Personalising the TV Experience using Augmented Reality

An Exploratory Study on Delivering Synchronised Sign Language Interpretation

Vinoba Vinayagamoorthy
Maxine Glancy
BBC R&D
London, United Kingdom
vinoba.vinayagamoorthy@bbc.co.uk
maxine.glancy@bbc.co.uk

Christoph Ziegler
Richard Schäffer
IRT
Munich, Germany
ziegler@irt.de

ABSTRACT

Augmented Reality (AR) technology has the potential to extend the screen area beyond the rigid frames of televisions. The additional display area can be used to augment televisions (TVs) with extra information tailored to individuals, for instance, the provision of access services like sign language interpretations. We invited 23 (11 in the UK, 12 in Germany) users of signed content to evaluate three methods of watching a sign language interpreted programme – one traditional in-vision method with signed programme content on TV and two AR-enabled methods in which an AR sign language interpreter (a '*half-body*' version and a '*full-body*' version) is projected just outside the frame of the TV presenting the programme. In the UK, participants were split 3-ways in their preferences while in Germany, half the participants preferred the traditional method followed closely by the 'half-body' version. We discuss our participants reasoning behind their preferences and implications for future research.

CCS CONCEPTS

• **Human-centered computing** → **Mixed / augmented reality**; **Empirical studies in accessibility**.

KEYWORDS

Accessibility, Augmented Reality, BSL, SSE, Companion Screen, Connected Experiences, DGS, HbbTV 2.0, HoloLens, Interaction Techniques, Personalisation, Second Screen, Sign Language, Synchronisation, Television

ACM Reference Format:
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 *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland Uk*. ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3290605.3300762

1 INTRODUCTION

Researchers in the media sector have worked on solutions to enable communication between personal devices and a connected television (TV). The aim is to extend the viewers' "screen" beyond the TV as part of a personalised companion experience in order to aid content discovery, to provide interactive services, or enable the collection of analytics [10, 44]. Augmented Reality (AR) technologies offer the potential to create new display surfaces in the living room for placing content such as an interactive information graphics for election coverage on the coffee table, a three-dimensional model of a dinosaur skeleton on a nearby wall while watching an archaeological documentary, or a presenter on the sofa next to the viewer explaining a Shakespearean drama. As part of an exploratory study, we used an optical head mounted display (HMD), in conjunction with a TV, to explore a way to deliver synchronised sign language interpretations.

The distribution of signed content forms part of the access services offered by broadcasters. In the UK, British Sign Language (BSL) [12, 15] is the main language used in signed content while in Germany, Deutsche Gebärdensprache (DGS) [17–19] is used. Traditionally, signed content is delivered as a picture-in-picture video (*in-vision*) with a signer placed on the bottom right corner of the TV. Although, functionally

ACM ISBN 978-1-4503-5970-2/19/05...$15.00
https://doi.org/10.1145/3290605.3300762

fit for purpose, the in-vision delivery method reduces the display area available for the presentation of the main programme and does not allow for the personalisation of the service to fit user preferences or needs.

We present an exploratory study in which we collect participant responses, in the UK and Germany, to three methods of delivering sign language interpretations of a factual documentary on TV: the traditional picture-in-picture '*in-vision*' method and two AR enabled methods using the designs of two AR interpreters proposed in a previous paper [45].

2 RELATED WORK

The idea of augmenting real world environments to meet the accessibility needs of individuals in a personalised way is of particular interest to media content and service providers.

There are two wide-spread forms of accessibility solutions for individuals with hearing impairments: subtitles (captioning) and sign language interpretation. The National Theatre, in the UK, provides hearing-impaired audiences with 'closed-captioned' glasses presenting the play's dialogue to the wearer [8]. Tomoki et al. [30] proposed a system for captioning 'real-world conversations' next to the speakers face. Peng et al. [37] evaluated the use of 'SpeechBubbles' next to a speaker, through the HoloLens, to dynamically caption group conversations in real-time. Rzayev et al. [41] studied different forms of presentation of text information, unattached to objects in the real-world and Dhruv et al. [24] presented design factors for real-time captioning, both using the HoloLens. A significant source of video content is the TV, so content providers make a valuable effort to make it accessible through augmentations such as subtitles and sign language interpretations [7]. On the TV itself, there has been research into improving the presentation of subtitles [13, 23], the evaluation of subtitles [5] and adding emotional tones to subtitles [3]. There has also been work on dynamic subtitles in 360 [14] and cinematic virtual reality [40].

Montagud et al. [33] describes the challenges faced in providing access services in immersive media, particularly the importance of synchronisation and effective system integration. These become even more interesting when sign language interpretations are considered. Unlike subtitles, sign language interpretations are considerably more complicated (and costly) to generate. Sign languages use gestures, facial expressions and body language to visually communicate information and emotional tones. The British Sign Language (BSL) [12, 15] and Deutsche Gebärdensprache (DGS) [17–19] are languages with their own grammar and vocabulary, independent of their written/spoken counterpart in the UK or Germany. Other forms of communication used, often interchangeably with BSL/DGS, in the hard-of-hearing/deaf communities, include Sign Supported English (SSE) and finger spelling (spelling words using hand movements). Even discourse functions such as turn taking are regulated through indicators such as head movement and eye gaze, or the use of 'resting' positions, such as putting the hands on the lap. It's clear that sign language is nuanced and complex.

There is active work into automatically generating sign language interpretations and presenting them on a 3D humanoid model. SignTime's SiMax [43] provide a service which translates content onto a 3D sign language interpreter and the Microsoft's HoloHear [22] translates the spoken word into American Sign Language (ASL), presented by a 3D model, through a HoloLens [32]. However, studies report different performances of systems with regards to comprehension and acceptance by target users. Ebling et al. [16] identified features (facial expressions, head & shoulder movements, fluidity of movement) that needed improvement in interpretation to the Swiss German Sign Language (DSGS). Kacorri et al. [25] found that the acceptance of computer generated sign language interpreters depended on demographic factors and technology-experience levels. Kipp et al. [27, 28] reported on issues with comprehensibility and acceptability in animated DGS interpreters. Adamo-Villani and Anasingaraju [2] found that a 3D sign language interpreter presented to children through an optical HMD was fun and easy but they do not address the issue of comprehensibility. Given these issues, we did not want to risk diluting the focus of our study through the use of a computer generated sign language interpreter.

In the current market, for the purposes of interpreting video content on a TV, the synchronised video presentation of a precaptured interpreter still offers advantages of familiarity, reliability, comprehensibility and fidelity. Additionally, regulatory guidelines [36] for sign language interpretation of video content favour visual consistency between content and interpreter, easier achieved when captured in video. However, sign language interpretations of content on TV, is often presented as a composite picture-in-picture (in-vision) video of the programme with an interpreter on the bottom right corner. In order to accommodate the interpreter, the content on the TV is often shrunk to about $5/6^{th}$ of the screen. One option to free screen space on the TV, is to offload the sign language interpretation to a personal device (phone/tablet).

Augmenting the TV to personalise the viewers experience has been implemented through a range of methods from utilising secondary personal devices [4, 9, 42, 46, 48] through to designing purpose built systems [1, 10] and offering commercial services [6]. Kožuh et al. [29] used AR-enabled phones/tablets to assess the efficiency of using AR for learning sign language. However, studies show that, other than in experiences with audio description [45], users must constantly shift attention between the TV and their personal device especially when the rate of flow of information and visual complexity of the companion content is high [34]. Delivering functional access services on a companion device pushes the limits of a suitable user experience. Kercher et al. [26] presented a prototype in which an optical HMD is used to present a video of a sign language interpreter in a planetarium type experience to children. The AR interpreter is kept in front of the child at all times and the child is hands-free.

3 RESEARCH OBJECTIVES

We primarily aim to gain an early understanding of how members in the BSL and DGS communities respond to accessing sign language interpretations, through an optical HMD (HoloLens [32]), in AR just outside the TV frame in comparison to the existing in-vision system experienced with traditional TV consumption. We used the HoloLens since it was the first commercially available device on market. We evaluate two designs of AR interpreters which could be used as part of a sign language interpreted experience in a TV+HoloLens system. The main research question was 'how *usable* and *enjoyable* the TV+HoloLens experience presented was?'. We were also interested in any differences in user responses across two countries – the UK & Germany, in addition, to confirming the reproducibility of our studies.

4 DESIGN & IMPLEMENTATION

We studied three methods (conditions) of delivering signed content and each participant involved in the study was shown all three methods (within-subject design). Our control condition (T) was the traditional way of delivering signed content through a TV using in-vision signing (Figure 1). The other two conditions involved viewing a video stream of a sign language interpreter, just outside the TV frame, synchronised to the programme content on the TV, viewed through a HoloLens – an AR sign language interpreter. We used two of our designs for AR sign language interpreters [45] – referred to as '*half-body*' (H) and '*full-body*' (F) AR interpreters (Figure 1).

Since the spoken language and sign language were different in both countries, either programme would not have translated well to the other language. We choose similar programmes, from the content available to us, in both the UK and Germany. Two easy to follow factual documentaries based in the Sahara: the BBC's '*Morocco to Timbuktu: An Arabian Adventure*' and Bayrischer Rundfunk (BR)'s '*Herausforderung Wüste: Sinnsuche in der Sahara*'. The BSL interpreter in the UK was female and the DGS interpreter in Germany was male to match the gender of the programme narrator. We selected a ~6-minute self-contained scene, from each programme, and cut it into three segments. Each of the cut ~2-minute segment was used in combination with the three test conditions. This resulted in 9 experiences of signed content in BSL and 9 in DGS. Participants were then

Figure 1: Conditions in our study (Left column: The UK, Right column: Germany; Top images: traditional control in-vision condition (T), Middle images: 'half-body' AR conditions (H), Bottom images: 'full-body' AR conditions (F)

presented a set of the three conditions, each with a different method of delivering the signed interpretation. We permuted the sequence T, H, F across participants, which resulted in six unique (counter-balanced) sequences. Across the participants, in the UK and Germany, each sequence was used twice, except sequence (F, H, T), which was used in the UK only once due to human error. Participants watched all three ~2-minute segments in the right original narrative sequence.

Designing the Content and the AR Interpreters

For the control condition (T), in the UK, we retrieved and stripped content from archived original broadcast transport streams in both its signed and unsigned forms. We captured videos of sign language interpretations of the chosen content, as presented in [45], for use in the 'half-body' (H) and 'full-body' (F) AR test conditions. We edited the captured videos of interpreters down to a video resolution of 960x1080 pixels, Chroma-keyed and placed on to a black background. The programme content and cut-out video of the interpreter were aligned to a common timeline, using the audio channels of the recordings, in order to be played in synchrony in the 'TV+HoloLens' system. We were not able to source the original German broadcast in signed form so we also made an in-vision style composite for our control condition (T).

We ended up with three videos, each of 2-minute signed content (in-vision in BSL and DGS). We had three video cut-outs of the upper half-bodies of each of our BSL and DGS interpreters and three video cut-outs of each of their full bodies as well. The control in-vision conditions were in keeping with broadcast standards [36]. The 'half-body' and 'full-body' AR interpreters were implemented using our own proposed guidelines [45] (Figure 1):

1. Heads of AR interpreters aligned to the top of the TV,
2. Bottom edge of 'half-body' AR interpreter aligned with the bottom of the TV almost sitting on the TV stand,
3. Feet of the 'full-body' AR interpreter grounded in the physical room,
4. Direction of gaze horizontally facing towards TV while in resting phase and directly towards the viewer during the interpreting phase, and
5. Positioned to slightly overlap the TV.

The 'TV+HoloLens' System

In both countries, we used a 40-inch HD TV to present the programme (signed or unsigned) and a Microsoft HoloLens [32] to present the AR interpreter (in AR conditions).

In the UK, we created a playback application for the HoloLens which displayed a video cut-out of the AR BSL interpreter on a 2D-plane using the Unity Video player; a web browser based video player application to present a video segment on the TV (via a laptop & an HDMI cable); a web browser application to enable precise control over the virtual position of the 2D-plane (through the HoloLens) and trigger synchronous video playback on both the TV (laptop) and HoloLens; and a 'node.js' application to serve the content and handle communication between the applications.

In Germany, we used a similar set-up except the laptop ran a DVB-CSS emulator which simulated a HbbTV 2.0 capable device [21] and the HoloLens ran a 'Universal Windows Platform App' which displayed a video cut-out of the AR DGS interpreter. The HoloLens application, became a 'companion screen', able to communicate with the TV via the network using the DVB-CSS protocols.

Although the means by which the 'TV+HoloLens' system was kept synchronised differed, the end effect of the experience to the participants "looked" very similar.

Pilot Study

We ran twelve 40-minute pilot study sessions (8 in the UK, 4 in Germany) to run through the procedure, test the 'TV+HoloLens' system and get feedback on the content presented. Participants in the pilot did not identify as hard of hearing. In the UK, a participant who was fluent in BSL, and in Germany, a DGS interpreter, were asked to evaluate the quality of signing captured.

During pilots, the attributes of the two AR interpreters – 'half-body' and 'full-body', the set-up of the apparatus and the design of the questionnaires were fine-tuned. AR interpreter attributes included: the placement (position and rotation) of the AR interprets relative to the TV, the scale of the AR interpreter with focus on the vertical position of their head relative to the top of the TV, the direction of gaze of the AR interpreter in both their interpreting and resting phase, and the brightness of the AR interpreter relative to the lighting in the room and brightness of the TV display.

Finally, we designed a training phase based on the two most popular issues pilot participants highlighted: a) managing pitfalls when fitting the HoloLens on participants unfamiliar to the device and b) designing the questions which allowed us to manage participant expectations about a novel device versus the limitations of the 'TV+Hololens' system like the added weight and limited field of the view of the headset.

5 METHOD

We divided each session in the study into four phases. In the first phase we collected consent and demographics through online questionnaires. In the second phase we conducted some training with each participant to familiarise them with the HoloLens. In the third phase, the participant experienced three variants of signed content delivery interspersed with post-condition questionnaires. In the final phase, we collected their comparative evaluations of the three conditions through a semi-structure interview.

The Training

Since AR devices are relatively new to the consumer market, participants were given 'training' (~5 minutes) to familiarise with the HoloLens. In addition, to letting us manage the effect novelty on participants, it gave participants the chance to ask questions. In the UK, we reused content from a previous project [11] and in Germany we utilised content produced by ZDF [47] both showing the weather. During the presentation of the weather on the TV, a half-bodied AR interpreter translated the content on the left hand side of the TV. None of the content used in the training was reused.

Apparatus

Our user lab spaces were set up as two areas. The first area recreated a living room in keeping with TV viewing conditions found in the UK [35]; where participants would view the HD TV sometimes while wearing a HoloLens. The second had a desk area where the participant would use a laptop (connected to a secondary mirroring monitor) to complete online questionnaires. The experimenters monitored the progress the participant made with the questionnaires, on the mirror monitor, and offered assistance as needed.

Questionnaire

We utilised three online questionnaires [39] for this study: a demographics questionnaire, a post condition questionnaire and a post study interview pacer questionnaire.

Our pre-study demographics questionnaire (page 1-4, [39]) included questions to ascertain participant fluency in both spoken and sign language; their level of computer literacy; their usual behaviour with personal devices while watching TV; their familiarity with AR apps (viz. Pokémon Go [38]) and AR devices (viz. the HoloLens [32]); their familiarity with catch-up and/or streaming services and their viewing tastes. We also captured information about how participants used lip-reading skills and access services (signing/subtitles) while watching TV both alone and with others.

The post condition questionnaire was administered to the participant after they saw each test condition while the post study questionnaire was used to conduct a semi-structured interview. The post condition questionnaire [39] consisted of 12 questions rated on a 7-point Likert-type scale. The questions gathered the participant's perception of the interpreter and their quality of experience (page 5, [39]).

The post study interview focused on what participants thought of the idea in general and invited to free-form responses to questions (page 6, [39]). We looked at their preferences between the three test conditions, reasons behind their choices, their self-evaluation of how wearing a HoloLens changed their TV watching behaviour, their suggestions of how each AR test condition might be designed better, their thoughts on what else could be delivered through the HoloLens, their evaluation of the AR interpreters' signing quality and their thoughts on the design of the whole system.

Procedure

At least two experimenters and one sign language interpreter facilitated the study in the UK while in Germany one experimenter facilitated the study. Participants were video recorded through the one-hour study. On arrival, the participant was assigned a predefined participant ID and briefed. In the UK, the interpreter facilitated most of the communication between the experimenters and the participant. In Germany, communication between the participant and the experimenter happened mainly in written form. The participant was asked to sign a consent form and fill in a demographics questionnaire.

Participants were then guided through the training phase, in which they were seated in front of the TV and fitted with the HoloLens. The interpreter helped us coax the participant to express their thoughts about the HoloLens including the limitations such as its weight, its effect on their head posture, the narrow field of view (30^{o} horizontal and 16^{o} vertical, [20]) which needed to encompass both the TV and the AR interpreter just outside the TV frame, and the time taken to acclimatise to the experience.

After training, the participant was asked to continue with the test conditions. As mentioned, although the method of delivering sign language interpretation was randomised, participants saw all three 2 minute segments of the chosen scene in the original narrative order. The participant sat in front of the TV with the interpreter, while assisting experimenters sat behind the participant. Once the first condition ended, the participant was asked to fill in an online post-condition questionnaire. This was repeated twice more to cover all three conditions.

After completing all three conditions, the participant was asked to go through an online post study questionnaire with the help of an experimenter and the interviewer. This questionnaire acted as a pacer for the semi-structure interview covering their overall thoughts about all three test conditions. After the interview, the participant was debriefed.

Participants

In the UK, we used an agency to recruit 12 participants who regularly needed/used BSL interpreted content for accessibility needs. In Germany, we recruited 13 participants through a mix of 'word of mouth' and directly contacting various associations [17–19]. However, we were only able to collect a complete data-set for 11 participants in the UK and 12 participants in Germany due to human error.

In the UK, we had 7 females and 4 males between the ages of 28 and 71 (M=54, SD=13). In Germany, we had 5 females and 7 males between the ages of 23 and 66 (M=46, SD=14). In the UK, 4 participants were profoundly deaf, 4 were either deaf or had hearing impairments, while 3 indicated no uncorrected hearing impairments. In Germany, 7 were deaf, 2 severely hard of hearing, 2 hard of hearing and one gave no indication of a lack of hearing impairment.

In the UK, 6 participants were fluent in English, while 3 were proficient and 2 had a basic command of English; 9 participants were fluent in BSL while 1 was proficient and 1 had a basic command of BSL; 6 participants were fluent in SSE, 2 were proficient in SSE, 1 had a basic command of SSE and 2 reported not knowing SSE. Of the 12 participants, 3 used BSL as a first language, 1 considered English as their first language but BSL was their preferred language and they used BSL 90% of the time for communicating, 6 used BSL and English to the same extent and 1 had some knowledge of BSL but hardly used it for communication. In Germany, 6 participants self-assessed as having a good command of German, the other 6 reported their command of German as very good; 3 had a good knowledge of DGS, the remaining 9 had a very good knowledge of DGS. Of the 12 participants, 3 reported using DGS more often than German while the remaining 9 used both languages to the same extent.

All our participants watched at least an hour of TV per day and none of them had a bias against factual documentaries. In the UK, while watching TV alone, 9 participants reported using subtitles – out of which 4 used signed content and 3 used signed content with subtitles. The remaining 2 participants preferred using signed content with subtitles always. However, 3 out of the 11 participants reported only using subtitles if they watched TV with others. In Germany, 7 participant preferred subtitles and 5 preferred both signed content with subtitles when watching TV alone. While watching with others, 8 participants used subtitles, 3 used signed content with subtitles and 1 participant used neither. There was a general preference for subtitles because it was seen as accurate and as a way to improve written language skills. Signed content was seen as more comprehensive and helpful if participants had a basic command of the written language.

In the UK, 9 participants could lip read while 2 reported that success in lip reading depended on the other person. In Germany, 7 participants could lip read, 4 reported being able to lip read under good conditions while 1 did not know how to lip read. In the UK, 3 sometimes lip read while watching TV, 1 reported lip reading on TV under good conditions, 1 reported lip reading if the person on TV faced the camera, 5 never lip read while watching TV and 1 reported as not consciously lip reading. In Germany, 5 participants lip read while watching TV, 1 reported lip reading on occasion and 6 never lip read while watching TV. Unsurprisingly, participants mentioned it was easier to lip read people on TV (or real life) if they were facing them head on.

Most of our participants (9 in the UK, 8 in Germany) self-assessed as 'expert' with regards to their familiarity with computers. Few of our participants reported expert experience with video games (1 in the UK, 5 in Germany) and even less reported expert experience with AR apps (1 in the UK, 1 in Germany). In the UK, 10 participants reported heavy use of touch screen devices while only 2 did so in Germany. In the UK, more participants reported getting easily distracted by content on their mobile device (7 compared to 3 in Germany) and being easily absorbed by content on their TV (7 compared to 4 in Germany).

In the UK, participants were paid £70.00 and in Germany, they were paid €10.00 in Amazon vouchers.

6 QUANTITATIVE RESULTS

We ran statistical tests across participant responses to the post-condition and the post-study questionnaires.

Post-Condition Questionnaire

Two types of statistical tests were applied to the post-condition questionnaire responses. The first set of tests examined the influence of the three methods studied – traditional (T), 'half-body' AR (H) and 'full-body' AR (F) interpreter on the responses of the UK and of the German participants. The second set of tests examined whether the participants in either country responded differently. We used non-parametric tests, since Shapiro-Wilk tests revealed that, for most of our samples, we cannot assume normally distributed populations. Figure 2 show descriptive statistics on participant responses.

Effect of the Method on Participant Responses. We examined the effect of the methods on participants' responses to the post-condition questionnaire by means of the Friedman test. Solid bordered boxes in figure 2 show where tests found significant differences between participant responses.

A significant effect of the method on responses to q2 "*I felt I was missing part of the programme on the TV because I had to pay attention to the signer.*" was found for both participant groups (in the UK: χ^2=8.45, df=2, p=0.0146; in Germany: χ^2=6.0556, df=2, p=0.048). A post-hoc analysis should tell us between which methods the differences in participant's responses occurred. For the pairwise comparisons of samples, we used the Wilcoxon signed-rank test and corrected p-values with the Bonferroni method. In the responses of the UK participants, the pairwise comparison found a significant difference between the traditional interpreter and the 'half-body' method (T vs. H: p=0.0467, T vs. F: p= 1, H vs. F: p=0.3622). Based on Cohen's classification of the effect size $r = z/\sqrt{N} = 0.43$, we conclude a medium effect of the method on the difference in ratings. In the responses of German participants, the test found no sample pair with a significant difference of the central tendencies (T vs. H: p=0.1588, T vs. F: p=0.2994, H vs. F: p=1).

In the group of German participants, Friedman tests uncovered a significant difference in the responses to q6 "*I felt the signer obscured too much of the programme content.*" for the different methods (χ^2 =7.3125, df=2, p=0.0258). The pairwise comparison found a significant difference (medium effect, r=0.41) between participant responses to the traditional interpreter and the 'full-body' AR interpreter (T vs. H: p=0.1259, T vs. F: p=0.0447, H vs. F: p=1) and no differences between the other sample pairs.

In the group of UK participants, an effect of the methods on responses to q7 "*I felt there was a delay between the narrator in the programme and the signer*" was found (χ^2= 6.3, df=2, p=0.0429). However, the post-hoc test did not reveal between which sample pairs the difference was significant (T vs. H: p= 1, T vs. F: p=0.1793, H vs. F: p=0.2669).

German vs. UK Participant Responses. We ran Mann-Whitney U tests to investigate if responses to the post-condition questionnaire differed between participants in the UK and Germany. Dashed boxes in figure 2 show for which combination of method and question, participants gave significantly different responses. Differences where found for 2 of the 11

		q1			q2			q3			q4		
		T	H	F	T	H	F	T	H	F	T	H	F
UK participants	max	7	7	7	6	7	7	7	7	7	7	7	7
	3rd	7	6	6	4.5	7	5.5	6.5	7	7	6	6	6
	2nd	6	6	5	3	6	5	6	7	6	5	5	5
	1st	5	5	4.5	2	4.5	3	6	6	4.5	4	4.5	3
	min	1	3	4	1	3	1	2	3	2	3	1	3
German participants	max	7	7	7	7	7	7	7	7	7	7	7	7
	3rd	7	7	7	5	7	7	7	7	7	7	6.3	6.3
	2nd	6	6.5	6.5	4.5	6	6	6.5	7	7	6	6	6
	1st	5	5	4.8	3.8	3.8	4	6	6	6	5	5.5	5
	min	1	1	4	1	1	1	5	5	5	4	3	1

		q5			q6			q7			q8		
		T	H	F	T	H	F	T	H	F	T	H	F
UK participants	max	7	4	5	5	7	5	6	7	4	7	7	7
	3rd	4	4	4	4.5	5	4	4	4	3.5	5.5	6.5	6.5
	2nd	4	4	4	3	3	2	4	4	2	5	5	5
	1st	3	3.5	3	2	1.5	2	1.5	1	1	3.5	5	4
	min	3	2	2	1	1	1	1	1	1	1	4	4
German participants	max	6	6	6	7	6	5	-	-	-	7	7	7
	3rd	4.3	4	5	5.3	2.8	2	-	-	-	7	7	7
	2nd	4	4	4	5	1	1	-	-	-	6	6	7
	1st	4	4	3.5	1.8	1	1	-	-	-	4	5	4
	min	3	3	1	1	1	1	-	-	-	1	4	3

		q9			q10			q11			q12		
		T	H	F	T	H	F	T	H	F	T	H	F
UK participants	max	4	6	6	7	6	7	7	7	7	7	6	7
	3rd	2.5	4.5	5	6	4.5	6.5	4.5	6.5	5.5	6.5	5.5	6.5
	2nd	2	4	3	6	4	6	2	5	3	6	5	5
	1st	1	3	2	5	2	3.5	2	2.5	1.5	4.5	4.5	4.5
	min	1	2	1	2	2	1	1	2	1	3	4	4
German participants	max	4	6	6	7	7	7	7	7	7	7	7	7
	3rd	2.3	3.3	4.3	7	6.3	6.3	6	6.3	7	7	6.3	7
	2nd	1	1.5	1	6	5	5.5	5	6	5	7	5.5	5.5
	1st	1	1	1	4	2.8	4.3	3.8	5	4.8	4	4	3.5
	min	1	1	1	3	1	1	1	4	2	3	2	1

Figure 2: Descriptive statistics on responses to post condition questions [39]: max: maximum, 3rd: third quartile, 2nd: median, 1st: first quartile, min: minimum. Solid boxes indicate where Friedman test indicates an influence of method (T: traditional, H: half-body, F: full-body). Dashed boxes indicate where Mann-Whitney U test indicate a difference across countries.

questions answered by all participants in the UK and Germany. The first difference (medium effect, r=0.44) was found for the responses to q6 "*I felt the signer obscured too much of the programme content.*" on the full-body method (W=98.5, p=0.037). The second difference (medium effect, r=0.48) was found between responses to q9 "*I felt frustrated with the whole experience*" on the 'half-body' method (W=102, p=0.025).

Post-study questionnaire

We observed an equal split of responses by participants in the UK when asked which interpretation method they liked best (T: 4, H: 4, F: 3). In Germany half of the participants preferred the AR interpreter, of whom only 1 preferred the 'full-body' interpreter (T: 6, H: 5, F: 1). Results of Kolmogorov-Smirnov tests indicate that the distributions of the responses do not differ significantly between countries (D=0.6667, p=0.5176) – see Figure 3a.

Figure 3b illustrates participant responses to questions on the acceptance of the TV+HoloLens system. The responses participants in UK and Germany gave on the question "*I liked the design of the whole system (TV+HoloLens)...*" did not differ significantly in their central tendencies (median UK: 4, median Germany: 5.5, Mann-Whitney U test: W=36.5, p=0.0677). Same central tendencies were also found for their responses to the question "*I would like to have a TV and HoloLens system in my home!*" (median UK: 4.5, median Germany: 6, Mann-Whitney U test: W=56.5, p=0.5687).

Participants were asked if the HoloLens changed the way they watched television. Figure 3c) illustrates the distribution of responses of our participants. In UK and Germany, 8 participants each responded in a good way. One of these in the UK stated he expected that. In Germany, 4 participants expected the positive change. Three participants each, in the UK and Germany, stated the way they watched TV changed in a bad way. In the UK, one participant said he expected that. In Germany, one participant stated they expected that. One participant in Germany claimed not to have experienced any change. Results of a Kolmogorov-Smirnov test suggest that the samples come from populations with the same distribution (D=0.2, p=1).

7 QUALITATIVE RESULTS

We conducted a thematic analysis on participants responses to the questions of the post-study questionnaire and on participants statements from the post-study interview.

HoloLens – A Novel Technology

The HoloLens was a new technology for all the participants except for one in the UK. The delivery of sign language interpretations through the HoloLens was novel and impressive to all participants. Participants liked the idea of using AR for sign-language interpretation services. In comparison to the AR interpreter, the traditional in-vision way was seen as 'flat'. However, participants felt there was room for improvement before it was seen as an accessible tool with more research needed to understand: the optimal placement of the AR interpreter, the barriers to new technology and how it integrated with with the technologies users already use.

> "*I have always wanted to watch a hologram and there has to be something in which a hologram just happens. Something like a projection. Rather than something you have to put on.*" – P4 (UK)

Physical Discomfort

All participants discussed the physical effects of wearing the HoloLens. The *weight* and *bulk* of the HoloLens, the difficulty of fitting the HoloLens over glasses or cochlea implants, and the need to maintain a fairly stable head posture to avoid losing the optimum field of view were all mentioned. Participants (3 in the UK) were more tolerant of wearing optical HMDs at the theatre/cinema over day-to-day viewing at home but one UK and one German participant reported that they could not imagine watching a full movie with the HoloLens on. Another participant (Germany) said that the

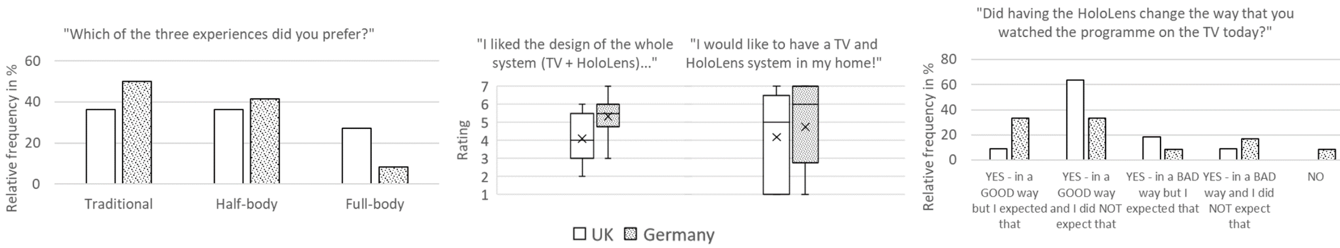

Figure 3: a) Participant preference of the methods; b) Responses on questions about the acceptance of the TV+HoloLens system; c) Responses to the question on whether the HoloLens changed the television experience.

HoloLens would cause him neck problems in the long run. Some participants (Germany) mentioned that they would like additional padding for the frame of the HoloLens.

> "*Having glasses on your head for two hours? No way! And I can imagine, you put them on and everyone will look at you and they will recognise you are deaf. I don't like showing the outside world that I am deaf. I feel too vulnerable*" – P10 (UK)

Some participants (UK) felt the HoloLens broke their connection with the AR interpreter and the TV content. They reflected that lighter headsets or adaptation to glasses already being used might be a better option.

> "*For a longer film the holograms are unsuitable, since eyes get tired fast. In this case subtitles are better. Maybe in the future holograms could be displayed without HoloLens or other glasses?*" – P1 (Germany)

However, some participants (Germany) stated they could imagine getting used to wearing the HoloLens while watching TV while some participants (UK) mentioned the limitations of the HoloLens and its effect on their behaviour forced them to concentrate on the content.

HoloLens – Limitations of the Technology

Due to the limited field of view of the HoloLens, the presentation of the AR interpreter was sometimes compromised. A change in the participants' head posture resulted in the AR interpreter dropping outside the field of view. Participants wondered how this might affect longer term viewing. A few participants (1 in the UK, 2 in Germany) also noticed colour separation effects when moving their head and while a few reported that some images were out of focus.

Position of the Interpreter

Participants understandably found the position of the AR interpreter, outside the frame of the TV, novel. In the traditional in-vision condition, the interpreter overlapped the TV content and could obscure up to 1/7th of the content. In the 'half-body' and 'full-body' conditions, the AR interpreter overlapped the content very slightly but the AR interpreter was transparent. Participants acknowledged that it is positive, that with AR interpreter conditions, the programme can occupy the entire TV screen real estate and does not have to be scaled down in favour of the sign language interpreter.

However, placing the interpreter outside the TV frame raised comments from participants, that it required more pronounced eye shifts and slight head movements thereby increasing the physical workload. This caused participants to feel a loss of connection between the AR interpreter and content since the focus of the participant shifted more towards to the AR interpreter due to not being able to use their peripheral vision as much.

> "*With traditional method, you can see everything within your eye line. You have to move attention from one to the other in the HoloLens conditions like a game of tennis. I wanted the signer on the picture even if it obscured the picture. It would be more acceptable*" – P2 (UK)

Most of the participants indicated that they would prefer to be able to determine the placement of the AR interpreter themselves.

Clarity of the Interpreter

All participants acknowledged that the sign language interpretations were of good quality. In the UK, some participants said this was because the AR interpreter and the clarity of the images through the HoloLens was good. However, some of the participants mentioned that a complex background or lack of contrast between AR interpreter and background can affect intelligibility.

> "[At home] *We do not watch television in a darkened room. The* [laboratory] *room has an optimal background. In private rooms it is often different. All in all, I would like the holograms to be sharper, richer in contrast and the distance between HoloLens interpreter and the television to be smaller.*" – P7 (Germany)

Participant Preferences

In the interviews, the participants gave expanded feedback and justifications for their individual preference.

Traditional (in-vision) vs. AR (HoloLens). Participants used descriptive terms such as 'flat' and '3D' while comparing the traditional interpreter and the AR interpreters.

> "*She was sharper and she felt closer to you. It felt like I had my own personal interpreter in the room with me. I wanted to offer her a cup of tea.*" – P7 (UK)

> "*I found* [the full-body interpreter] *exciting, because you think that the interpreter was there live.*" – P5 (Germany)

Participants who preferred the in-vision method of signed content delivery indicated that it was easier for them to follow the interpreter. One participant added that he found it annoying to have to look back and forth between the AR interpreter and the TV picture. Participants who preferred the AR interpreter said they liked the fact that the important part of the TV picture was not obscured by the AR interpreter. Participant talk about the disconnect between AR interpreter and the TV content in the AR conditions. Finger spelling gets difficult to 'read' when the distance between the interpreter and content gets wide.

> "*The HoloLens interpreter was clearer, quality of the sign language, than the one on the TV. I didn't like the narrow field of view since I couldn't watch both at the same time and it was sensitive to my head movement. I had to find the right angle.*" – P1 (UK)

Half-Body vs. Full-Body. Participants, in the UK, were split in their preference between the 'half-body' AR interpreter and the 'full-body' AR interpreter. In Germany, participants preferred the 'half-body' interpreter. Participants who preferred the 'half-body' AR interpreter stated that it contained all relevant information. Some participants (Germany) explicitly stated that the full-body interpreter contained "irrelevant information".

> "*It didn't matter whether it was half or full. I concentrated from the hips upwards.*" – P6 (UK)

> "*Half the body is enough; I don't necessarily have to see the legs. This would allow me to see my hands and face better. The whole body looks too big and the face would be too small.*" – P8 (Germany)

Participants who preferred the full-body found it "*more pleasant and beautiful*", better "*overall coverage*" or "more understandable and pleasant to read".

> "*I prefer the full body. I didn't really like the half. I felt like I was missing when Dominique* [the half-body interpreter] *used signs in the lower part of the signing space in the half body. With the full body, you could see more of the body language.*" – P11 (UK)

Different interpreters have different styles of signing especially with respect to how much 'formality' and 'constraint' they put into their interpretation. The preference for 'half-body' or 'full-body' interpreter might be dependent on the interpreter captured.

AR Interpreters and Subtitles in AR. Most participants indicated that subtitles were generally preferred to signed content. Three participants (2 in the UK, 1 in Germany) stated that they could imagine getting subtitles delivered on the HoloLens and that the subtitles should be fixed in the field of vision. Although most of our participants thought that the way subtitles are delivered currently is quite well done already. Another participant (Germany) said that it strongly depends on the skills of the interpreter whether he prefers sign language to subtitles.

Participants, in both countries, said that translations by non-hearing (deaf) interpreters – who use sign language as their mother tongue – are often noticeably different or better than those by hearing interpreters. Some interpreters are seen as more emotive while others are seen as more informative and the viewers' preferences often depend on the type of content.

Implications for the Future

Some of the participants made suggestions for extensions of the technology. One participant (Germany) indicated that he was interested in a 3D sign language interpreter. Another said that he was wondering if it made sense to display several AR interpreters to interpret conversations as options. This urge to control and fine tune the sign language interpreted experience was echoed by participants in the UK as well.

> "*Having control over where to place the signer would be good. People have different preferences and it would be good to adjust to type of programme, peoples' visual capability and wallpaper. I don't have time to make adjustments to such a system and fine tune things… There should be something universal.*" – P9 (UK)

However, there were concerns that the experience would not be easy to personalise or that it would take too much effort. Features participants wanted to control included: the choice between a 'half-body' or 'full-body' AR interpreter, the choice between a 'hearing' interpreter or a 'non-hearing' interpreter, the option to have subtitles, the functionality to choose the scale and placement of the AR interpreter.

8 DISCUSSION

We aimed to evaluate the potential of AR technologies to augment TV viewing with the delivery of synchronous sign language interpretations of the content. Our data reveals no outright preferences between the three methods of sign language delivery. However, we gained a deeper insight into our participants' perceptions and pointers for improving the design of AR interpreters [45].

Unsurprisingly, participants pointed out that the HoloLens was not quite suitable yet for this type of usage due to the technical limitations of the device. We tried to manage limitations such as the weight and narrow field of view in the study design and participant training, but there were also colour separation issues as reported by Microsoft [31]. Despite the limitations, participants liked the idea of using novel technologies to solve a real world problem in TV watching and were open to the idea of using AR technologies to receive synchronous sign language interpretations in the future (Figure 3b and Figure 3c). Unfortunately, we cannot predict if it is worth putting extra effort into the production of 'full-body' interpreter videos.

At the beginning of the study, we thought it had the potential to enrich the experience of viewing signed content. However, that is not how all participants felt. Some thought AR interpreters, particularly 'full-body' versions, were distracting while others thought it was natural and afforded a sense of completeness. The principles used in the design and management of interpreters for in-vision signed broadcast content, as regulated by Ofcom [36] in the UK, seem to apply to the building of AR interpreters. The important factors include clarity of interpreter, maintaining the connection between interpreter and content, placement of interpreter relative to the content, the balance between the interpreter overlapping over the content without obscuring the content.

Personalisation is key to delivering useful TV augments. Our results show that the demands made by deaf viewers on the sign-language service are very individual. Our observations indicate that a TV augment which proffers an AR sign language interpreting option should exploit the potential of the technology and offer sufficient degrees of freedom in the presentation of the AR interpreter to the viewers. We gathered from the participants that it would be desirable to be able to determine the position of the AR interpreter and that a changeable background or dress colour could improve the intelligibility of sign language in some viewing environments. There is an obvious advantage in being able to tailor the way in which individuals from various backgrounds across different age ranges and needs choose to get their sign language interpretation. This has the potential to be especially important for our younger audiences. Signed content play an important role in the education of children with hearing impairments especially where sign language is their only form of communication and a lot of incidental learning comes from signed video content. Using AR devices in this way could have a significant impact on how we might enable easier access to sign language interpretations for content.

9 CONCLUSIONS & FUTURE WORK

We conducted an exploratory study, in both the UK and Germany, to gauge participant responses to the concept of delivering synchronised sign language interpretations through an optical HMD while watching a factual documentary on a connected TV. There were two main differences, not by design, in the implementation of the studies in both countries. Firstly, there were at least two experimenters and a BSL interpreter to help conduct studies in the UK while in Germany we had one experimenter who used the written form of German for communicating with participants. This might account for the more verbose responses from our UK participants. We would recommend the use of a sign language interpreter if working with participants in the deaf community.

Secondly, in Germany, we used the same male sign language interpreter for all three conditions tested, however, in the UK, we used a female 'hearing' interpreter to build the AR interpreters in the two AR conditions, but we simply used the broadcasted signed content, with a female 'non-hearing' interpreter, for the control condition. This meant that there were two different interpreters used across the conditions tested in the UK with two distinct styles of interpreting. This led us to unpick participant preferences which depended on the style of interpretation of our two BSL interpreters. In future studies, it might be interesting to see how different styles of sign language interpretation affect signed experiences.

Would the experience be richer if the AR interpreter was truly 3D? Participants commented that the AR interpreters in the study looked 3D and in comparison, the in-vision interpreter looked flat. The AR interpreters used in our study were video cut outs projected on a 2D place (billboards). Would participants respond differently to an experience accompanied by a volumetric video capture of a sign language interpreter? Here, we focused on placing augments in front of the viewer, as proposed by Vinayagamoorthy et al. [45]. How far can we push the capabilities of optical HMDs in conjunction with a connected TV? When combined with frame-accurate synchronisation technologies like DVB-CSS [46], there is potential to place augments around the viewer in their living room space. In future studies, it would be interesting to explore how different types of users, including children, react to novel accessible personalised experiences designed specifically for a TV+HoloLens system.

ACKNOWLEDGMENTS

This research was funded through the European Union's H2020-ICT-2015 programme under grant agreement number 687655 (2-IMMERSE). We thank our colleagues from BBC R&D and IRT who advised on this project.

REFERENCES

[1] 2-Immerse. 2018. 2-Immerse home page. Retrieved March 27, 2018 from http://2immerse.eu

[2] Nicoletta Adamo-Villani and Saikiran Anasingaraju. 2016. Holographic Signing Avatars for Deaf Education. In *E-Learning, E-Education, and Online Training. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, Giovanni Vincenti, Alberto Bucciero, Markus Helfert, and Matthias Glowatz (Eds.). Springer International Publishing, Cham, 54–61.

[3] Diana Affi, Joël Dumoulin, Marco Bertini, Elena Mugellini, Omar Abou Khaled, and Alberto Del Bimbo. 2015. SensiTV: Smart EmotioNal System for Impaired People's TV. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '15)*. ACM, New York, NY, USA, 125–130. https://doi.org/10.1145/2745197.2755512

[4] Edward Anstead, Steve Benford, and Robert J. Houghton. 2014. Many-screen Viewing: Evaluating an Olympics Companion Application. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '14)*. ACM, New York, NY, USA, 103–110. https://doi.org/10.1145/2602299.2602304

[5] Mike Armstrong. 2013. The Development of a Methodology to Evaluate the Perceived Quality of Live TV Subtitles. In *IBC2013 Conference*. IET, Amsterdam. https://doi.org/10.1049/ibc.2013.0044

[6] Augmen.tv. 2018. Augmen.TV homepage. Retrieved September 21, 2018 from http://augmen.tv

[7] BBC. 2016. Access Services at the BBC: Subtitling, Audio Description & Signing - Inside the BBC. Retrieved September 11, 2018 from https://www.bbc.co.uk/corporate2/insidethebbc/whatwedo/access-services

[8] BBC News. 2017. National Theatre Specs create floating subtitles. Retrieved September 21, 2018 from https://www.bbc.co.uk/news/av/technology-41491953/national-theatre-specs-create-floating-subtitles

[9] BBC R&D. 2012. Companion Screens: Creating a viewing experience across more than one screen. Retrieved March 27, 2018 from http://www.bbc.co.uk/rd/projects/companion-screens

[10] BBC R&D. 2013. Unconventional Screens: Exploring the potential of future display technologies. Retrieved March 27, 2018 from https://www.bbc.co.uk/rd/projects/unconventional-screens

[11] BBC R&D. 2015. Forecaster: our experimental object-based weather forecast. Retrieved September 4, 2018 from https://www.bbc.co.uk/rd/blog/2015-11-forecaster-our-experimental-object-based-weather-forecast

[12] British Sign Language. 2018. Online resources, games & course. Retrieved September 19, 2018 from https://www.british-sign.co.uk

[13] Andy Brown, Rhia Jones, Mike Crabb, James Sandford, Matthew Brooks, Mike Armstrong, and Caroline Jay. 2015. Dynamic Subtitles: The User Experience. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '15)*. ACM, New York, NY, USA, 103–112. https://doi.org/10.1145/2745197.2745204

[14] Andy Brown, Jayson Turner, Jake Patterson, Anastasia Schmitz, Mike Armstrong, and Maxine Glancy. 2017. Subtitles in 360-degree Video. In *Adjunct Publication of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '17 Adjunct)*. ACM, New York, NY, USA, 3–8. https://doi.org/10.1145/3084289.3089915

[15] BSL Zone. 2018. Watch. Retrieved September 19, 2018 from https://www.bslzone.co.uk/watch/

[16] Sarah Ebling, Sarah Johnson, Rosalee Wolfe, Robyn Moncrief, John McDonald, Souad Baowidan, Tobias Haug, Sandra Sidler-Miserez, and Katja Tissi. 2017. Evaluation of Animated Swiss German Sign Language Fingerspelling Sequences and Signs. In *Universal Access in Human–Computer Interaction. Designing Novel Interactions*, Margherita Antona and Constantine Stephanidis (Eds.). Springer International Publishing, Cham, 3–13.

[17] Fachverband für Menschen mit Hör- und Sprachbehinderung e.V. 2018. Home page. Retrieved September 5, 2018 from http://www.blwg.eu/

[18] Gehörlose Bergfreunde München e.V. 2018. Home page. Retrieved September 5, 2018 from http://www.gbf-muenchen.de/

[19] Gehörlosenverband München und Umland e.V. 2018. Home page. Retrieved September 5, 2018 from https://www.gmu.de/

[20] Uwe Gruenefeld, Dana Hsiao, Wilko Heuten, and Susanne Boll. 2017. EyeSee: Beyond Reality with Microsoft HoloLens. In *Proceedings of the 5th Symposium on Spatial User Interaction (SUI '17)*. ACM, New York, NY, USA, 148–148. https://doi.org/10.1145/3131277.3134362

[21] HbbTV. 2018. *TS 102796 V1.5.1 – Hybrid Broadcast Broadband Television.* ETSI, Sophia Antipolis Cedex, France.

[22] HealthiAR. 2018. Microsoft HoloHear. Retrieved September 21, 2018 from https://healthiar.com/holohear-translates-the-spoken-word-to-asl-in-mixed-reality

[23] Chris J. Hughes, Mike Armstrong, Rhianne Jones, and Michael Crabb. 2015. Responsive Design for Personalised Subtitles. In *Proceedings of the 12th Web for All Conference (W4A '15)*. ACM, New York, NY, USA, Article 8, 4 pages. https://doi.org/10.1145/2745555.2746650

[24] Dhruv Jain, Bonnie Chinh, Leah Findlater, Raja Kushalnagar, and Jon Froehlich. 2018. Exploring Augmented Reality Approaches to Real-Time Captioning: A Preliminary Autoethnographic Study. In *Proceedings of the 2018 ACM Conference Companion Publication on Designing Interactive Systems (DIS '18 Companion)*. ACM, New York, NY, USA, 7–11. https://doi.org/10.1145/3197391.3205404

[25] Hernisa Kacorri, Matt Huenerfauth, Sarah Ebling, Kasmira Patel, Kellie Menzies, and Mackenzie Willard. 2017. Regression Analysis of Demographic and Technology-Experience Factors Influencing Acceptance of Sign Language Animation. *ACM Trans. Access. Comput.* 10, 1, Article 3 (April 2017), 33 pages. https://doi.org/10.1145/3046787

[26] Kellie Kercher and Dale C Rowe. 2012. Improving the learning experience for the deaf through augment reality innovations. In *Engineering, Technology and Innovation (ICE), 2012 18th International ICE Conference on*. IEEE, Munich, 1–11.

[27] Michael Kipp, Alexis Heloir, and Quan Nguyen. 2011. Sign language avatars: Animation and comprehensibility. In *International Workshop on Intelligent Virtual Agents*. Springer, Berlin, Heidelberg, 113–126.

[28] Michael Kipp, Quan Nguyen, Alexis Heloir, and Silke Matthes. 2011. Assessing the Deaf User Perspective on Sign Language Avatars. In *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '11)*. ACM, New York, NY, USA, 107–114. https://doi.org/10.1145/2049536.2049557

[29] Ines Kožuh, Simon Hauptman, Primož Kosec, and Matjaž Debevc. 2015. Assessing the efficiency of using augmented reality for learning sign language. In *International Conference on Universal Access in Human-Computer Interaction*. Springer, Cham, 404–415.

[30] Tomoki Kurahashi, Kazuki Suemitsu, Keiichi Zempo, Koichi Mizutani, and Naoto Wakatsuki. 2017. Disposition of captioning interface using see-through head-mounted display for conversation support. In *Consumer Electronics (GCCE), 2017 IEEE 6th Global Conference on*. IEEE, Nagoya, Japan, 1–4.

[31] Microsoft. 2018. Hologram stability. Retrieved September 20, 2018 from https://docs.microsoft.com/en-us/windows/mixed-reality/hologram-stability#color-separation

[32] Microsoft. 2018. Microsoft HoloLens. Retrieved March 28, 2018 from https://www.microsoft.com/en-gb/hololens

[33] Mario Montagud, Issac Fraile, Juan A. Nuñez, and Sergi Fernández. 2018. ImAc: Enabling Immersive, Accessible and Personalized Media Experiences. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '18)*. ACM, New York, NY, USA, 245–250. https://doi.org/10.1145/3210825.3213570

[34] Timothy Neate, Matt Jones, and Michael Evans. 2015. Mediating Attention for Second Screen Companion Content. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 3103–3106. https://doi.org/10.1145/2702123.2702278

[35] Katy C. Noland and Louise H. Truong. 2015. *A Survey of UK Television Viewing Conditions.* Technical Report WHP 287. BBC R&D. Retrieved December 19, 2018 from http://downloads.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP287.pdf

[36] Ofcom. 2017. The Ofcom Code on Television Access Services. Retrieved March 27, 2018 from https://www.ofcom.org.uk/tv-radio-and-on-demand/broadcast-codes/tv-access-services

[37] Yi-Hao Peng, Ming-Wei Hsi, Paul Taele, Ting-Yu Lin, Po-En Lai, Leon Hsu, Tzu-chuan Chen, Te-Yen Wu, Yu-An Chen, Hsien-Hui Tang, and Mike Y. Chen. 2018. SpeechBubbles: Enhancing Captioning Experiences for Deaf and Hard-of-Hearing People in Group Conversations. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 293, 10 pages. https://doi.org/10.1145/3173574.3173867

[38] Pokémon Go. 2018. Homepage. Retrieved September 3, 2018 from https://www.pokemongo.com/en-gb/

[39] BBC R&D. 2018. Companion Device Research - Information and Questionnaire Sheet. Retrieved December 19, 2018 from https://docs.google.com/forms/d/e/1FAIpQLSfX3WsFiY1Y5NMLY46UiydrHh-WsfWG3-cDJb2HSKwTJ3D8TA/viewform

[40] Sylvia Rothe, Kim Tran, and Heinrich Hußmann. 2018. Dynamic Subtitles in Cinematic Virtual Reality. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '18)*. ACM, New York, NY, USA, 209–214. https://doi.org/10.1145/3210825.3213556

[41] Rufat Rzayev, Paweł W. Woźniak, Tilman Dingler, and Niels Henze. 2018. Reading on Smart Glasses: The Effect of Text Position, Presentation Type and Walking. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 45, 9 pages. https://doi.org/10.1145/3173574.3173619

[42] Ege Sezen. 2015. Enhancing Watching Experience of Football Matches on TV via Modes of Interaction and Types of Visualisation of Match-Related Information on Second Screen. In *ACM TVX'2015 Doctoral Consortium.* ACM, Brussels.

[43] Simax. 2018. Home page. Retrieved September 21, 2018 from https://simax.media/?lang=en

[44] Vinoba Vinayagamoorthy, Penelope Allen, Matt Hammond, and Michael Evans. 2012. Researching the User Experience for Connected Tv: A Case Study. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems (CHI EA '12)*. ACM, New York, NY, USA, 589–604. https://doi.org/10.1145/2212776.2212832

[45] Vinoba Vinayagamoorthy, Maxine Glancy, Paul Debenham, Alastair Bruce, Christoph Ziegler, and Richard Schäffer. 2018. Personalising the TV Experience with Augmented Reality Technology: Synchronised Sign Language Interpretation. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '18)*. ACM, New York, NY, USA, 179–184. https://doi.org/10.1145/3210825.3213562

[46] Vinoba Vinayagamoorthy, Rajiv Ramdhany, and Matt Hammond. 2016. Enabling Frame-Accurate Synchronised Companion Screen Experiences. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '16)*. ACM, New York, NY, USA, 83–92. https://doi.org/10.1145/2932206.2932214

[47] ZDF Mediathek. 2018. Wetterinformationen. Retrieved September 5, 2018 from https://www.zdf.de/nachrichten/wetter

[48] Christoph Ziegler, Christian Keimel, Rajiv Ramdhany, and Vinoba Vinayagamoorthy. 2017. On Time or Not on Time: A User Study on Delays in a Synchronised Companion-Screen Experience. In *Proceedings of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '17)*. ACM, New York, NY, USA, 105–114. https://doi.org/10.1145/3077548.3077557