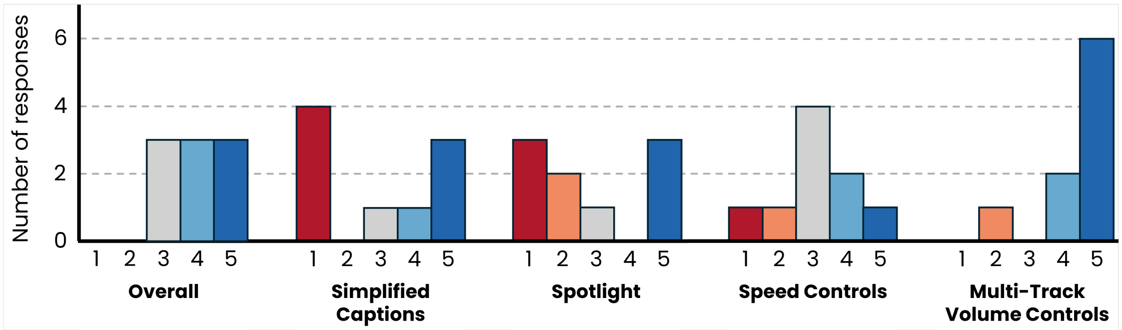

Fig. 3. Histograms of Likert scores for how helpful the overall system was, as well as for each individual accessibility feature.

feature – 8 participants rated it as *helpful* or *very helpful*. This aligns with usage data, as 92.7% of total watch time involved *multi-track volume controls*, primarily lowering background (90.4%) and music (88.6%) volumes – see Figure 5. *Speed controls* were considered a "nice-to-have" rather than essential, with limited impact on viewing experience. The *simplified captions* and *spotlight* features received mixed responses, as reflected by both these features being disabled for more than half of the total watch time (51.3% for captions and 56.2% for spotlight). For the *simplified captions*, 4 participants, who regularly use captions in their daily viewing, found them to be *helpful*, while the remaining 4 stated they were *not at all helpful* since they still found them too difficult to read. *Spotlighting* was seen as *helpful* by 3 participants, but *not at all helpful* by 5, often being described as distracting. Looking at the EMA ratings, participants' experiences fluctuated, increasing slightly over time as they familiarised themselves with the system – see Figure 4.

Analysing the participants' logs, we see that our participants watched an average of 4 hours and 18 minutes of the provided video content across the two week deployment. Over that watch time, participants used the *multi-track volume control* feature the most – see Figure 5. When watching with *multi-track volume controls* enabled, participants lowered the speaker volume to an average 86.9%, the music to 53.4%, and the background sounds to 57.0%. The *speed controls* was the next most used feature, enabled for 65.9% of the total watch time. Participants slightly favoured manual playback control (36.6%) compared to automated controls (29.3%). When using the manual playback controls, the average playback speed was slowed down by 21%. *Simplified captions* were enabled only 25.8% of the time, with standard captions turned on in 22.9%, and captions fully turned off in 51.3% of the time. Reflecting the negative sentiment towards the *spotlight*, this feature was only enabled for 43.8% of the watch time.

Looking at individual participant feature usage, we find that there were three main viewing patterns – passive engagement with the features, constant active engagement, and those who gradually converged on settings that best met their needs. An example of a participant that passively engaged with the features was **P6**, who did not adjust the settings much after initial tweaking – see Figure 6 for exemplar interaction with the settings of the *multi-track volume control* feature. We also see active engagement, such as with **P3**, who constantly adjusted feature settings as she watched the video content. Some participants, such as **P9**, changed the settings over time, trying to identify what settings work best for him.

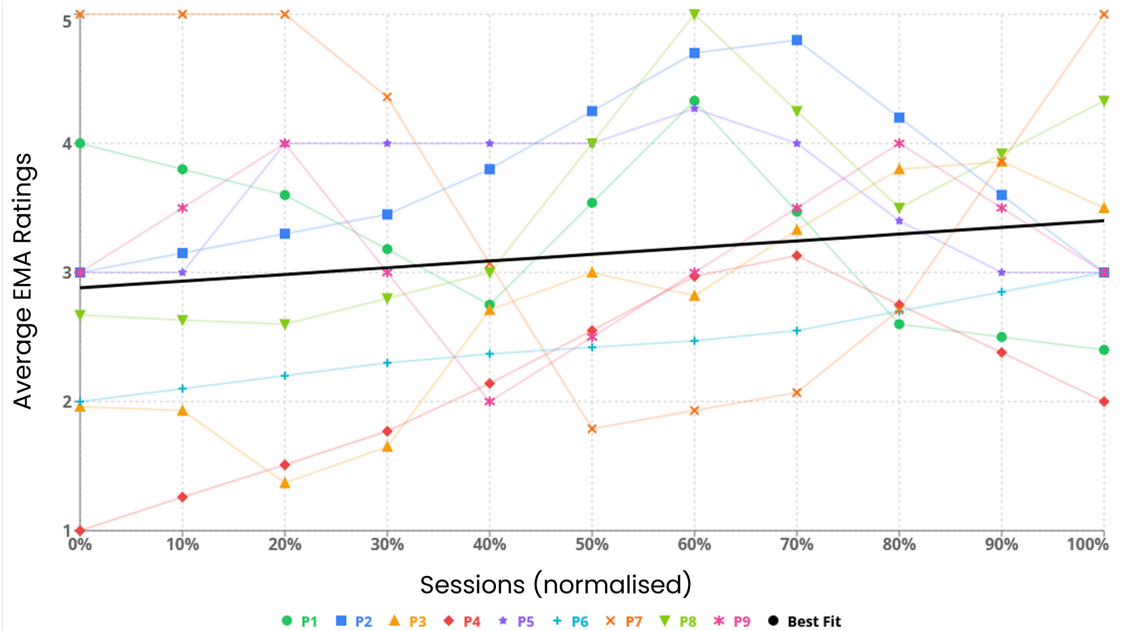

Fig. 4. Participants' self-reported EMA ratings across viewing sessions. The viewing sessions are normalised to prevent participants with anomalous session counts from skewing results.

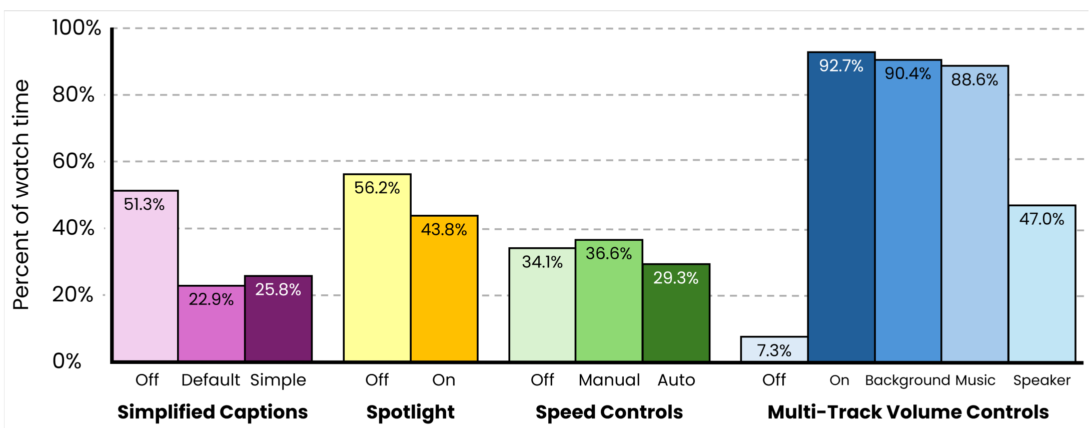

Fig. 5. The total use of the accessibility features and their settings as a percentage of the total watch time.

4.2 Qualitative Results

4.2.1 Patterns of Feature Adoption. When discussing how participants approached using the accessibility features, different patterns emerged, from initial exploration of the different types of content adaptation, to establishing habitual viewing routines.

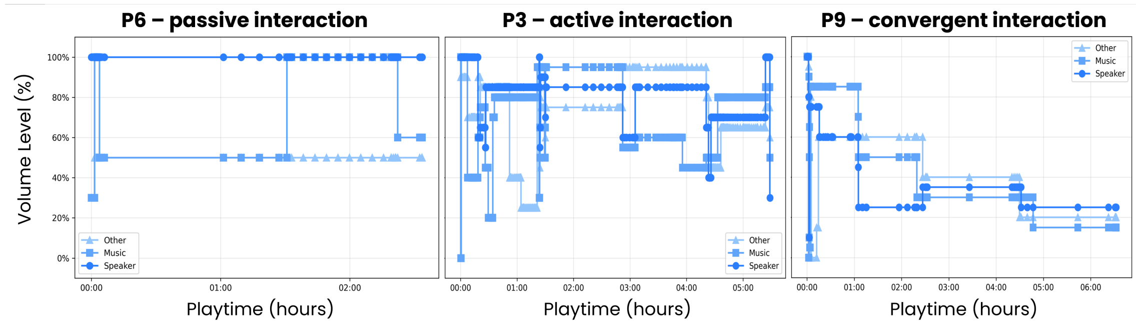

Fig. 6. Line charts showing the settings of the multi-track volume controls for three representative participants (P6, P3, P9), showing how much these participants interacted with the settings across their watch time.

Exploration Prior to Adoption. The first activity all participants did was experimenting with the four accessibility features, looking at what these features did and comparing them to the baseline viewing experience. For instance, **P1** suggested that at the start of the deployment he did "*experiment with various things*", or **P9** trying out the *spotlight* feature "*just to compare*", which is reflected by the initial high level of interaction in Figure 6. This was often done with each piece of content separately, seeing if certain features worked better in different contexts: "***P7**: I tried every program with it on. Didn't find it useful, but did try it on each program*". Often, participants engaged in a pattern of 'trial and error' when exploring the features and their settings, adjusting these for different content to see if they were useful: "***P4**: So, in the end we ended up with, after trial and error, we ended up with 65% [speed]. Spotlight on. Uh... volume, speaker, uh... a hundred percent, music... 50%, background... 50%. And then that's, that's the ones we... we tried*". This initial exploration played an important role in building up the participants' mental models of what the features did, and which ones aligned with their own personal needs and wants. Reflecting on this exploration, **P7** suggested that when using the simplified captions, he "*tried them and I think, no... that's not my disability. So I, um... switched them off*". On the other hand, when it came to the *multi-track volume controls*, **P7** experimented with the different track volumes to identify which track contained certain noises, as he "*hated it with the laughter on*", leading him to "*realise that the music and the laughter track are different things*".

Learning Curve to Adapt to Own Needs. While the initial feature exploration was crucial in understanding which features were useful, there was a learning curve as to how best to select these features, as well as how to navigate the system as a whole. When trying to use the features for the first time, **P6** struggled with knowing which feature did what, and initial exploration was difficult: "*Well, we didn't at all have any success the first time so I phoned up.*". Similarly, **P9** expressed initial confusion when using certain features: "*to use it now... turn off... turn on... sometimes this was a bit confusing*". After using the system and growing more familiar with the different features and the content adaptation they offer, participants grew more confident with using the features they needed: "***P3**: Because when you left... I... I'm... I'm confused. And then again and again and again and then... I got it*". This is somewhat reflected in the increase in EMA ratings over time. This early stage confusion introduced cognitive effort to engage with the complex personalisation options, at the cost of potential increased simplicity of the viewing experience, suggesting there is a need to reduce the initial onboarding complexity: "***P9**: I found it hard to activate it at the start [...] because I didn't know.*".

Curated Feature Use. Once participants grew more familiar with which features, they selected a sub-set of these they saw as facilitating their viewing experience, avoiding others they deemed not useful: "***P1**: Captions, no. It didn't* *seem to make any difference [...] uh... I couldn't see what it was trying to do*". This selection was dependent on which elements of the video content participants thought were necessary, such as **P6** seeing most audio as a distraction: "*Well, the speaker is always on. Music, I don't really care about the music, and the background is... No, I don't care about that.*". Additionally, the features had to provide a benefit to the participants' viewing experience while requiring a justifiable amount of additional work on their end, both interactional and cognitive: "***P5**: It was quite useful... But... But... Um... I don't know... Um... I could have done it, um... I could have done it... I could have done it... Without speeding up, I mean slowing down...*". The selection of which features to use was also dependent on the content: "***P1**: So for the Asia documentary with David Attenborough, his normal pace is [slow]... So slowing it down was a bit too slow.*".

Habitual Viewing Affecting Feature Use. The way participants habitually watch video content also affected what features they chose to incorporate into their viewing pattern. This was most prominent with the use of captions, an accessibility feature many people with aphasia find pointless due to the reading challenges they experience. Even when captions are enabled, many of our participants suggest that they would pause or rewind to be able to read them all: "***P2**: [Captions] on the screen, I didn't use that... because I can't keep up with them. Always known... If it's something that I've got... Um... I would have to go back and... and check... the writing... again*". On the other hand, some participants are comfortable with reading captions, using them at all times: "***P5**: I've always done it... Um... With the subtitles anyway. So, I couldn't ever watch TV without it*". **P8** also suggested that he would not use the *multi-track volume controls*, as his TV comes with a built-in feature that automatically normalises audio to improve speaker clarity: "*I tried [multi-track volume controls]... yes, but, in the end... the... no, no need... volume included... speaker, music and background, it is in... it is already included in... in the program, because when a speaker is... speaking, and the music and the background... it choose automatically*". Additionally, some participants did not want to stray from their habitual viewing pattern, such as not pausing the video to interact with the feature settings on the tablet while watching on the TV, suggesting the interaction should deviate little from habitual viewing: "***P1**: So, I would just watch it on the tablet for the rest of the day.*".

Passive or Active Engagement. There were two main ways our participants engaged with the accessibility features – passive or active engagement. When it came to passive engagement, participants would initially set-up the features in a way that worked for them, and then, such as **P2** stated, "*just leave it*". This viewing pattern was preferred by those participants that did not want to put unnecessary effort into their viewing experience, wanting to simply turn on and watch something: "***P4**: When, once you set it up and so you didn't have to go into it and make alterations, and then go and start watching it again*". This was also reflected in finding the features that facilitated viewing and enabling them on at all times: "***P5**: Way back when... Um... Two weeks [at the beginning of the deployment], so, at the first... But... I... So turning on the spotlight, I use once... But then turn it on and... Um... All the time*". In contrast, other participants embraced the ability to adapt the video content to their needs and actively engaged with the features, such as **P9** who "*[changed settings] quite a lot*". **P1** suggested that when watching, he would "*settled into watching the program I thought, what happens if this, then I've done this*", and then tweak settings based on the specific barrier he was experiencing.

4.2.2 Controls and Interaction Expectations. This theme looks at participants' preferences for how they interacted with content personalisation, and how they envisioned interacting with such a system.

Default to Simple and Intuitive Interactions. Participants consistently emphasised the importance of simplicity and intuitive controls when interacting with the accessibility features. From the beginning, the system needed to clearly convey what each feature did, reducing the cognitive effort required to engage with the system. As **P1** explained, the design of the control itself shaped their perception of agency: "*I suppose [the slider] gives me the impression that I can control it more*". Similarly, **P7** highlighted how the tactile and visual affordances of a slider made the functionality more transparent and appealing: “*Because then I can control what I’m seeing, but the actual feeling of doing a slider rather than a button [...] Because the slider, I can control what I hear [...] Because I can put it all the way down or all the way up or half*”. This sense of control was not merely aesthetic, it influenced whether participants engaged with the features at all, since when the perceived benefit of an interaction was immediately apparent, participants were more likely to explore and adopt the feature: “***P9**: No, it doesn’t think [layered interaction] worked well for me, because I couldn’t understand.*”. Automated defaults or sensible presets were therefore seen as valuable, reducing the initial burden of configuration while still allowing for later personalisation. As **P8** suggested when reflecting on playback speed adjustments: “*Let’s say... this is like normal speed, and then you have numbers that... just like... put like this and then you pick the number, rather than the writing going up and down.*”.

Minimising Interaction Latency. Participants expressed frustration with having to interrupt their viewing experience to adjust the accessibility features – a design choice we made. While the controls themselves were generally considered usable, the need to leave the content and navigate to a separate screen introduced friction that disrupted immersion: “***P1**: It was a bit frustrating. You had to stop... and go back to the... the controls were all frustrating... purely because... you have to go to a separate screen... The controls themselves were okay... but... just having to go back... all the time*”. Similarly, **P3** described the repetitive nature of these interruptions: “*Oh god. And then you go back. And then play. So, if you don’t want... If you don’t want, then click in. This is the questions. [...] And then... Again. And then... press. And then pause. And then you can play.*”. These interaction delays acted as an accessibility barrier in their own right, with participants noting that breaking the flow of viewing was frustrating, even when the features improved comprehension. To mitigate this, participants called for more accessible controls that allowed real-time adjustments without leaving the content: “***P4**: People find it more satisfying if they could actually do the changes as they’re watching, rather than going into the screen and sort of having the screen somehow, you know, as a bit of a sound effect or something*”.

Differing Levels of Control. Once participants became familiar with the accessibility features, some expressed a desire for greater control over the ones they found useful, such as **P1** who suggested adding more granular adjustments to the *spotlight*: “*Um, although I think the highlight could have been more intensified. I suppose you gave it a slider on that as well*”. Reflecting on the *simplified captions*, **P5** suggested more control over the presentation of the captions, such as “*Uh, maybe, um... Big or small letters*”. While some participants were unsure what additional options they might want, they still expressed interest in having the possibility: “***P5**: More [control], but I don’t know what they would be. Um... I mean... I don’t know*”. In contrast, others preferred fewer controls and minimal complexity, emphasising that the existing options were sufficient: “***P6**: This is enough. This is good [...] This is already plenty*”. **P2** warned that adding more audio tracks “*would make it more difficult*”, and **P3** suggesting “*Fewer [volume control tracks], I think speaker and music, that’s it*”. These contrasting preferences highlight a tension between simplicity and flexibility. Participants initially wanted the technology to ‘just work’ with minimal effort required to start using it: “***P4**: I would say, you know, sort of plug and play*”. Designing for this spectrum of needs requires offering a clear, simple default experience, while enabling optional layers of personalisation for those who seek it.

4.2.3 Enhancements and Constraints to Personalised Viewing. This theme explored the ways in which participants’ viewing experience got affected through adapting the content, along with aspects that prevented them from doing so.

Tailoring to Personal Needs. Participants valued the ability to adapt the video content to suit their individual needs and preferences, particularly to reduce distractions and prioritise speech clarity: “***P2**: I could adjust it. It made things, the background, ground noise adjustable. And that was good [...] You know, the, um... things can distract you from hearing.*”. **P5** described how adjusting the relative levels of music and background noise enabled him to focus on the speaker: "*Well... Music... I like it... 30? And background... 25? But, um... it enables me to listen to the speaker. Um... as opposed to music, which is not... that... not that... crucial*". Some participants were willing to invest time in configuring these settings to achieve an optimal experience: "***P6**: I had to pause for 10 minutes on the tablet just before starting a new program*". This was often done at the beginning of a watching session, fine-tuning the settings to the content and the viewing context: "***P3**: 55. Not bad. And then... you can... play... you can... set it up [further]... like this*". Through this adjustment, participants gained a sense of agency, allowing them to make the content 'theirs'.

Enhanced Understandability. Participants reported that the accessibility features helped make content easier to follow, improving both confidence and comfort during viewing, as **P1** suggested when asked if the features improved his understanding: "*Yeah. I might not, but [chuckles] I could watch and listen easily*". For **P2**, the *multi-track volume controls* "*made... made my watching good... better*", and **P3** thought the *speed controls* made the video content "*clear [...] it's clearer for me*". For some, these adaptations enabled experiences that were previously challenging, such as **P4**, who typically struggled to keep up with captions, noted that she "*read all the subtitles on when I was showing them, I got through all the words*". When watching the Weakest Link, a quiz show with fast paced questions, **P4** could slow down the playback speed, and "*then I could understand what they was asking... and I tried to answer*". The *spotlight* also helped **P4** follow the fast paced action: "*Oh, yes, because see, even though you can't understand them, you, you know who's talking*". Reducing distractions was also key to improve focus: "***P8**: The focus... focus to... to... to... to talk or and watch... without directing distractions... no distractions*". Adaptive features can reduce cognitive load and restore a sense of agency, allowing participants to engage more fully with the content rather than expending effort on overcoming accessibility barriers.

Contextual Barriers and Breakdowns. While many participants found the accessibility features helpful, there were instances where the features did not work as expected or failed to meet expectations. **P9** noted confusion when adjusting audio layers and expressed uncertainty about the available options: "*I think there were, like, the background sound and background sound and the speaker*". In some cases, features introduced new issues, as **P1** described how the speaker highlight, while useful, was not always synchronised due to occasional delay caused by the Chromecast [4]: "*That means that when I used the highlighter, it was good, but it wasn't always appearing at the right time [...] counterproductive*". Likewise, **P4** found that slowing down playback sometimes caused unexpected problems: "*The slower you made it, the worse some of the little quirks happened*". Technical difficulties further compounded these frustrations, often leaving participants questioning whether the problem was with the system or themselves, such as **P1**, who asked us if "*someone else had this issue*", and later adding that issues he experienced were his fault: "*Probably, probably means that I didn't set it properly*". These reflections can have a significant negative impact on our participants, ingraining the idea that adopting new technologies is too difficult for them, making it more likely that they give up on adopting useful assistive technologies in the future: "***P6**: I nearly gave up. I was trying to phone up somebody... Not clear. Nearly gave up.*".

Emotional Highs and Lows. Participants' experiences with the accessibility features were often accompanied by strong emotional reactions, ranging from pride and satisfaction to frustration and abandonment. Successfully using the features brought a sense of accomplishment and confidence: "***P3**: I'm very, very... I was so surprised [...] It was nice. It was very nice. I'm proud of myself*". For **P4**, successfully using the accessibility features deepened her engagement with the content and enabled a meaningful moment with her grandson: "*And then [grandson] come in and sat with me and we were talking about what we was watching [...] Yeah, and I, he, he, uh... just sat there and listened to what I was saying, and that, that, that made me happy*". Conversely, moments when participants encountered difficulties led to

[4]This issue was fixed for P3 onwards

frustration, with **P3** stating "*when... first started, it wasn't getting to my head*", and **P4** getting "*very frustrated, because we couldn't get back on it*". In some cases, repeated failures pushed some participants in wanting to abandon using the system altogether: "***P6****: Not clear. I nearly gave up [...] Yes. Very frustrated... I couldn't get it worked.*". These contrasting experiences suggest that there are important emotional stakes when it comes to adopting new accessibility technologies, where failure risks reinforcing feelings of inadequacy and deterring future engagement.

4.2.4 Social Negotiating in Shared Viewing. In this theme we describe how shared viewing experiences shaped the use of content adaptation, how feature use was negotiated, and how communication played a crucial role.

Dependency on Viewing Partner. Participants' ability to use the accessibility features varied, with some drawing on support from others when troubleshooting or navigating the system. This often reflected an asymmetry in perceived competence, particularly when comparing themselves to their viewing partners: "***P4****: I know how to turn the TV on, uh... because I know what button I press. But to, um... get... go... go to any... anything that's not on the normal TV, I can't do, because I don't know how to do it, how to do it [...] because they... they got the controller, and they know how to turn it on. Even my grandchildren know how to put all different things on the, uh... on the, um... the, what's it called? Um, channels*". Indeed, **P4** did not interact with the system at all, preferring to have her husband set up the system for her: "***P4's husband****: P4 and technology, if she'd been on her own, she wouldn't have known what to do, or when to do it*". Other participants similarly described moments when they needed assistance, such as **P3**, who recalled calling us to troubleshoot an issue she was having: "*On Friday, last week, it came dark. And I said, oh my God, where can... What can I do? And then you told me? And then she... He told me*". At the same time, other participants demonstrated that independence was possible with familiarity and practice. For instance, **P6** initially relied on the support from his wife to set up the system to start watching, learning how to use it independently, as his wife describes: "*The last time I wasn't even at home and I walked in and you had the program going. Um, so you had managed perfectly well on your own by then*".

Negotiating Collaborative Viewing Adaption. Watching with a partner influenced how participants engaged with the accessibility features, often requiring negotiation to balance individual needs with shared preferences. For example, **P4**'s husband reflected on wanting to adjust the playback speed to meet his preferences: "*I think if P4 had been at 85% rather than 65%, then it would have been more comfortable for me as well. But, you know, as I say, this was an experiment to see what... What would suit P4, not for me*". These compromises highlight the social dynamics of shared viewing, where accessibility adjustments for one person can affect the viewing experience of others. Participants also made assumptions about their viewing partners' reactions, such as interpreting silence as approval: "***P1****: She... She didn't make comments. I... I... Take it as good*". In other cases, participants defaulted to more 'normal' viewing patterns to avoid disrupting others, even when this meant not using accessibility features that could have improved their experience: "***P8****: With kids... wouldn't simplify [captions]... they're still learning... when interrupts them... causes confusion*". This illustrate the tension between individuals' accessibility needs and collective viewing norms – designing for shared viewing contexts requires accessibility features that are minimally disruptive and transparent to all viewers, enabling accessibility without compromising the social fabric of co-viewing.

Communication Challenges Affecting Social Viewing. Negotiating the use of accessibility features during shared viewing was often complicated by the communication challenges associated with aphasia, as **P5** stated: "*speaking is very hard*". Participants described needing extra time and patience to express their preferences: "***P3****: You've got to be patient [...] because, me I... I can understand it you know... slowly by slowly or slowly... to... to... understand what... what you mean*". These challenges were often compounded by social dynamics in group conversations: "***P6****: I need silence* *because people wouldn't do anything, I can't get a word in*". Communication during shared viewing could also introduce stress and distractions, particularly when participants felt rushed or talked over, with **P8** expressing this frustration: "*I need... to be heard... time... to construct my thought and speak out... don't like people... talking over*". Additionally, **P8** also described how viewing with his young children introduced additional pressures as he was expected to be able to entertain them and answer their questions: "*The moment... what... with the kids, don't focus and too much questions*". Designing for shared viewing contexts requires more than adapting the content, requiring supporting the negotiation process itself.

5 Discussion

Our findings show that while participants valued the ability to adapt content, their experiences revealed two core challenges: reducing the effort required to personalise viewing and managing the social dynamics of doing so in shared contexts. These challenges highlight the distinction between *customisation*, that is user-driven adjustments to existing content, and *personalisation*, where user or system-driven adaptations create new experiences based on viewer data. The following sections explore these themes by first examining how systems can evolve from simple, low-effort defaults to richer personalisation without disrupting viewing, and then considering how accessibility adaptations become socially negotiated in co-viewing scenarios.

5.1 From 'Plug-and-Play' to Personalisation

The first of the two challenges highlighted in our findings concerns the effort required to personalise viewing. Participants wanted systems that were immediately usable, described by **P4** as 'plug-and-play', yet still capable of growing with them over time. This tension between simplicity and flexibility reveals a path for personalisation, beginning with effortless defaults, moving through simple automation, and eventually supporting more advanced adaptation. Thus, early interactions with the system must feel effortless and rewarding to build trust and encourage repeated use. Beginning with default adaptations that 'just work' can reduce the cognitive overhead of first encounters, especially for viewers who are hesitant to engage with unfamiliar technology. Such clear and context-aware defaults can therefore significantly increase the likelihood that personalisation is adopted, without removing viewer choice [13]. Defaults should also be paired with straightforward opt-out mechanisms to preserve autonomy, ensuring that viewers are not locked into configurations they do not want. For example, automatically lowering background music in dialogue-heavy scenes provides an immediate benefit, while still allowing viewers to disengage the feature with minimal effort. Such defaults act as scaffolds that ease viewers into new forms of access, preventing them from reverting to 'normal' viewing patterns that may have lower interaction costs, but remain inaccessible.

While automation can reduce the effort of repeated adjustments, our participants preferred manual control, as long as it offered high controllability – afforded by fast, easy, and fine-grained adjustments. Automated adaptations were seen as useful only when they were predictable, transparent, and easily reversible, aligning with findings on automation and controllability by Roy et al. [81]. Automation could offer suggestions or temporary adjustments, but viewers needed to be able to correct or undo these decisions without breaking the flow of watching. For instance, slowing down a quiz show segment was helpful only if participants could immediately restore normal playback when the pace relaxed. Without this kind of control, automation risked worsening the viewing experience. This echoes patterns seen with other accessibility tools, such as predictive text [77] or speech recognition [82], where users accept automation only if error correction is quick and obvious.

As viewers gained confidence using the system, they often expressed a desire for finer-grained options, such as more nuanced audio controls or larger caption text. Showing this complexity too early, however, risks alienating those who described themselves as 'not technical', a common issue among people with disabilities [25, 32, 67, 72]. Progressive disclosure provides a pathway through this tension [27, 46], supporting both novice and expert users by initially surfacing only simple adjustments and gradually revealing more advanced ones. This staged approach not only prevents overload, but also builds confidence and might reduce abandonment by demonstrating the payoff of personalisation over time [90]. Mixed-initiative approaches extend this by offering context-aware prompts (e.g., recommending a background volume reduction during noisy scenes) while leaving the final decision to the viewer. These prompts can be based on metadata embedded in the content following a flexible media formats, which can provide detailed information about video content to suggest contextually relevant accessibility adjustments. For instance, labelling different audio tracks with their narrative importance could allow such a system to suggest lowering the volume of less narratively important audio tracks during loud moments. Additionally, advancements in small, locally-ran machine learning models could enable learning from viewers' behaviours, suggesting content adaptations in real-time, while preserving viewer privacy and trust by not sharing user data [60, 95]. Since the disclosure is interactive, it can also support shared viewing contexts by prompting the group to discuss potential adjustments, without forcing them into disruptive setup.

Crucially, such scaffolding strategies also address interaction latency – our participants repeatedly emphasised the frustration of having to pause and navigate away from the content to make adjustments. This latency acted as an accessibility barrier in its own right, pushing people away from using features that might otherwise have supported them. Progressive, just-in-time disclosures reduce this disruption by enabling lightweight, real-time adjustments at natural transition points [13]. To be effective, these features must be transparent about why a change is being suggested and how it will help, otherwise they risk eroding trust. They must also remain reversible, giving viewers the option to undo adaptations on the fly. For people with aphasia, experiencing an adaptation is often more effective than imagining its impact [71, 98], making reversible, experiential trials especially valuable.

5.2 Socially Negotiated Personalisation

Accessibility in shared viewing experiences emerges as a relational and collective performance rather than an individual one, where co-viewers mutually produce access through everyday collaboration. When it comes to using accessibility features, especially those that adapt the underlying content, what matters is not only whether features exist, but how they are socially taken up and sustained. Therefore, feature use negotiation is not a failure to be hidden, but a central accessibility practice – personalisation is co-created and interdependent, where one person's adaptation reshapes the experience for everyone else [5, 8, 85]. This negotiated interdependence is also emotionally demanding. Some participants described being self-conscious about the adaptations they were making, worrying they were being burdensome. These emotional costs were often perceived as outweighing the accessibility benefits of personalisation, leading to silent compromise, even when this compromise meant that the viewing experience was significantly worse. The burden of negotiation was intensified by two main factors. First, communication challenges experienced by participants with aphasia made it harder to articulate preferences and repair misunderstandings in the moment. Asking for changes could itself be perceived as difficult. Second, the interactional cost of making adjustments meant that using a feature required pausing and navigating away from the content, breaking the flow of viewing, thus promoting a passive 'default' of inaction. Viewers often abandoned otherwise helpful features when many or slow steps broke the flow of watching, consistent with broader observations that high setup or adjustment costs suppress adoption and persistence [74].

This analysis contributes a re-framing of accessibility negotiation as a legitimate and routine social practice. Designing for shared use, therefore, means reducing both communication and interaction load so the accessible option is not the demanding one. Recognising negotiation in this way repositions accessibility work as 'normal' in homes and public spaces, rather than as an add-on or private workaround. This also helps explain why some participants sometimes chose to watch passively when co-viewing, reverting to normative viewing patterns due to stigma and the internalised pressure of people with complex communication needs to 'fit in'. Prior research has found that people with disabilities broadly avoid assistive technologies in social settings, wanting to avoid being perceived as different [87]. Internalised ableist notions about their abilities compound this effect, leading to significant self-doubt, fear of burdening others, and withdrawal from seeking improved access [12, 41]. Concealment of accessibility can also reinforce the idea that access is exceptional. Our findings suggest a more constructive path by normalising accessibility through being explicit in naming it, placing it alongside standard media controls, and using neutral language that moves away from 'accessibility-only' affordances [86, 87]. This aligns with universal design principles that emphasise creating features that benefit everyone, rather than marking certain functionalities exclusively for people with disabilities [76]. Pullin [76] demonstrates how re-framing assistive technologies as 'aspirational' or desirable consumer features can reduce stigma, as well as positioning devices as premium or fashionable rather than medical can make them more socially acceptable to use. Recent work by Curtis et al. [20, 21] demonstrate how this framing can be put into practice, co-designing AAC devices with people with aphasia that were perceived as empowering and helping reduce stigma. This approach, however, presents a tension around the visibility of access – whilst aspirational framing can reduce stigma, it may also risk making accessibility needs less visible and understood by others [21]. Moreover, unless personalisation is deliberately designed to be appropriable, transparent, and open-ended for people with complex communication needs, universalist framings risk reinforcing normative assumptions about viewing, rather than supporting the diverse, context-specific ways people adapt and negotiate to achieve access [89, 99]. In practice, making the state of the current level of personalisation transparent and mutually legible would help all viewing participants by promoting shared sense-making and reinforcing viewer agency. Indeed, interfaces that offer visibility, awareness, and accountability reduce negotiation overheads because everyone starts from the same information [26]. Small, reversible trials of features (e.g., temporary toggling) allowed participants to experiment with an adjustment without committing to it. This reduced decision anxiety, supported quick repair when adaptations did not work, and anchored accountability without shaming [26].

Our contribution here highlights design strategies for lowering interaction latency. Participants wanted to adjust while watching, making fast, in-flow, and obviously reversible adjustments essential, ensuring that experimentation remains safe. Conversely, delays and mode switching led to abandonment, aligning with evidence that difficulty reconfiguring on-the-fly drives disengagement in accessibility tools [73]. In shared viewing, a collective benefit framing helps avoid issues related to groupware, where one person tends to bear most of the labour for the group's gain [31]. Additionally, assigning the responsibility for the interaction effort solely to the viewer with aphasia mirrors the invisible labour many people with disabilities perform simply to co-exist in shared spaces [9]. A more equitable practice begins with the access needs of the person most at risk of exclusion, then invites co-viewers to accept or soften specific elements of the adaptation, reducing the personal burden of accessibility work – the system must invite this behaviour through its design. Consensus prompts should propose rather than impose, so that co-viewers can collectively share the labour and outcomes without undermining the baseline personalisation [31]. This collective framing can also extend to hardware. For example, acoustically transparent bone-conducting headsets preserve situational awareness and social inclusion while delivering personalised audio enhancements without altering others' experiences [57]. Such devices further reduce stigma and the self-perceived burden associated with using personalisation for accessibility [58].

5.3 Limitations and Future Work

This study has several limitations that inform directions for future work. First, our sample was small and lacked certain diversity – only 2 of our participants were women, the age skewed towards older adults, and our participants come from 'Western, Educated, Industrialised, Rich and Democratic' (WEIRD) [52] countries. This reflects the practical challenges of recruiting for in-home deployments involving people with aphasia. This was made harder as we excluded individuals who reported no difficulty watching video, which allowed us to focus on those most likely to benefit. Future research should explore broader recruitment strategies and extend the scope to other groups with complex communication needs, for whom similar personalisation features may be valuable. Second, insights into social viewing were drawn from a subset of participants who used the system with others, highlighting the need for studies explicitly designed to examine co-viewing dynamics. Third, our prototype was intentionally minimal, serving as a probe rather than a fully featured product. Future iterations should incorporate the design implications identified in our discussion, including progressive disclosure, in-flow adjustments, and transparency, while expanding the feature set to support richer personalisation and appropriation over time. Finally, while our two-week deployment provided situated insights, it could not capture long-term patterns of adoption or abandonment. Longer, less researcher intensive deployments – potentially through releasing the system as an open tool – would allow us to study sustained engagement and evolving practices in everyday contexts.

6 Conclusion

Ensuring access to audiovisual media for people with complex communication needs requires approaches that move beyond language-based accessibility features. We addressed this challenge by designing and deploying a custom video player that enabled four personalisation features in nine participants' homes for two weeks, allowing them to interact with four features based on prior envisioning work – simplified captions, speaker spotlight, playback speed controls, and multi-track volume controls. This approach provided ecologically valid and situated insights, especially on the interaction costs and their role in shaping the adoption of personalisation, how accessibility features are appropriated in everyday viewing routines, and how shared viewing introduces negotiation and compromise. These findings highlight that leveraging content personalisation for the purpose of accessibility is not only about feature availability, but also about reducing initial adoption frictions, supporting feature appropriation, and embedding personalisation into social viewing practices. We identify that future personalisation systems should prioritise having effortless and context-aware defaults, progressive personalisation affordances, minimise interaction latency, and make content adaptations transparent and socially legible. Advances in flexible media formats and lightweight locally-ran machine-learning models offer promising directions for enabling real-time, context-aware adjustments that respect creative intent, while making accessibility an ordinary part of viewing experiences.

References

[1] Meryl Alper. 2021. Critical media access studies: Deconstructing power, visibility, and marginality in mediated space. *International Journal of Communication* 15 (2021), 22.

[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] Stroke Association. 2019. Lived experience of stroke report. https://www.stroke.org.uk/lived-experience-of-stroke-report

[4] Erin Beneteau, Ashley Boone, Yuxing Wu, Julie A. Kientz, Jason Yip, and Alexis Hiniker. 2020. Parenting with Alexa: Exploring the Introduction of Smart Speakers on Family Dynamics. 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, 1–13. https://doi.org/10.1145/3313831.3376344

[5] Cynthia L. Bennett, Erin Brady, and Stacy M. Branham. 2018. Interdependence as a Frame for Assistive Technology Research and Design. 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, 161–173. https://doi.org/10.1145/3234695.3236348

[6] Karianne Berg, Jytte Isaksen, Sarah J. Wallace, Madeline Cruice, Nina Simmons-Mackie, and Linda Worrall. 2020. Establishing consensus on a definition of aphasia: an e-Delphi study of international aphasia researchers. *Aphasiology* 36, 4 (Dec. 2020), 385–400. https://doi.org/10.1080/02687038.2020.1852003

[7] 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 *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Yokahama, Japan) *(CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 277, 22 pages. https://doi.org/10.1145/3706598.3714000

[8] Stacy M. Branham and Shaun K. Kane. 2015. Collaborative Accessibility: How Blind and Sighted Companions Co-Create Accessible Home Spaces. 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, 2373–2382. https://doi.org/10.1145/2702123.2702511

[9] Stacy M. Branham and Shaun K. Kane. 2015. The Invisible Work of Accessibility: How Blind Employees Manage Accessibility in Mixed-Ability Workplaces. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility* (Lisbon, Portugal) *(ASSETS '15)*. Association for Computing Machinery, New York, NY, USA, 163–171. https://doi.org/10.1145/2700648.2809864

[10] 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*. American Psychological Association, USA, 57–71. https://doi.org/10.1037/13620-004

[11] Shelagh M. Brumfitt and Paschal Sheeran. 1999. The development and validation of the Visual Analogue Self-Esteem Scale (VASES). *British Journal of Clinical Psychology* 38, 4 (Nov. 1999), 387–400. https://doi.org/10.1348/014466599162980

[12] Fiona A. Kumari Campbell. 2008. Exploring internalized ableism using critical race theory. *Disability and Society* 23, 2 (March 2008), 151–162. https://doi.org/10.1080/09687590701841190

[13] Ana Caraban, Evangelos Karapanos, Daniel Gonçalves, and Pedro Campos. 2019. 23 Ways to Nudge: A Review of Technology-Mediated Nudging in Human-Computer Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3290605.3300733

[14] Jade Cartwright and Kym A. E. Elliott. 2009. Promoting strategic television viewing in the context of progressive language impairment. *Aphasiology* 23, 2 (2009), 266–285. https://doi.org/10.1080/02687030801942932

[15] Deborah Chambers. 2021. Emerging temporalities in the multiscreen home. *Media, Culture and Society* 43 (2021), 1180–1196. Issue 7. https://doi.org/10.1177/0163443719867851

[16] Frederic Chaume. 2018. Is audiovisual translation putting the concept of translation up against the ropes? *Journal of Specialised Translation* 30 (2018), 84–104. Issue 30.

[17] Frederic Chaume. 2019. *Localizing media contents: Technological shifts, global and social differences and activism in audiovisual translation* (1st edition ed.). Routledge, New York, NY, USA, 320–331.

[18] Shauna Concannon, Natasha Rajan, Parthiv Shah, Davy Smith, Marian Ursu, and Jonathan Hook. 2020. Brooke Leave Home: Designing a Personalized Film to Support Public Engagement with Open Data. 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/10.1145/3313831.3376462

[19] Michael Crabb, Rhianne Jones, Mike Armstrong, and Chris J. Hughes. 2015. Online News Videos: The UX of Subtitle Position. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers and Accessibility* (Lisbon, Portugal) *(ASSETS '15)*. Association for Computing Machinery, New York, NY, USA, 215–222. https://doi.org/10.1145/2700648.2809866

[20] Humphrey Curtis, Adam D G Jenkins, Seray B Ibrahim, and Timothy Neate. 2024. Looking Past Screens: Exploring Mixed Reality and Discreet AAC Devices. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility* (St. John's, NL, Canada) *(ASSETS '24)*. Association for Computing Machinery, New York, NY, USA, Article 42, 22 pages. https://doi.org/10.1145/3663548.3675655

[21] 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* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 52, 19 pages. https://doi.org/10.1145/3544548.3580936

[22] Caluã de Lacerda Pataca, Saad Hassan, Lloyd May, Michelle M Olson, Toni D'aurio, Roshan L Peiris, and Matt Huenerfauth. 2025. Tactile Emotions: Multimodal Affective Captioning with Haptics Improves Narrative Engagement for d/Deaf and Hard-of-Hearing Viewers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 68, 17 pages. https://doi.org/10.1145/3706598.3713304

[23] Caluã de Lacerda Pataca, Matthew Watkins, Roshan Peiris, Sooyeon Lee, and Matt Huenerfauth. 2023. Visualization of Speech Prosody and Emotion in Captions: Accessibility for Deaf and Hard-of-Hearing Users. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 831, 15 pages. https://doi.org/10.1145/3544548.3581511

[24] Charles Ellis, Rose Y. Hardy, Richard C. Lindrooth, and Richard K. Peach. 2017. Rate of aphasia among stroke patients discharged from hospitals in the United States. *Aphasiology* 32, 9 (Oct. 2017), 1075–1086. https://doi.org/10.1080/02687038.2017.1385052

[25] Kirsten Ellis, Emily Dao, Osian Smith, Stephen Lindsay, and Patrick Olivier. 2021. TapeBlocks: A Making Toolkit for People Living with Intellectual Disabilities. 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 280, 12 pages. https://doi.org/10.1145/3411764.3445647

[26] Thomas Erickson and Wendy A. Kellogg. 2000. Social translucence: an approach to designing systems that support social processes. *ACM Trans. Comput.-Hum. Interact.* 7, 1 (March 2000), 59–83. https://doi.org/10.1145/344949.345004

[27] Leah Findlater and Joanna McGrenere. 2004. A comparison of static, adaptive, and adaptable menus. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vienna, Austria) *(CHI '04)*. Association for Computing Machinery, New York, NY, USA, 89–96. https://doi.org/10.1145/985692.985704

[28] Liam Funnell, Isabel Garriock, Ben Shirley, and Tracey Williamson. 2019. Dementia-friendly design of television news broadcasts. *Journal of Enabling Technologies* 13 (9 2019), 137–149. Issue 3. https://doi.org/10.1108/JET-02-2018-0009

[29] Walter Gantz and Lawrence A Wenner. 1995. Fanship and the Television Sports Viewing Experience. *Sociology of sport Journal* 12, 1 (1995), 56–74.

[30] William Gaver. 2012. What should we expect from research through design?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 937–946. https://doi.org/10.1145/2207676.2208538

[31] Jonathan Grudin. 1994. Groupware and social dynamics: eight challenges for developers. *Commun. ACM* 37, 1 (Jan. 1994), 92–105. https://doi.org/10.1145/175222.175230

[32] Laxmi Gunupudi, Maryam Bandukda, Giulia Barbareschi, Tigmanshu Bhatnagar, Aanchal Singh, Satish Mishra, Amit Prakash, and Catherine Holloway. 2024. Scaffolding Digital Literacy Through Digital Skills Training for Disabled People in the Global South. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility* (St. John's, NL, Canada) *(ASSETS '24)*. Association for Computing Machinery, New York, NY, USA, Article 29, 14 pages. https://doi.org/10.1145/3663548.3675666

[33] Alexis Hiniker, Jon E. Froehlich, Mingrui Zhang, and Erin Beneteau. 2019. Anchored Audio Sampling: A Seamless Method for Exploring Children's Thoughts During Deployment Studies. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300238

[34] Anthony Hornof, Haley Whitman, Marah Sutherland, Samuel Gerendasy, and Joanna McGrenere. 2017. Designing for the "Universe of One": Personalized Interactive Media Systems for People with the Severe Cognitive Impairment Associated with Rett Syndrome. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 2137–2148. https://doi.org/10.1145/3025453.3025904

[35] Elfed Howells and David Jackson. 2021. Object-Based Media Report. https://www.ofcom.org.uk/research-and-data/technology/general/object-based-media

[36] Katherine Isbister, Peter Cottrell, Alessia Cecchet, Ella Dagan, Nikki Theofanopoulou, Ferran Altarriba Bertran, Aaron J. Horowitz, Nick Mead, Joel B. Schwartz, and Petr Slovak. 2022. Design (Not) Lost in Translation: A Case Study of an Intimate-Space Socially Assistive "Robot" for Emotion Regulation. *ACM Trans. Comput.-Hum. Interact.* 29, 4, Article 32 (March 2022), 36 pages. https://doi.org/10.1145/3491083

[37] Lucy Jiang, Crescentia Jung, Mahika Phutane, Abigale Stangl, and Shiri Azenkot. 2024. "It's Kind of Context Dependent": Understanding Blind and Low Vision People's Video Accessibility Preferences Across Viewing Scenarios. 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, Article 897, 20 pages. https://doi.org/10.1145/3613904.3642238

[38] Lucy Jiang, Woojin Ko, Shirley Yuan, Tanisha Shende, and Shiri Azenkot. 2025. Shifting the Focus: Exploring Video Accessibility Strategies and Challenges for People with ADHD. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 561, 16 pages. https://doi.org/10.1145/3706598.3713637

[39] Lucy Jiang and Richard Ladner. 2022. Co-Designing Systems to Support Blind and Low Vision Audio Description Writers. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility* (Athens, Greece) *(ASSETS '22)*. Association for Computing Machinery, New York, NY, USA, Article 74, 3 pages. https://doi.org/10.1145/3517428.3550394

[40] Lucy Jiang, Mahika Phutane, and Shiri Azenkot. 2023. Beyond Audio Description: Exploring 360° Video Accessibility with Blind and Low Vision Users Through Collaborative Creation. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (New York, NY, USA) *(ASSETS '23)*. Association for Computing Machinery, New York, NY, USA, Article 50, 17 pages. https://doi.org/10.1145/3597638.3608381

[41] Ásta Jóhannsdóttir, Snæfríður Þóra Egilson, and Freyja Haraldsdóttir. 2022. Implications of internalised ableism for the health and wellbeing of disabled young people. *Sociology of Health and Illness* 44, 2 (Jan. 2022), 360–376. https://doi.org/10.1111/1467-9566.13425

[42] Aura Kagan. 1998. Supported conversation for adults with aphasia: methods and resources for training conversation partners. *Aphasiology* 12, 9 (Sept. 1998), 816–830. https://doi.org/10.1080/02687039808249575

[43] Hyunjin Kang and S. Shyam Sundar. 2016. When Self Is the Source: Effects of Media Customization on Message Processing. *Media Psychology* 19, 4 (2016), 561–588. https://doi.org/10.1080/15213269.2015.1121829

[44] Daniel Killough and Amy Pavel. 2023. Exploring Community-Driven Descriptions for Making Livestreams Accessible. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (New York, NY, USA) *(ASSETS '23)*. Association for Computing Machinery, New York, NY, USA, Article 42, 13 pages. https://doi.org/10.1145/3597638.3608425

[45] Nanum Kim, Sangsu Jang, Hansol Kim, Dayoung Shin, and Young-Woo Park. 2025. Lino: An Interactive System for Daily Mood Recordings Supporting Meaning-Making through Single Stroke Drawing Approach. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference* (Madeira, Portugal) *(DIS '25)*. Association for Computing Machinery, New York, NY, USA, 2255–2269. https://doi.org/10.1145/3715336.3735675

[46] Junhan Kong, Jacob O. Wobbrock, Tianyuan Cai, and Zoya Bylinskii. 2025. Supporting Mobile Reading While Walking with Automatic and Customized Font Size Adaptations. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Yokahama, Japan) *(CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 1007, 13 pages. https://doi.org/10.1145/3706598.3713367

[47] Kuno Kurzhals, Fabian Göbel, Katrin Angerbauer, Michael Sedlmair, and Martin Raubal. 2020. A View on the Viewer: Gaze-Adaptive Captions for Videos. 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, 1–12. https://doi.org/10.1145/3313831.3376266

[48] Dawon Lee, Jongwoo Choi, and Junyong Noh. 2025. OptiSub: Optimizing Video Subtitle Presentation for Varied Display and Font Sizes via Speech Pause-Driven Chunking. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 799, 12 pages. https://doi.org/10.1145/3706598.3714199

[49] Chu Li, Katrina Oi Yau Ma, Michael Saugstad, Kie Fujii, Molly Delaney, Yochai Eisenberg, Delphine Labbé, Judy L Shanley, Devon Snyder, Florian P P Thomas, and Jon E. Froehlich. 2024. "I never realized sidewalks were a big deal": A Case Study of a Community-Driven Sidewalk Accessibility Assessment using Project Sidewalk. 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, Article 969, 18 pages. https://doi.org/10.1145/3613904.3642003

[50] Franklin Mingzhe Li, Cheng Lu, Zhicong Lu, Patrick Carrington, and Khai N. Truong. 2022. An Exploration of Captioning Practices and Challenges of Individual Content Creators on YouTube for People with Hearing Impairments. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW1, Article 75 (apr 2022), 26 pages. https://doi.org/10.1145/3512922

[51] Martin Valdemar Anker Lindrup, Arjun Rajendran Menon, and Aksel Biørn-Hansen. 2023. Carbon Scales: Collective Sense-making of Carbon Emissions from Food Production through Physical Data Representation. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (Pittsburgh, PA, USA) *(DIS '23)*. Association for Computing Machinery, New York, NY, USA, 1515–1530. https://doi.org/10.1145/3563657.3596043

[52] Sebastian Linxen, Christian Sturm, Florian Brühlmann, Vincent Cassau, Klaus Opwis, and Katharina Reinecke. 2021. How WEIRD is CHI?. 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 143, 14 pages. https://doi.org/10.1145/3411764.3445488

[53] Yixin Liu, Kai Zhang, Yuan Li, Zhiling Yan, Chujie Gao, Ruoxi Chen, Zhengqing Yuan, Yue Huang, Hanchi Sun, Jianfeng Gao, Lifang He, and Lichao Sun. 2024. Sora: A Review on Background, Technology, Limitations, and Opportunities of Large Vision Models. https://doi.org/10.48550/ARXIV.2402.17177

[54] Mariana Lopez, Gavin Kearney, and Krisztian Hofstadter. 2021. Enhancing Audio Description: Inclusive Cinematic Experiences Through Sound Design. *Journal of Audiovisual Translation* 4, 1 (Oct. 2021), 157–182. https://doi.org/10.47476/jat.v4i1.2021.154

[55] Michal Luria, Rebecca Zheng, Bennett Huffman, Shuangni Huang, John Zimmerman, and Jodi Forlizzi. 2020. Social Boundaries for Personal Agents in the Interpersonal Space of the Home. 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, 1–12. https://doi.org/10.1145/3313831.3376311

[56] Emma J McDonnell, Tessa Eagle, Pitch Sinlapanuntakul, Soo Hyun Moon, Kathryn E. Ringland, Jon E. Froehlich, and Leah Findlater. 2024. "Caption It in an Accessible Way That Is Also Enjoyable": Characterizing User-Driven Captioning Practices on TikTok. 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, Article 492, 16 pages. https://doi.org/10.1145/3613904.3642177

[57] Mark McGill, Stephen Brewster, David McGookin, and Graham Wilson. 2020. Acoustic Transparency and the Changing Soundscape of Auditory Mixed Reality. 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, 1–16. https://doi.org/10.1145/3313831.3376702

[58] Mark McGill, Florian Mathis, Mohamed Khamis, and Julie Williamson. 2020. Augmenting TV Viewing using Acoustically Transparent Auditory Headsets. In *Proceedings of the 2020 ACM International Conference on Interactive Media Experiences* (Cornella, Barcelona, Spain) *(IMX '20)*. Association for Computing Machinery, New York, NY, USA, 34–44. https://doi.org/10.1145/3391614.3393650

[59] 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

[60] Kalyan Nakka, Jimmy Dani, and Nitesh Saxena. 2024. Is On-Device AI Broken and Exploitable? Assessing the Trust and Ethics in Small Language Models. https://doi.org/10.48550/arxiv.2406.05364

[61] Rosiana Natalie, Ruei-Che Chang, Smitha Sheshadri, Anhong Guo, and Kotaro Hara. 2024. Audio Description Customization. In *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility* (St. John's, NL, Canada) *(ASSETS '24)*. Association for Computing Machinery, New York, NY, USA, Article 39, 19 pages. https://doi.org/10.1145/3663548.3675617

[62] 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

[63] Alexandre Nevsky, Filip Bircanin, Elena Simperl, Madeline N Cruice, and Timothy Neate. 2025. To Each Their Own: Exploring Highly Personalised Audiovisual Media Accessibility Interventions with People with Aphasia. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference* (Madeira, Portugal) *(DIS '25)*. Association for Computing Machinery, New York, NY, USA, 1826–1843. https://doi.org/10.1145/3715336.3735771

[64] 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

[65] 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

[66] 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

[67] OFCOM. 2020. Internet users' experience of potential online harms: summary of survey research. https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/online-research/online-harms/2020/concerns-and-experiences-online-harms-2020-chart-pack-accessible.pdf?v=324903

[68] 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

[69] Pilar Orero, Marta Brescia-Zapata, and Chris Hughes. 2021. Evaluating Subtitle Readability in Media Immersive Environments. In *9th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-Exclusion* (Online, Portugal) *(DSAI 2020)*. Association for Computing Machinery, New York, NY, USA, 51–54. https://doi.org/10.1145/3439231.3440602

[70] Patrick B O'Sullivan and Caleb T Carr. 2017. Mass personal communication: A model bridging the mass-interpersonal divide. *New Media and Society* 20, 3 (Jan. 2017), 1161–1180. https://doi.org/10.1177/1461444816686104

[71] Carolyn Pang, Zhiqin Collin Wang, Joanna McGrenere, Rock Leung, Jiamin Dai, and Karyn Moffatt. 2021. Technology Adoption and Learning Preferences for Older Adults: Evolving Perceptions, Ongoing Challenges, and Emerging Design Opportunities. 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 490, 13 pages. https://doi.org/10.1145/3411764.3445702

[72] Andrew Perrin and Sara Atske. 2021. Americans with disabilities less likely than those without to own some digital devices. https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/online-research/online-harms/2020/concerns-and-experiences-online-harms-2020-chart-pack-accessible.pdf?v=324903

[73] Helen Petrie, Stefan Carmien, and Andrew Lewis. 2018. *Assistive Technology Abandonment: Research Realities and Potentials*. Springer International Publishing, New York, NY, USA, 532–540. https://doi.org/10.1007/978-3-319-94274-2_77

[74] Betsy Phillips and Hongxin Zhao. 1993. Predictors of Assistive Technology Abandonment. *Assistive Technology* 5, 1 (June 1993), 36–45. https://doi.org/10.1080/10400435.1993.10132205

[75] Mark Priestley, Martha Stickings, Ema Loja, Stefanos Grammenos, Anna Lawson, Lisa Waddington, and Bjarney Fridriksdottir. 2016. The Political Participation of Disabled People in Europe: Rights, Accessibility and Activism. *Electoral Studies* 42 (June 2016), 1–9. https://doi.org/10.1016/j.electstud.2016.01.009

[76] Graham Pullin. 2009. *Design Meets Disability*. MIT Press, London, England.

[77] Philip Quinn and Shumin Zhai. 2016. A Cost-Benefit Study of Text Entry Suggestion Interaction. 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, 83–88. https://doi.org/10.1145/2858036.2858305

[78] 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

[79] Pablo Romero-Fresco and Frederic Chaume. 2022. Creativity in Audiovisual Translation and Media Accessibility. *The Journal of Specialised Translation* 38 (1 2022), 75–101. Issue 1.

[80] Tanya A. Rose, Linda E. Worrall, Louise M. Hickson, and Tammy C. Hoffmann. 2012. Guiding principles for printed education materials: Design preferences of people with aphasia. *International Journal of Speech-Language Pathology* 14, 1 (Dec. 2012), 11–23. https://doi.org/10.3109/17549507.2011.631583

[81] Quentin Roy, Futian Zhang, and Daniel Vogel. 2019. Automation Accuracy Is Good, but High Controllability May Be Better. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–8. https://doi.org/10.1145/3290605.3300750

[82] Korok Sengupta, Sabin Bhattarai, Sayan Sarcar, I. Scott MacKenzie, and Steffen Staab. 2020. Leveraging Error Correction in Voice-based Text Entry by Talk-and-Gaze. 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, 1–11. https://doi.org/10.1145/3313831.3376579

[83] Shannon M. Sheppard and Rajani Sebastian. 2020. Diagnosing and managing post-stroke aphasia. *Expert Review of Neurotherapeutics* 21, 2 (Dec. 2020), 221–234. https://doi.org/10.1080/14737175.2020.1855976

[84] Saul Shiffman, Arthur A. Stone, and Michael R. Hufford. 2008. Ecological Momentary Assessment. *Annual Review of Clinical Psychology* 4, 1 (April 2008), 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

[85] Kristen Shinohara, Cynthia L. Bennett, Wanda Pratt, and Jacob O. Wobbrock. 2018. Tenets for Social Accessibility: Towards Humanizing Disabled People in Design. *ACM Trans. Access. Comput.* 11, 1, Article 6 (March 2018), 31 pages. https://doi.org/10.1145/3178855

[86] Kristen Shinohara and Josh Tenenberg. 2009. A blind person's interactions with technology. *Commun. ACM* 52, 8 (Aug. 2009), 58–66. https://doi.org/10.1145/1536616.1536636

[87] Kristen Shinohara and Jacob O. Wobbrock. 2011. In the shadow of misperception: assistive technology use and social interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 705–714. https://doi.org/10.1145/1978942.1979044

[88] Laurianne Sitbon, Ross Brown, and Lauren Fell. 2019. Turning Heads: Designing Engaging Immersive Video Experiences to Support People with Intellectual Disability when Learning Everyday Living Skills. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, New York, NY, USA, 171–182. https://doi.org/10.1145/3308561.3353787

[89] Katta Spiel, Fares Kayali, Louise Horvath, Michael Penkler, Sabine Harrer, Miguel Sicart, and Jessica Hammer. 2018. Fitter, Happier, More Productive? The Normative Ontology of Fitness Trackers. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3170427.3188401

[90] Aaron Springer and Steve Whittaker. 2020. Progressive Disclosure: When, Why, and How Do Users Want Algorithmic Transparency Information? *ACM Trans. Interact. Intell. Syst.* 10, 4, Article 29 (Oct. 2020), 32 pages. https://doi.org/10.1145/3374218

[91] Abigale Stangl, Shasta Ihorn, Yue-Ting Siu, Aditya Bodi, Mar Castanon, Lothar D Narins, and Ilmi Yoon. 2023. The Potential of a Visual Dialogue Agent In a Tandem Automated Audio Description System for Videos. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility* (New York, NY, USA) *(ASSETS '23)*. Association for Computing Machinery, New York, NY, USA, Article 32, 17 pages. https://doi.org/10.1145/3597638.3608402

[92] 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

[93] Vinoba Vinayagamoorthy, Maxine Glancy, Christoph Ziegler, and Richard Schäffer. 2019. Personalising the TV Experience Using Augmented Reality: An Exploratory Study on Delivering Synchronised Sign Language Interpretation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300762

[94] Agnieszka Walczak. 2017. Audio Description on Smartphones: Making Cinema Accessible for Visually Impaired Audiences. *Universal Access in the Information Society* 17, 4 (Aug. 2017), 833–840. https://doi.org/10.1007/s10209-017-0568-2

[95] Xubin Wang, Zhiqing Tang, Jianxiong Guo, Tianhui Meng, Chenhao Wang, Tian Wang, and Weijia Jia. 2025. Empowering Edge Intelligence: A Comprehensive Survey on On-Device AI Models. *ACM Comput. Surv.* 57, 9, Article 228 (April 2025), 39 pages. https://doi.org/10.1145/3724420

[96] 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

[97] 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

[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] Myriam Winance. 2014. Universal design and the challenge of diversity: reflections on the principles of UD, based on empirical research of people's mobility. *Disability and Rehabilitation* 36, 16 (July 2014), 1334–1343. https://doi.org/10.3109/09638288.2014.936564

[100] 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

[101] Lan Xiao, Maryam Bandukda, Katrin Angerbauer, Weiyue Lin, Tigmanshu Bhatnagar, Michael Sedlmair, and Catherine Holloway. 2024. A Systematic Review of Ability-diverse Collaboration through Ability-based Lens in HCI. 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, Article 961, 21 pages. https://doi.org/10.1145/3613904.3641930

[102] Tianze Xie, Xuesong Zhang, Feiyu Huang, Di Liu, Pengcheng An, and Seungwoo Je. 2025. VRCaptions: Design Captions for DHH Users in Multiplayer Communication in VR. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (Yokahama, Japan) *(CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 757, 18 pages. https://doi.org/10.1145/3706598.3714186

[103] 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