Deliverable

Project Acronym:	**IMAC**
Grant Agreement number:	761974
Project Title:	*Immersive Accessibility*

D2.1. User Centered Design

Revision: 2.0

Authors: Anna Matamala & Pilar Orero (UAB)

Delivery date: M13

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 761974

Dissemination Level		
P	Public	x
C	Confidential, only for members of the consortium and the Commission Services	

Abstract:

D2.1. documents all activities and their results in the development of ImAC user scenarios and provides detailed user cases for the selected scenarios.

REVISION HISTORY

Revision	Date	Author	Organisation	Description
0.1	22-11-2017	Pilar Orero	UAB	Template
0.2	27-11-2017	Anna Matamala	UAB	Template and some sections
0.3	02-12-2017	Francesc Mas	CCMA	Revision + 1 sub-section
0.4	02-12-2017	Pilar Orero	UAB	Revision
0.5	04-12-2017	All	All	Consolidation
0.6	09-12-2017	Pilar Orero	UAB	Adding RBB info
0.7	14-12-2017	Pilar Orero	UAB	Adding RNIB & USAL info
0.8	18-12-2017	Enric Torres	ANGLA	Revised v. with comments
1.0	19-12-2017	Pilar Orero	UAB	Final version
1.1	02-07-2018	Matamala & Orero	UAB	1st version of 2nd iteration
1.2	19-07-2018	Matamala & Orero	RBB, CCMA, UAB	Revised version for review
1.3	30-07-2018	Enric Torres	ANGLA	Revised version with comments
1.4	31-07-2018	Matamala & Orero	UAB	Final revised version
1.5.	21-101-2019	Matamala &Orero	UAB	Changes suggested by review
1.6	27-02-2019	Anna Matamala	UAB	Revised version for review.
1.7	04-03-2019	Enric Torres	ANGLA	Revised v. with corrections and comments
1.8	05-03-2019	Anna Matamala	UAB	Final revised version
1.9	25-04-2019	Anna Matamala	UAB	Final revised version with copyright-related changes
2.0	19-07-2019	Zora Schärer	RBB	Revision of Conclusions

EXECUTIVE SUMMARY

ImAc follows a user centered design approach, i.e. the project developments are driven by real user needs, continuously involving users in every step. D2.1. describes the user centered design approach in the context of previous accessibility-related research projects and taking into account human rights obligations by EU countries.

First of all, it defines the four areas where development is expected to take place in ImAc: content management, content production, content delivery, and presentation. Next, it defines the process of user identification, in which professional users and advanced home users are differentiated. The interaction of these two user profiles with the different ImAc technological components allows to identify a series of user scenarios, which are the foundation of the ImAc user centered design roadmap. The challenges of accessibility in immersive media are also discussed as compared to traditional access services.

The document reports on the qualitative methodological approach for the focus groups, developed as part of the first iteration. More specifically, it discusses the results for: focus groups on audio description and audio subtitling in the UK, Spain and Poland, focus groups on sign language in Germany and Spain, and focus groups on subtitling in Spain and Germany.

The document also reports on the qualitative pre-pilot actions performed in the second iteration, aimed at fine-tuning the results of the first iteration and as an actual preparation for WP5 pilots. Results are discussed for: focus groups on audio description in the UK and Spain, interviews on subtitling in Germany and Spain, and interviews on sign language in Germany.

A total of 77 users have contributed to define user scenarios, needs, and requirements. Professional users have provided suggestions for professional tools. Regarding subtitles, advanced home users have provided suggestions on subtitle location, non-speech information representation, guiding mechanisms, and comfort field of view, among other features. Concerning audio description and audio subtitling, advanced home users have made recommendations on content selection and presentation, audio features, player interaction with voice commands, and head-mounted device usage, among others. Finally, as far as sign language users are concerned, they have highlighted their needs mostly concerning signer location, comfort field of view, customization, and guiding mechanisms.

.

CONTRIBUTORS

First Name	Last Name	Company	e-Mail
Anita	Fidyka	UAB	anita.fidyka@uab.cat
Anna	Matamala	UAB	Anna.matamala@uab.cat
Belén	Agulló	UAB	belen.agullo@uab.cat
Chris	Hughes	USAL	C.J.Hughes@salford.ac.uk
Doreen	Ritter	RBB	Doreen.Ritter@rbb-online.de
Enric	Torres	ANGLA	enric@anglatecnic.com
Francesc	Mas	CCMA	fmas.z@ccma.cat
Olga	Soler	UAB	Olga.soler@uab.cat
Pilar	Orero	UAB	Pilar.orero@uab.cat
Sergi	Fernández	I2Cat	sergi.fernandez@i2cat.net
Sonali	Rai	RNIB	Sonali.Rai@rnib.org.uk
Sven	Glaser	RBB	sven.glaser@rbb-online.de
Zora	Schärer	RBB	zora.schaerer@rbb-online.de

CONTENTS

TABLES OF FIGURES AND TABLES

LIST OF ACRONYMS

Acronym	Description
AD	Audio Description
AST	Audio Subtitling
ST	Subtitling
SL	Sign Language
ACM	Accessibility Content Manager
CRPD	Convention of Rights of Persons with Disabilities
DoA	Description of the Action
DTV4ALL	Digital Television for All
FOV	Field of View
HbbTV	Hybrid Broadcast Broadband TV
HBB4ALL	Hybrid Broadcast Broadband Television for All
HoH	Hard of Hearing
HMD	Head-mounted Display
VDPC	Video Disk Control Protocol
VR	Virtual Reality

1. INTRODUCTION

This introduction describes the purpose of this deliverable, scope, status, and relationship with other ImAc activities.

1.1. Purpose of this document

This deliverable documents in detail the end user needs, which have an impact on the technology (WP3 and WP4) and on the scenarios for piloting the services (WP5). It describes how user input has been gathered in two iterations through focus groups and pre-pilot actions.

1.2. Scope of this document

Access services in immersive environments may have different behaviors depending on the viewing context, the viewing experience, and the user experience with the service itself. Subtitles may be “burnt” or located according to the viewer’s field of view, or placed in determined regions within the visual content. Audio description would have to be produced depending on the information/directionality given by the sound, the end user requirements, and the way the audio description is received by the end user: as an atmospheric sound or with an ear piece. There are many technical solutions and it is important to understand those who will allow for mainstreaming accessibility.

Working and testing with persons with disabilities requires specific considerations and a qualitative research approach is the best suited to address the different needs and expectations. Still, organising focus groups and interviews with persons with disabilities requires additional preparation [1] [2] with smaller user participation [3] [4]. The reason for this is the time needed to inform everyone [6] and to allow time for presenting the questions and for participants to process information and questions. Also, more time is allowed for participants to express their opinions and thoughts. Balancing time is crucial since participants might experience fatigue [3]. Also, presenting stimuli for persons with disabilities is crucial and requires time and preparation [5]. In the case of access services for immersive media –where there is no available solution at hand—it is even more challenging. For this reason, during the focus group and pre-pilot sessions, different approaches were taken to deal with available materials.

This document details the different stages and efforts taken to define end users, and avoiding “professional informers” [6] while looking for industry professionals: a fine balancing act. Given the lack of stimuli and real life simulation, a great effort was needed to prepare meaningful actions and think of possible scenarios to fulfil a truly user centered methodological approach. These departing scenarios set the needs for deliverables D2.2 and D2.3 and ImAc roadmap. The document describes the project testing workflow, it identifies the technologies to be tested, end user classification, user testing approaches, and finally the results from the focus groups and pre-pilot actions.

1.3. Status of this document

This document is the second iteration of D2.1, with revised content after the review. A first iteration was produced in month 03. This document includes the same content as the first iteration, with some minor revisions, and adds the user centered activities developed after month 03 as part of Task 2.1. It also includes some content suggested during the review. Overall the document presents the user centered design followed in ImAc and the steps taken to identify user needs through a series of actions.

1.4. Relation with other ImAc activities

D2.1. is part of T2.1. (User Centered Design) and feeds directly on T2.2 User Requirements (and its corresponding deliverable, D2.2 User Requirements), which also impacts on T2.3. Platform Specification (and its deliverable, D2.3. Platform Specification). User needs and requirements are the basis for D5.1. Pilot operation plan and D5.2. Pilot evaluation and methodology and plan, which are part of T5.1. Execution and evaluation plan.

2. USER CENTERED APPROACH IN IMAC

Immersive environments (such as virtual reality and omnidirectional video) are often assumed to be of little or no interest for persons with disabilities. This position reveals a lack of knowledge of the Human Rights obligations (CRPD[1]) ratified by all EU countries, as well as overlooking the versatility of access service and the resourcefulness and creativity of end users. Immersive environments exist and they should be accessible for all. During ImAc some test cases will be deployed, showing the available accessibility possibilities and offering real life solutions.

During many previously EU funded projects, access services have been tested to adapt to new technologies such as the switch from analogue to digital broadcast (DTV4ALL) and the new connected media ecosystem (HBB4ALL). As this last project has demonstrated, it is of great interest to all stakeholders within the value chain, that access services are tested while defining and developing the new technology. HBB4ALL started with HbbTV1.0 and finished with HbbTV1.5, enabling the project results to be implemented and providing recommendations for future development in HbbTV 2.0. Impact from the project contributed to standardization bodies, creating an exchange of information and common understanding to meet the needs of the manufacturers. The successful deployment of access services by national broadcasters has also proven that it is possible a common shift in attitudes regarding the importance of access services. ImAc is seeking the same success story within the new immersive environment. This time departing from HBB4ALL and adding the expertise from across the ImAc partners in order to achieve a similar situation. As with HBB4ALL the end user will be at the centre of the project, following a user centered design. The user being both the audience but also the broadcasters or content developer.

ImAc will follow a bottom-up approach, following a user centered methodology. The interdisciplinary team has been chosen from complementary partners with a common aim: **making immersive environments accessible for all**. ImAc groups leaders of previous EU funded projects in order to secure an organised and efficient use of resources. Partners have also been chosen to complement their expertise in the value chain, from development to production, distribution and exploitation. The project has an important Social Science element, looking at innovation and impact beyond technical solutions towards a wider social spectrum. ImAc believes in the role played in society by technology, and how good technical solutions can offer massive benefits in terms of social inclusion and cohesion.

While media companies are still experimenting a lot with new user experiences, ImAc focuses on the development of strategies and tools for enhancing access services. Particularly investigating the characteristics of 360° experiences. During the project the development of new trends in the immersive media market will be monitored, as this will also affect the corresponding access services. If required, the ImAc project will modify its work plan in line with industry to ensure that the final results fit with the market situation.

ImAc aims at developing and integrating and end to end system to create, distribute and display accessibility content as part of omnidirectional video. ImAc development can be grouped in four main sectors where consultation with end users is required: Content Management, Content Production, Content Delivery, and Presentation.

In the first iteration, the workflow was defined as in Figure 1.

[1] http://www.un.org/disabilities/convention/conventionfull.shtml

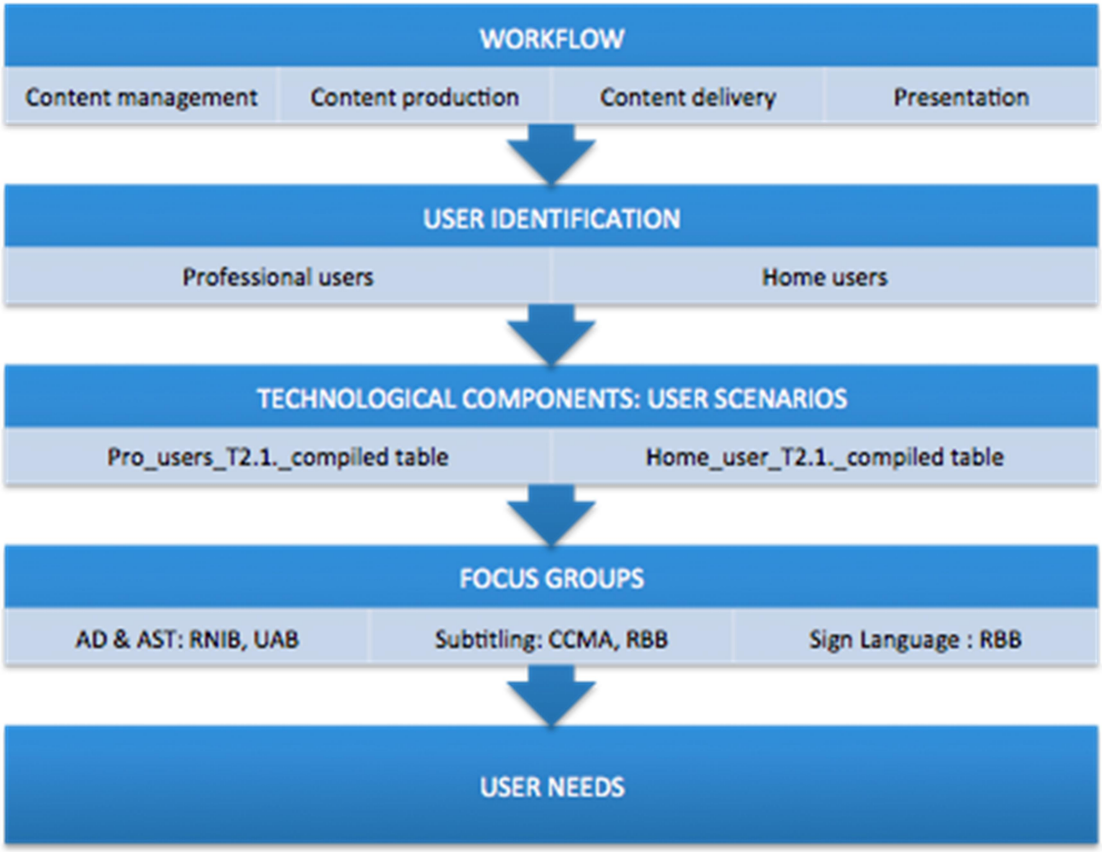

Figure 1 ImAc user centered workflow

In the second iteration, a second round of user centered actions were planned, namely "pre-pilot actions", in the form of both focus groups and qualitative interviews.

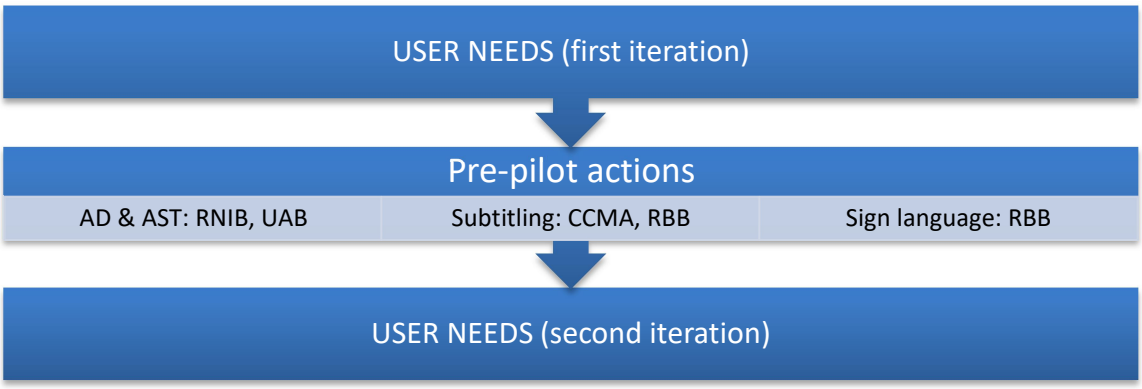

Figure 2 ImAc user centered workflow

2.1. Workflow

First of all, in ImAc we defined the four areas where development will take place longitudinally in the life of the project. These areas are the following.

2.1.1. Content management

In order to handle the storage requirements of the accessibility content and data ImAc will need to deliver a bespoke Accessibility Content Manager (ACM), made available via IP to both direct users and external systems. By hosting the content on a central file system, it will provide control such as logging and access permissions to the content and related data. A web interface will provide users with a mechanism to interact with the content directly (such as a subtitler who needs to upload the content ahead of broadcast) and a web service will provide automated access to processes (such as streaming platforms and playout systems). The ACM will also make it possible to automate tasks centrally such as notifications and make use of protocols used through the broadcast industry. The ACM will provide an essential backbone to the ImAc platform.

2.1.2. Content production

As broadcasters increasingly integrate web distribution channels, new requirements have to be considered in content production. Additional information and metadata about the content is needed to know how to feed each distribution channel simultaneously. ImAc will address this need, firstly with the specification of suitable formats, and secondly with editor and conversion tools that allow the production of accessibility content for all distribution channels and user platforms.

New formats for subtitle, audio description and sign language services will be required, since the current standards only cater for traditional TV broadcast. ImAc will develop solutions that allow for the production and distribution of accessibility content across various channels with the same or minimal extra resources needed on the editor side.

Nowadays, production workflows are designed to supply TV broadcast distribution. ImAc will push towards an inclusion of new services in the production chain. It is important that new services integrate in the current access service workflows to achieve a successful deployment. Format conversion tools developed in ImAc can help when adapting and integrating with existing broadcast systems. Both the developed production tools and the specified formats will meet the requirements defined within the project in order to support the developed services. This follows the understanding that content should be produced only once for economic efficiency. Therefore, all new formats for accessibility data must include enough metadata to supply all distribution channels, new services as well as traditional broadcast services. ImAc will push for rich content annotation during production processes and will provide the tools to do so. For all formats, a very close relation to existing standards is aspired, as it enhances system interoperability and dependability. Both are crucial aspects in broadcasting.

ImAc will evolve production solutions that do not establish additional workflows but instead include enhanced access services for existing system environments. Attractive use cases can then be a driver to develop more enhanced production tools that will support the authoring of enriched content. That content can provide additional value to enhanced platforms while simultaneously serving traditional broadcast playout.

2.1.3. Content delivery

Developments in the project will strongly rely on existing technology and standards, and especially on web technologies which are a central enabler when realizing new user experiences. This is due to high flexibility, the fast-developing market and the cross-platform design of web technologies. Internet streaming has now become a main distribution channel for media content.

Delivering new content types on new platforms require new formats. While audio and video have been in production, accessibility and immersive aspects require extensions to the existing workflows and technologies. This need for new formats is an opportunity to create relevant requirements for standards and for convergence in formats between the production and the end user delivery which, for historical reasons, all use different formats. The approach of ImAc is to leverage the latest standards to cover these new use-cases. ImAc will address the openness of the solution by contributing to the standards and will extend or create tools so that the production content is delivered accurately to the end users, whatever the intended playout.

2.1.4. Presentation

Inclusive TV Content (e.g. with subtitles, audio description, sign language) is usually consumed with a traditional television and via second screen displays. Displaying immersive (virtual or omnidirectional) content has been achieved using players built for specific platforms, mostly Android (Samsung Gear VR, Android DayDream), Windows (HTC Vive, Oculus Rift) or with web players (Facebook, Youtube). ImAc will provide a step forward integrating both experiences from traditional TV viewing and adapted content display. This project will use omnidirectional video enriched with novel techniques for inclusive audiovisual content to deliver immersive content, matching the demands of persons with disabilities when using immersive displays which can be consumed through tablet, mobiles, HMD and traditional TV. A new layer with content adapted to each type of user need (blind, low-sighted, deaf, low hearing, elderly, etc) will provide a personalised experience. For this purpose, we will integrate additional layers into immersive experiences. ImAc will base its developments on ImmersiaTV's player, where it is possible to render an omnidirectional video stream overlapping with complementary and independent augmented layers (including AD), guided with recommendations from the platform itself based upon the user's requirements or by selecting from default presets streams. Audio content (enhanced audio services) will add an additional challenge, as never before have immersive environments been created for blind people. Position based audio fragments and audio indications are the starting point for this innovation.

2.2. User identification

From the four identified areas, it was decided to define who will be the user in the different iterations of focus groups, and in the pilots in WP5. In the national pilots (WP5) it was specified in the proposal that at least 80 users will be tested. From previous EU funded projects DTV4ALL and HBB4ALL, where there was an important effort dedicated to testing component, it was easy to define who will be the end user to be consulted in WP5, and how they will be recruited. In ImAc we have a partner RNIB who will not only help out recruiting, but also advising on methodology and accessibility issues when end user interaction is required.

For the focus groups around 10 persons should be consulted. But these 10 people had to be defined further and differentiated from those in the pilots. To this aim we created two different profiles: that of the professional users and that of the advanced home users.

- Professional users included: IT, graphic designer, subtitler, audio describers, sign language interpreters (signer).
- Advanced home users referred, for instance, to deaf, hard-of-hearing, blind, low vision users, the elderly.

It was agreed that the end users to be included in focus group would be experienced or advanced end users. This meant that, besides their condition of regular lay users, they would also have some knowledge on the technologies that will be developed in the project. In the focus groups there is a fluent conversation to advance their expectations to match the innovation. It would make no sense to consult end users with no knowledge or experience with neither functional diversity nor technological background since at this stage what we require is not their acceptance of the final service, but issues related to technology development.

Some considerations when drafting the information and designing user requirements were related to the following areas.

While some technological components are language agnostic, others related to user interfaces will be decisive to tag at this stage the language dependency. For this project we have the following language classification, which also defines the end user:

1. Oral/written languages:
    a. Catalan
    b. German
    c. Spanish
    d. English

2. Visual-gestural languages:
    a. Catalan Sign Language
    b. German Sign Language
    c. Spanish Sign Language

Hence language definition at this stage is considered to be crucial to understand end user group for testing in the future, the services to be tested and the final (WP5) pilots.

The choice of participants is also dependent of the service for which they will be tested:

- Audio description
- Audio subtitling
- Sign Language
- Subtitling

The end user should also be defined for their sensorial functionality:

- Deaf

- Hard-of-hearing
- Blind
- Low vision

And finally age could also be a factor, since it also often linked to some degree of hearing or vision loss, hence the elderly were also considered.

Further from these considerations, it was crucial to understand from the technological components the type of end user required, which is explained in the next section.

2.3. Technological components: user scenarios

Once we had identified two different end users (advanced home users and professional users), it was required to fine tune end user profiling further and link it to user experience. One table (see Table 1) was created from the previous stage. In this table partners had to identify the technological components to be developed for each WP system side, the end users that would be interacting with the component: professional or advanced home users. And the user scenario they envisaged, i.e. what the already identified user would be experiencing and how.

System side	Technical Component	User: professional user or advanced home user	User scenario: indicate what the previous user will experience and how	input from
Player for mobile phone, TV, head mounted display	ImAc player	Advanced home user	The user will enjoy the experience in the tv, tablet or head mounted display or any combination of these devices. The experience will be synchronized across devices. The player will be available as a web application, so the user won't need to download anything. The tv will start the show, the complementary 360 video is distributed synchronized to the main show, the other devices will show new Audio description and subtitle services adapted to each user impairment The user will access to the contents published in the server and will enjoy an experience adapted to the device where he/she is consuming the content. The user will adapt the device that will store his/her preferences	i2CAT

Table 1: Sample of table structure

Partners contributed with their expertise to identify technological components, users and related user scenarios, that is, what the already identified users will experience and how for each system side and technical component. Each partner uploaded one document with its input and a first version of the compiled table was created (see Annex I)

The table was discussed in a dedicated meeting. To make the discussion more productive, the previous table was divided into two compiled tables depending on the user profile. The meeting was used to further refine the tables, to which partners added extra input in a second revision.

The revised compiled tables can be found in Annex II for professional users and in Annex III for advanced home users

2.4. Challenges of accessibility in immersive media

Access services in immersive media are practically non-existent nowadays: immersive content, including cinematic virtual reality, is still in its infancy and research on this new audiovisual language is ongoing [7, 8, 9, 10]. Broadcasters are experimenting with this medium and, according to a report on virtual reality by the European Broadcasting Union [11] 49% of its members are developing or plan on developing immersive content. The role of the viewer in this new environment is different: the audience does not face a flat screen but is immersed in a sphere where content is all around and can be accessed by simply moving the head. There is a lack of control over viewer's gaze directional behavior because viewers can look at any point in the viewing sphere. This new medium requires a new grammar of filmmaking. As Dooley rightly points out [8]: "Just as filmmakers of the late nineteenth century took some time to experiment with screen grammar and establish the rules of narrative storytelling on the two-dimensional screen, so too are VR developers now exploring a new screen grammar for the 360-degree, interactive space". The techniques for direction attention include movement, sound and lightning cues [7], and are in the process of being developed. Immersive content is also characterised by the absence of a defined frame or shots controlled by the director, which has an impact on access services.

Accessibility and guidance in immersive media are critical for all users, and this is especially relevant for those who cannot access the audio or visual cues. When including access services, the main challenges are mostly related to the fact that users can actually look anywhere and that sounds can come from different places depending on the head position. Some general considerations are included next for each of the services, comparing traditional practices with accessible immersive media.

Traditional subtitling for the deaf and hard-of-hearing reproduces information in the auditory channel that cannot be accessed by certain users. These items include oral verbal elements such as speech or lyrics, but also non-speech information such as sound effects, music and paralinguistic information. Identifying the speaker is also relevant, especially when the character is off screen. In this regard, action in traditional media can happen on or off-screen. In subtitling for the deaf and hard-of-hearing in immersive content, the challenges increase as the on/off-screen dichotomy becomes more complex and audiences can look all around. The main challenges in this regard are then related to where to position the subtitles and how to indicate where the speaker is, while keeping the immersive experience.

Regarding audio description [12, 13], it can be considered a way of retelling a story: the visuals are translated into words and included in the silent gaps, so that an audiovisual content can be enjoyed only auditorily. Critical elements are describing the characters, the actions and the spatial-temporal settings. The changes in the way stories are told in immersive environments

will also imply a change in the way stories are retold in audio description, and this is one seminal aspect that ImAc will tackle through the use of phrasing and sound. Instead of focusing on a 2D scene, audiences can wander around 360º, focusing either on the main story or on secondary actions, while the temporal space available for audio descriptions is still restricted by the temporal development of the film. Therefore, deciding what to describe, when to describe it and how are critical elements that will need to be thoroughly discussed.

Concerning audio subtitling [14, 15] they serve those audiences who cannot access the written subtitles and cannot understand the language of the original by providing an oral version of the written interlingual subtitles. The challenge in immersive media is to define how these audio subtitles can enhance the user experience, by making the most of sound technologies.

Sign language interpreting is one of the three TV mature access services along subtitling and audio description. The challenge of defining where to position the signer in traditional broadcasting takes a new dimension in the immersive world, and new possibilities may arise. The search for existing 360º solutions has given a unique response, that of a videogame using a mouse who communicates with sign language (see https://kotaku.com/people-are-falling-in-love-with-a-video-game-mouse-who-1797534518). Apart from this anecdote, there is no other example, neither as an avatar or real sign language interpreter, and display solutions need to be identified and tested.

There are many challenges that need to be addressed during the technological development and service implementation, but a necessary first step in a user-centric project is asking users and involving them. Users at ImAc go beyond the public, since solutions across the broadcast chain flow will be developed, and will need to be tested towards final exploitation in the industrial broadcasting and media content sectors.

Media access services for video format in a convergence ecosystem present a wealth of opportunities. From the hybridisation of services (audio subtitles, easy to read subtitles, object based audio description, etc) to the interaction modes, and the personalisation opportunities.

On the one hand, it will define the many solutions and opportunities to deploy access services in the new 360º media environment. This stays at paper prototype level, with descriptive features and requirements. A second intervention will identify the most challenging services, and will match the user expectations with the technical possibilities towards testing a first prototype. A pilot will test mature technology and content.

2.5. Focus groups: understanding user needs (first iteration)

Once we had the two groups of end users (advanced home and professionals) the technology they will be testing, the user scenarios and the challenges immersive media pose to accessibility, there was the need to define:

- which partner would conduct the focus groups, with
- which end users and how to interact with them, and
- the procedure.

Based on the existing accessibility enhancements for TV-centered and web based services, the aim of the focus group was to gather information about user needs, expectations and wishes regarding the various technological components to be defined. Taking into account that access services are almost non-existent in immersive media, it was considered fundamental to gather feedback from end users for these new scenarios through this qualitative method.

Various methodological tools such as interviews, questionnaires or focus groups could be used to gather user feed-back, but at this stage of the project focus group with experienced users was favoured.

2.5.1. Focus group: partner and service distribution

It was decided that focus groups would be conducted replicating as much as possible the division made for piloting in WP5. It would make no sense for example for RBB to do a focus group on audio description since they will not pilot the service in WP5.

To this aim the division is by services:

Service	Partner conducting FG	Participating partners
Audio description	RNIB (English) UAB (Catalan and Spanish)	ANGLATECNIC, CCMA, IRT, RNIB, UAB
Audio subtitling	RNIB (English) UAB (Catalan and Spanish)	ANGLATECNIC, CCMA, IRT, RNIB, UAB
Sign Language	RBB, CCMA (German, Catalan and Spanish Sign Language)	RBB, CCMA
Subtitling	RBB (German) CCMA (Catalan and Spanish)	ANGLATECNIC, CCMA, IRT, RBB, UAB

Table 2: Focus group by service and participating partner

In Spain there are six official languages: Spanish, Spanish Sign Language, Catalan, Catalan Sign Language, Basque and Galician. Tests in ImAc are performed in Catalan or in Spanish or both. Users were accessed through different user association, with no specific requirements in terms of language, respecting the official bilingualism in Catalonia. CCMA content is produced in Catalan, but tests led by CCMA were prepared with both Catalan and Spanish documentation. All participants agreed to fill in the test documents in Catalan, although Catalan and Spanish were used indistinctly to interact and communicate between organisers and participants.

Beyond the scope of the project, a focus group was also conducted in Poland for audio description and audio subtitling by UAB. Although not initially included in the proposal, it was considered that involving users from a voice-over country such as Poland could yield additional results.

2.5.2. Focus group: end user profiling

Concerning end users, it was agreed that focus group would gather input from both advanced home users and professional users.

Profiling end users was considered necessary, and a specific short questionnaire to be used during the project was developed and translated into the end user languages. The

questionnaire includes some basic questions, validated with end users, that can be expanded in future tests but which offer some shared basic information about the users.

The questionnaire was produced first in English (see Annex IV) and translated into German, Catalan and Spanish.

2.5.3. Focus group: procedure

ImAc produced a shared common procedure for Focus Groups administration and management. A special form was created with general instructions to all partners in English managing a Focus Group to be found in Annex V.

It was decided to hold a focus group with reduced numbers of participants, following the recommendations on qualitative research design [16], as this facilitates their participation. The approach taken was therefore qualitative rather than quantitative. In the first stage, it was also decided not to use ImAc prototypes but various types of materials (for instance, mock-ups, spoken descriptions or visualisations of existing access services from the TV domain) to gather creative ideas from users and not direct them only towards ImAc solutions.

All partners had to write down a procedure following a specific model to be found in Annex VII.

The focus group development was as follows:

- Welcome participants
- Information sheet and consent forms to be read and signed
- Demographic information: short questionnaire administered
- Group discussion following questions from compiled tables (see Annex VI)
- Summary of focus group conclusions, to be approved by all participants.

At all times in ImAc testing, piloting or consulting end users follow the Ethical Considerations drafted in a separate document D.1.2 Ethical Considerations and Data Protection Management, to be found also in Annex VII.

2.6. Fine-tuning user needs: pre-pilot actions (second iteration)

Many user needs, requirements and wishes arose in the first iteration through the focus group but there was a recurrent setback: users gave suggestions based on their intuitions rather than real experience. Many users found it difficult to imagine the actual implementation of their proposals. This is why before the actual pilots, it was agreed to carry out a series of pre-pilot actions aimed at gathering more extensive feed-back on specific issues that arose during the focus groups. The aim was to offer participants actual ImAc prototypes that could be used for further discussion. Their feed-back would allow to fine-tune user requirements and to narrow down the list of possible testing options in the actual pilots. A qualitative approach was again favoured, and a reduced sample of users was used. This is why results are presented in a descriptive way and no inferential statistics are performed on the data.

2.6.1. Pre-pilot actions: partner and service distribution

It was agreed that the partner and service distribution would be identical as in the focus group (see Table 2). It was agreed that that users should be asked for more feed-back concerning the following aspects.

Regarding subtitles:

- Comfort viewing field, by providing examples in six field of view levels in a 16:9 ratio (30% to 80%) (see Figure 3).
- Guiding user to speaker, by using three strategies: sided text, arrow and compass.

Regarding sign language:

- Comfort viewing field, by providing examples in six field of view levels, similar to subtitling test (see Figure 4).
- Guiding user to speaker using arrow below signer video window and signer video left/right. Plus forced perspective: automatic change to speaker only once in the beginning.

Regarding audio description (and audio subtitling):

Three versions of audio description, namely: AD placed on the action (privilege of sound), AD anchored to head position (voice of God), and AD anchored to soundscape (first person, past tense).

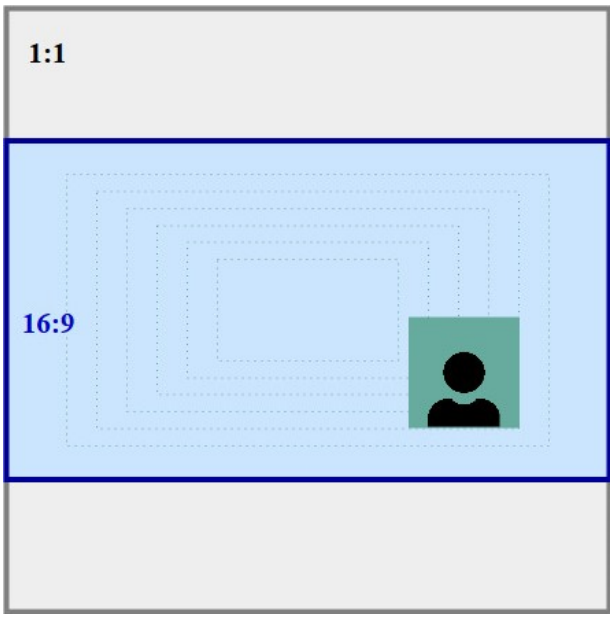

Figure 3 Comfort field of view for sign language

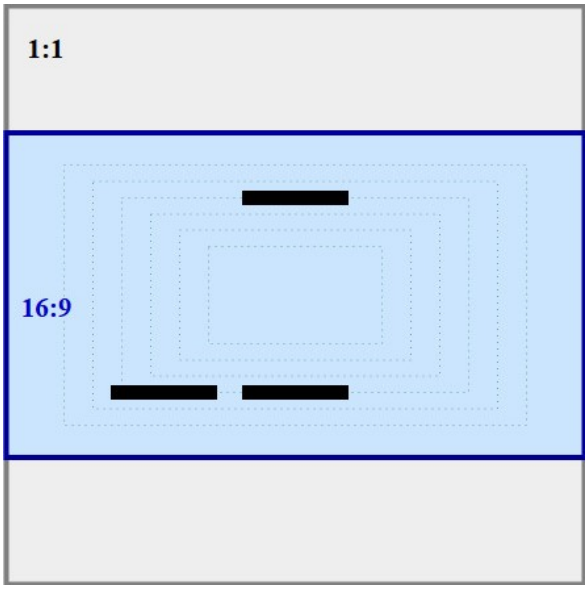

Figure 4 Comfort field of view for subtitling

2.6.2. Pre-pilot actions: end user profiling

Due to the specificities of the methodology used to assess each access service it was decided that for audio description home users would be involved in a focus group.

Regarding subtitles and sign language, a qualitative semi-structured interview was prioritized with some stimuli followed with questions to trigger discussion. The same users as in the first iteration were targeted, although an additional pre-pilot was planned at UAB on subtitling in which the users were hearing persons with different degrees of technical capabilities but no hearing impairment. The aim of such pre-pilot action was to provide additional feed-back and to assess whether a user profiling based on capabilities rather than on disabilities would work better. To that end the demographic questionnaire had to be adapted (see Annex XIV).

In both cases a qualitative approach was favoured, with a suggested number of participants around 5. Users were profiled using the same questionnaire as in the first iteration (Annex IV).

2.6.3. Pre-pilot actions: procedure

Pre-pilot actions were conducted in April and May 2018. Two approaches were adopted: for subtitling and sign language testing took place on an individual basis, due to the methodological constraints, and for audio description the focus group format was kept.

Different types of videos were prepared to gather feed-back: i2Cat prepared five videos for subtitling and sign language for the different test conditions (interview with Young Hurn during 1st "Rapzember" event) and IRT prepared three different audio mixes for the AD test [private_content][2] based on the audio descriptions provided by RNIB in English and translated into Catalan by UAB.

[2] The name of the video used has been substituted in all public documents delivered by ImAc due to copyright issues.

For the subtitling pre-pilot tests, two videos per language were prepared. They included still images to guide the user through the test condition and asked them to rate each condition on a 1 to 5 scale. For the sign language pre-pilot tests, three videos were prepared, including also still images to guide users. For the audio description pre-pilot tests, three versions per language were prepared. A blank image at the beginning was included so that participants could have time to put on the cardboard glasses and adjust them.

Specific instructions, with a detailed experimental procedure, were produced for each of the access services. They are available in Annex VIII (subtitling), IX (sign language) and X (audio description). Reporting templates were also produced for all three services (Annexes XI, XII and XIII).

Snapshots of the prototypes are provided below.

Figure 5 Guiding mechanisms with arrow

Figure 6 Guiding mechanism with compass

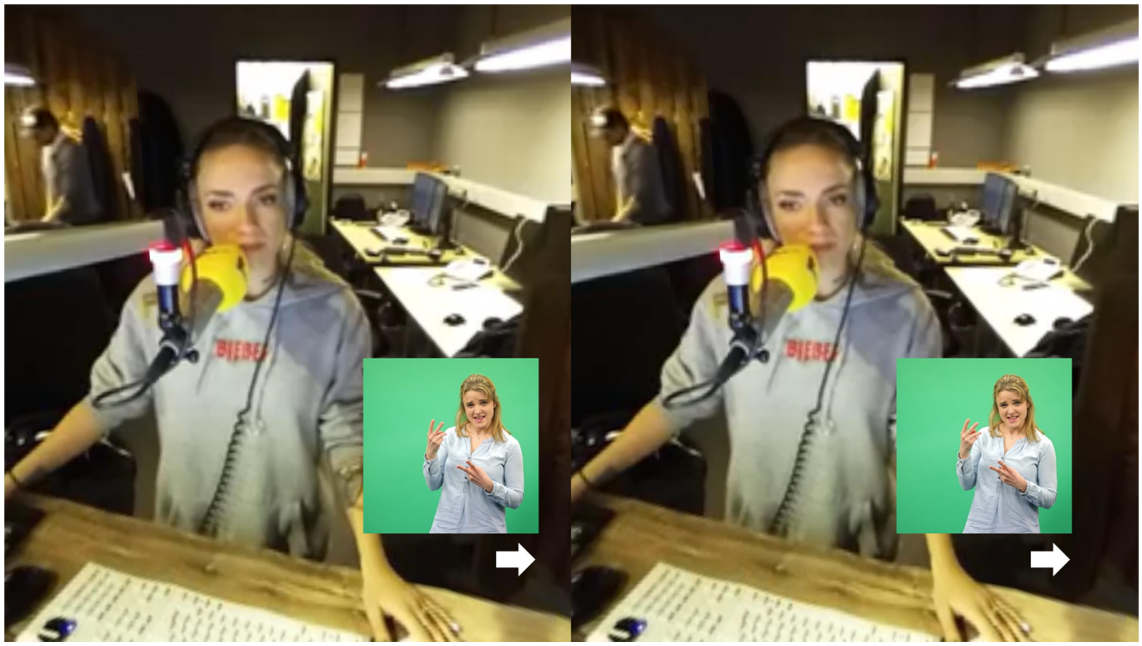

Figure 7 Sign language pre-pilot snapshot

3. RESULTS: USER NEEDS

The results of both iterations are presented in this section. Results are presented according to the service, providing first the results of the first round of focus group and then the results of pre-pilot actions. Both iterations have identified end user interest, needs and expectations from the services and technologies to be developed in ImAC.

3.1. Audio description and audio subtitling

To mirror the piloting in WP5 three different focus groups and two pre-pilot actions were formed to gather information regarding audio description and audio subtitling.

3.1.1. First iteration: AD and AST in Catalonia

A press release of the Focus Group can be consulted here: http://blogs.uab.cat/blogdtieao/
The focus group report, with all the notes and comments from participants, is available here: https://drive.google.com/open?id=18c-67PGrtlFg8bwCwbLTmP7Asw8RSD4O

Figure 8 Participants in UAB Focus Group

Focus group general information

- Partner responsible for the workshop: UAB.
- Place and date: UAB (Bellaterra, Barcelona), 22.11.17
- Access service(s) discussed: AD and AST.

Participants profile

- Number of home users: 2.
- Number of professional users and profile: 4 (3 audio describers, 1 technical expert).
- Demographics for users:
    1) Sex: 4 “male”, 2 “female”, 0 “other”, 0 “prefer not to reply”.
    2) Age: 25, 25, 30, 34, 51, No Reply.
    3) Main language of the participants: 1 Spanish, 1 Sp/Cat, 3 Catalan.
    4) Level of finished studies: 0 “no studies”, 0 “primary education”, 0 “secondary education”, 1 “further education”, 5 “university”.
    5) (Only for home users): “I define myself as a…” 0 “blind person”, 2 “low vision person”, 0 “deaf person”, 0 “hearing impaired person”, 0 “deaf-blind person”.

6) (Only for home users): Age in which your disability began: 0 “From birth”, 1 “0-4”, 0 “5-12”, 1 “13-20”, 0 “21-40”, 0 “41-60”, 0 “More than 60”.
7) What technology do you use on a daily basis? You can select more than one. 4 “TV”, 2 “PC”, 5 “Laptop”, 4 “Mobile phone”, 3 “Tablet”.
8) Do you have any device to access virtual reality content? 0 “Yes”, 6 “No”, = “I don’t know or I don’t want to reply”.
9) (Only for home users): Which of the following do you use on your connected devices to access the above content? 1 “Magnification”, 1 “Screen reader”.
10) (Only for home users): Which of the following controls would you like to use with your screen reader/magnification tool when watching content online? 1 “Identify content”, 2 “Functions such as play, stop, pause, forward, rewind”, 2 “Switch AD/AS on and off”.

Summary of participants’ profile

Six users took part in the focus group (2 home users, 3 audio describers and 1 technical expert). They were 4 males and 2 females, with ages ranging 25-51 and one participant not providing this data. Five participants had university studies and 1 had further education studies. None of them reported having a device to access virtual reality content. A laptop was the most used technology by the participants on a daily basis (5), followed by TV and mobile phone (4), tablet (3) and PC (2). The home users were low vision participants (vision impairment started between 0-4 and 13-20, respectively). One used a magnification tool to access content and the other one, a screen reader. Both participants identified “functions such as play, stop, pause, forward, rewind” and “switch AD/AST on and off” as needed to watch content online, and one also considered “Identify content” as needed.

Conclusions from professional users

Professionals think that AD in immersive environments is challenging because there is more information than in a standard non-immersive film. It is challenging to decide what to audio describe, especially when there is not enough time for the AD. Time management will also be different, because a user can spend more time watching a film.

Professionals think that the space should be described first but it would be interesting to offer simultaneous audio descriptions of different sections of the scene. This will increase the number of audio description units, which will not always be activated, and this may raise some financial issues.

Professionals consider that the type of content will determine what is feasible: a virtual space gives more time for audio description but standard films are already challenging, so immersive films will also be even more difficult.

The audio describer needs to have a general view of the scene (plain view) and the possibility to select different sections (minimum of 4, generally 6, maybe more), which can be opened in new windows.

Audio describers think it would be useful to watch the content with glasses first but do not want to work with an immersive editor, just to check the final outcome.

Concerning current software, audio describers indicate that some functions are only used once and could be deleted from the general view (for instance, voice calibration). Others functions are not used by our professional users (“*thermometers*”, “check if mike works”).

Conclusions from advanced home users

Concerning the interaction: accessing the services

Users in our focus group consider that voice interaction would be better than video commands, especially if the user is in a private environment, which is where normally this product consumption takes place.

Users in focus group consider that it would be useful that the system identifies user preferences and parameters. Some of these preferences could be transferred between devices but not all of them because the user may have different preferences for a smaller or bigger screen.

Regarding accessing immersive content with or without glasses, users indicate that it depends a lot on the home user.

Concerning the services

Advanced home users consider that 360 immersive content is an interesting technology and implementing accessibility is needed. It is an actual improvement to the content.

Advanced home users consider that enlarging images is a requirement.

Advanced home users think that immersive sound could be useful to position yourself and identify where action is happening. Information could be prioritized according to the volume. Information could be given like “headlines” and then, if you are interested, you turn your head to that area and the volume then increases.

Advanced home users favour the AD position linked to action being described. Users want a main audio description of the main action (so that they can follow the plot), but they also want to be able to choose different secondary audio descriptions for additional action, even if this means the film will last longer. They explain the possibility of watching the film different times and choosing different paths. They also indicate the usefulness of having two different voices: one for the main action and one for the secondary actions.

Advanced home users do not want more information than the one received by persons with no visual impairments, so professionals do not think the action happening outside the view of the user should be described unless specific action is taken by the user (moving the head).

Advanced home users indicate the challenge of moving away from the main action and going back to it. They think a specific sound effect (“beep”) could be good to position themselves and know they are back to the action.

Advanced home users express the need for audio subtitling (spoken subtitles), better than zooming in on the subtitles.

3.1.2. First iteration: AD and AST in the UK

The whole report is available here: https://drive.google.com/open?id=15F4go5ko-IC_HUiDVDZ8zLqaW34-zRdJ

This report concerns the accessibility of 360 degree content for audio description, as tested by a focus group of blind and partially sighted people. The aim of the focus group was to gather feedback from regular users of audio description on viewing and interacting with 360 degree content in an immersive environment.

A review of immersive environments outlined four key areas on which feedback was sought during the focus group. On the form of the audio description track: linear, based on the traditional approach where viewer sees what the director wishes to show on the screen and non-linear, wherein view changes when viewer interacts with the environment, and using audio to enhance the immersive experience. Lastly, issues around accessing the 360 degree content were addressed.

In the absence of any audio described 360 degree content, live description was delivered on a clip available on Youtube in a 360 degree format. We wished to explore how a small group of participants would respond to the material.

Focus group general information

- Partner responsible for the workshop: RNIB and USAL.
- Place and date: RNIB (London, UK) 7 December 2017
- Access service discussed: AD.

Participants profile

- Number of home users: 7.
- Number of professional users and profile: 2 (1 audio describer, 1 technical expert).
- Demographics for home users:
    1) Sex: 4 “male”, 3 “female”, 0 “other”, 0 “prefer not to reply”.
    2) Age: 26, 28, 32, 36, 39, 49, 53.
    3) Main language of the participants: 7 English.
    4) Level of finished studies: 0 “no studies”, 0 “primary education”, 0 “secondary education”, 0 “further education”, 7 “university”.
    5) “I define myself as a…” 5 “blind person”, 2 “low vision person”, 0 “deaf person”, 0 “hearing impaired person”, 0 “deaf-blind person”.
    6) Age in which your disability began: 7 “From birth” (they said the level of sight they had, had deteriorated over the years), 0 “0-4”, 0 “5-12”, 0 “13-20”, 0 “21-40”, 0 “41-60”, 0 “More than 60”.
    7) What technology do you use on a daily basis? You can select more than one. 7 “TV”, 1 “PC”, 6 “Laptop”, 7 “Mobile phone”, 6 “Tablet”.
    8) Do you have any device to access virtual reality content? 0 “Yes”, 7 “No”, 0 “I don’t know or I don’t want to reply”.
    9) Which of the following is your preferred device for watching online video content (i.e., Youtube, Vimeo, Netflix, Amazon Prime, broadcast catch up service etc.)? 1 “PC”, 6 “a combination of smartphone and tablet”.
    10) Which assistive technology do you use? 1 “Magnification (i.e. Zoomtext)”, 5 “Screen reader (i.e. Zoomtext)”, 1 “a combination of magnification and screen reader”.
    11) Which of the following controls would you like to use with your screen reader/magnification tool when watching content online? 7 “Identify content”, 7 “Browse content library”, 7 “Functions such as play, stop, pause, forward, rewind”, 7 “Switch AD on and off”.
    12) Which of the following describes what you are able to see? 1 “well enough to recognise a friend at arm’s length”, 1 “well enough to recognise a friend if they got close to his or her face”, 3 “the shapes of the furniture in a room”, 5 “can tell by the light where the windows are”, 2 “cannot see anything at all”.
    13) Which of the following barriers do you encounter when watching TV? 7 “difficulty seeing buttons on the remote control”, 5 “difficulty seeing the picture on the TV screen”, 7 “difficulty seeing the fine detail on the TV screen”,

7 “difficulty seeing text on the TV screen”, 5 “see the light of the TV screen”, 2 “cannot see anything on the TV screen”, 7 “participants find it difficult to follow what is going on, on the screen”.

14) How do you currently watch or follow a programme or film on television? 1 “uses residual sight to watch”, 1 “sits closer to the TV screen”, 6 [not the preferred option] “ask their friends or family members to assist by explaining what happens on the screen”, 6 [not the preferred option] “try to pick up as much as they can from the sound of the film or programme”, 6 “use audio description to explain to them what happens on the screen”.

Summary of participants’ profile

Seven home users took part in the focus group and 2 professional users (1 audio describers and 1 technical expert). The home users were 4 males and 3 females, with ages ranging 26-53. All home users had university studies. None of them reported having a device to access virtual reality content. TV and mobile phone were the most used technology by the home users on a daily basis (7), followed by laptop and tablet (6), and only (1) used PC. 5 home users were blind from birth and 2 had low vision also from birth (the level of sight they had, had deteriorated over the years). Most home users preferred to watch online video content using a combination of smartphone and tablet (6) and only 1 preferred PC. To access content 1 home user used a magnification tool, 5 a screen reader and 1 a combination of magnification and screen reader. All home users identified “Identify content”, “Browse content library”, “functions such as play, stop, pause, forward, rewind” and “switch AD/AST on and off” as needed to watch content online. 2 home users cannot see anything at all and 5 can tell by the light where the windows are, 3 can see the shapes of the furniture in a room, 1 can recognise a friend if they got close to his or her face and 1 can see well enough to recognise a friend at arm’s length. All home users have difficulty seeing buttons on the remote control, seeing the fine detail on the TV screen and find it difficult to follow on the screen what is going on, of them 5 have difficulty seeing the picture on the TV screen and the light of the TV screen and 2 cannot see anything on the TV screen. Finally, 6 home users use audio description to explain to them what happens on the screen, and also 6 home users, although not the preferred option, ask their friends or family members to assist by explaining what happens on the screen or try to pick up as much as they can from the sound of the film or programme, 1 uses residual sight to watch and 1 sits closer to the TV screen.

Results

Task 1

Aim of task 1: Understand how AD fits with linear storytelling and then how the interactive aspects of 360 degree videos impact the AD track.

(Cue: In conventional television we can assume that people are looking directly at a TV screen in front of them and the audio description describes what’s on the screen. However, what do we do with the audio description in a 360° experience if the viewer can be looking in any direction?)

Content
Title: Attenborough and the Giant Dinosaur - BBC One
Length: 4 mins (approx.)
Link: https://www.youtube.com/watch?v=rfh-64s5va4
Description delivered live by Describer Roz Chalmers

Feedback
Participants praised the live audio description, it was agreed that the track comprised all those aspects of the video that had lacked audio clues and completed the picture.
“The important thing is the description complemented the narration, it wasn’t repetitive and it wasn’t overly descriptive.”

“I think it managed to describe everything that was on the screen even though there wasn’t much time.”
Although the description was appreciated in terms of making the story clear, some felt it lacked the elements to build the atmosphere needed for an immersive experience.
“What I didn’t get from the description was the ambience – you hear the rustling of the foliage, large clomping feet, what’s the weather like, what birds can you see in the sky? This is where your imagination comes in.”
“I still can’t get a proper visual picture of the dinosaur in my head, so the colour, has it got a spikey back?”
Here the describer pointed out that the brief gaps in narration meant there was a need to prioritise what could be described. “You got what the sighted viewer was getting.” It must be noted here that NO ONE wanted the audio description to go over the voiceover.

Task 2
Aim of task 2: Understand how AD fits with the interactive aspects of 360 degree videos.

(Cue: In conventional television we can assume that people are looking directly at a TV screen in front of them and the audio description describes what’s on the screen. However, what do we do with the AD in a 360° experience if the viewer can be looking in any direction?)

Content
Title: Attenborough and the Giant Dinosaur - BBC One
Length: 4 mins (approx.)
Link: https://www.youtube.com/watch?v=rfh-64s5va4
Description delivered live by Describer Roz Chalmers

In order to emulate a 360 environment, cursor provided within the video was used to shift focus to different views. For example, instead of watching David gaze at the dinosaur during the first break in voiceover, view panned left to the mountains in the distance and consequently the audio description changed as it referred to the on-screen elements now in view. This resulted in participants mostly feeling disconnected with the storyline and found the description to be repetitive and missing important details.

“I thought that sounded disjointed and for me if it is a description of 360 then it would take a lot more than just words. This didn’t give me anything really.”

While the 360 degree movement seemed to enhance the overall immersive experience visually, it was almost impossible to simulate that experience in audio given the brief gaps in voiceover. Some participants commented that it was difficult to comprehend a 360 view.

“I don’t understand, was it vertical now, like a vertical axis, looking down? Was that the view? Was it from the perspective of the dinosaur?”

“For some people who can’t see and never have, it is already a challenge to understand 360.”

"I don't believe this! Where is the TV screen? I can't get my head around it."

Participants agreed that a lot more would be needed than just audio description to make the environment more immersive for people with significant sight loss and that getting the right balance without information loss would be difficult.

Task 3
Aim of task 3: Use spatial audio to create the illusion of sounds all around the subject thereby creating a 360 degree environment.

Content
Title: Virtual Barber Shop
Length: 5 mins (approx.)
Link: https://www.youtube.com/watch?v=IUDTlvagjJA
Speakers were used to allow the focus group to participate as a group.

Participants responded enthusiastically to this clip, which had no visual content:

"It's great! I've heard it before and it's a million times better with headphones!"

"For me this was immersive!"

"You're actually in the room!"

"It was very impressive that last one, I found it very interesting. I know who's around this table because I'm in the room. Felt as if it was happening around me!"

"Honestly that's what I need to be completely immersed."

It was noted that the narration in the clip integrated elements of audio description within it. For example, "now I'm moving to the right." "Look at my pair of scissors". This combined with the perception of depth delivered in spatial audio enhanced the experience for the focus group.

Task 4
Establish user preference for tools required to access 360 degree content.

Most participants agreed that a HMD would be unnecessary as audio would be the key feature for them. However technical expert participating in the group pointed out the significance of the headset which may be used to track head movements and subsequently trigger specific descriptions.

The issue of integration with assistive technology tools such as speech readers and magnification were also discussed briefly. It must be noted here that all participants were regular users of video on-demand services such as Netflix, Amazon Videos, BBC iPlayer which are set up to work with assistive technology. Participants reiterated that this was essential to allow independent access.

However, in view of the characteristics of the immersive environment that may not allow control via traditional tools such as keyboard and mouse, voice control was considered most

appropriate. Once again, members of the focus group have previously used voice control on smartphones i.e., VoiceOver on iOs and Talkback on Android devices.

Discussion

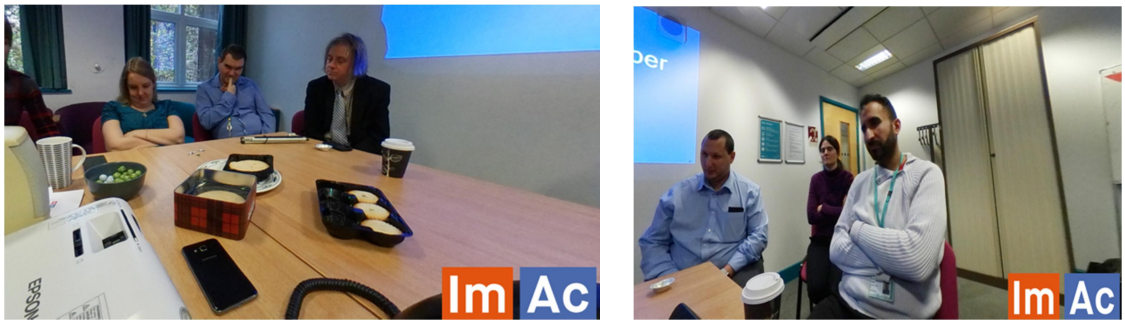

Figure 9 Focus group in UK

For 360 Image: http://imac.crazysandbox.co.uk/360photo/?img=360_0262

Overall, the focus group agreed that an audio led immersive environment was easier to comprehend than an audiovisual environment for dedicated users of audio description. It was felt that the visual display of a 360 degree environment was somewhat irrelevant without elements of it figuring in the audio description track. Head mounted displays were regarded as unnecessary by the group with the exception of one participant who is not a regular audio description user.

Participants strongly felt that the description track needed to complement the main narrative and any deviation from the primary storyline would lead to unnecessary disorientation. The five key elements of audio description - who, what, why, where and when - were prioritised over the description of 360 degree elements by the participants. These were considered more acceptable when offered in the immersive audio environment of the Virtual Barber Shop where sound design was used to set the scene in a 360 degree view but the script followed a single narrative with integral clues on the setting.

It was also felt that since such content was meant to be consumed independently therefore the audio description script could be written in second person i.e., you're only a few paces away from the stage where the band is playing, etc. Professional describer added that this may help pull listeners into the scene and enhance the immersive experience.

The group discussed various factors that could contribute to immersive experience including how many voices would be considered too many in an immersive environment and whether directionality and placement of the audio description would impact the viewer experience for people with sight loss i.e., audio description could be the voice in your ear or coming from somewhere behind you. However, no definite conclusions could be reached in the absence of samples.

On the subject of accessing content, there was a consensus on using a combination of voice control and integration with assistive technology tools – magnification and speech readers. Technical expert in the group however pointed out that immersive experiences are unlikely to support the use of traditional equipment such as keyboard/mouse therefore voice control may be the only way to access content. It must be noted here that some participants had previous

experience of using voice controlled environments such as Amazon Fire TV, Alexa and Google Home.

Further research, which includes specially produced content and a wider focus group comprising of people with different sight levels could clarify the importance of sound in an immersive environment.

3.1.3. First iteration: AD and AST in Poland

Figure 10 Focus group in Poland

The focus group in Kraków was conducted on 28 December 2017. UAB was the partner responsible for the workshop.

Participants profile

- Number of expert end users: 3. Number of professional users: 3 audio describers.
- Demographics for users:
  1) Sex: 2 “male”, 4 “female”, 0 “other”, 0 “prefer not to reply”.
  2) Age: 46, 37, 25, 31, 43, 33.
  3) Main language of the participants: 6 Polish.
  4) Level of finished studies: 6 “university”.
  5) (Only for expert end users): “I define myself as a…” 3 “blind person”, 0 “low vision person”, 0 “deaf person”, 0 “hearing impaired person”, 0 “deaf-blind person”.
  6) (Only for expert end users): Age in which your disability began: 2 “From birth”, 0 “0-4”, 1 “5-12”, 0 “13-20”, 0 “21-40”, 0 “41-60”, 0 “More than 60”.
  7) What technology do you use on a daily basis? You can select more than one. 1 “TV”, 3 “PC”, 5 “Laptop”, 6 “Mobile phone”, 2 “Tablet”.
  8) Do you have any device to access virtual reality content? 1 “Yes”, 5 “No”, 0 “I don’t know or I don’t want to reply”.
  9) Which of the following is your preferred device for watching online video content (i.e., Youtube, Vimeo, Netflix, Amazon Prime, broadcast catch up service etc.)? You can select more than one.
     1 “PC”, 4 “Laptop”, 2 “Smartphone”, 1 “Tablet”, 0 “I don’t watch online video content”, 2: Smart TV, External monitor connected to laptop “Others”
  10) (Only for expert end users):Which of the following do you use on your connected devices to access the above content? “Magnification”, 2 “Screen reader”, 1 “Both”, 0 “None”.
  11) (Only for expert end users): Which of the following controls would you like to use with your screenreader/magnification tool when watching content online? 2 “Browse content library”, 2 “Identify content”, 3 “Functions such as play, stop, pause, forward, rewind”, 3 “Switch AD/AS on and off”.

Summary of participants' profile

Six users took part in the focus group (3 end users, 3 audio describers). They were 2 males and 4 females, with ages ranging 25-46. All participants had university studies. One of them reported having a device to access virtual reality content. Mobile phone was the most used technology by the participants on a daily basis (6), followed by laptop (5), PC (3), tablet (2), and TV (1). The end users were blind participants: vision impairment started from birth (2) and between 5-12 (1). Two blind participants reported using screen reader to access content and the third one both magnification and screen reader. All participants identified functions such as "play, stop, pause, forward, rewind" and "switch AD/AST on and off" as needed to watch content online, and two of them also considered "Browse content library" and "Identify content" as needed.

Workshop conclusions from professional users

Professional audio describers think that AD in immersive environments is more challenging than AD in standard non-immersive video materials. They consider that immersive environments are interactive, meaning that the user can choose which parts of the visual landscape to consume and that blind users should also be able to consume the contents in such a way.

Professional audio describers consider that the visual scene is much larger. There is much more possible information to convey than in a standard non-immersive video content. From professional point of view, it will considerably increase the amount of work needed. They also consider that providing audio description to immersive content will be much more challenging from financial perspective, as there will be more possible audio description units.

Professional audio describers consider that the question of how to remunerate audio describers writing 360 contents is crucial.

Professional audio describers think that audio description of the main action should be provided as a priority to allow the blind and partially-sighted to follow it, but audio description of the surrounding visual scene should also be offered. During the focus group, three possible ways of implementing AD in immersive environments were suggested by professional audio describers, but no technical solutions were proposed in this regard:

1. Pausing the video: to offer audio description of the main action with the possibility to pause the video material and listen to audio descriptions of the surrounding visual scene. It creates the possibility to watch the film many times and listen to different AD units in a personalized way. It will result in different time management, as user can spend much more time watching the film. This is why this possibility will mostly concern video materials consumed at home, not in institutions such as museums. This will also increase the number of audio description units which not always will be activated.
2. Simultaneous AD: to offer audio description of the main action with simultaneous audio descriptions of the surrounding scene. It raises the issue of how to come back to the main action.
3. Visiting paths: professional audio describers added that, taking into consideration the amount of possible work, the video materials should first be carefully selected and audio description should be produced in close cooperation with the creators of 360 contents. It was suggested that making such video materials accessible should be an obligation of the content creators. Visiting paths should be prepared: Users should have the possibility to choose the option ´guide me´ which would guide them through the main action.

Professional audio describers think that they would not like to work in an immersive editor (because of the difficulties related to writing on keyboard and marking timecodes). A general view of the visual scene would be needed in a plain screen, with a possibility to write audio description in windows linked with different sections of the visual scene.

Text-to-speech module that would allow proof-reading with glasses: As in Poland audio description is recorded and read aloud by a professional, *lektor*, audio describers point to the necessity of having a module with a text-to-speech software in the Accessibility Content Manager to check the final outcome of audio description in glasses before it is recorded.

Concerning current programs, professional audio describers use in their daily work Microsoft Office and AD Maker (.doc and .srt files)

Audio describers consider that information could be prioritized according to the volume (audio description of the main action could be played louder). Professional audio describers suggested that two voices could be used: male and female, respectively for main action and surrounding visual scene. Professional audio describers suggest that AD could be given like headlines and then the user can turn their heads towards it, but end users were more interested by being guided by binaural sound.

Professional audio describers consider that they would need to gather information about all possible types of video content (documents, fiction films, simulations) and venues for immersive contents (use at home, museums, training) to gather all challenges and determine what is feasible in each case. It was suggested that, while watching a film, the most relevance is placed upon what the film director wants to convey and not to look around. In simulations or city tours – looking around is important.

Workshop conclusions from advanced end users

Two groups of conclusions have been drafted as follow:

Concerning the interaction: accessing the services

Screen readers, screen magnification and voice commands should be implemented in immersive technology to allow blind and partially-sighted users to access video materials. End users should be able to mark them in check-boxes. Users in the focus group consider that their preferences and parameters should be replicated between devices. Regarding accessing immersive content with or without glasses, users indicate that it depends on the end user.

Concerning the services

Users and audio describers consider that implementing access services to 360 immersive content is needed: if this technology is developed, access services should also be offered.

Binaural sound: users think that it could be useful to orient yourself in the scene. It is also needed to know where to turn your head to receive audio description. They consider that it deepens the sensation of being in the centre of the action. They expressed an opinion that object-based audio description could be very interesting, as it would deepen the experience of being inside the film.

Users favour the AD position linked to action being described. They are also interested in looking around and trigger secondary audio description tracks (e.g. to turn around and listen about what is behind them).

Returning to the main action: both end users and audio describers indicate the challenge of moving away from the main action and going back to it. They think that the information concerning the main action could be given (1) by the different voice, or (2) by a different volume of the sound.

Users (and audio describers) are strongly in favour of dubbing instead of voice over or subtitles when consuming foreign-language film.

Sensor with eye-tracker: end users think that this technology should respond when a user closes his or her eyes (or moves his or her head because the content provokes disgust or anxiety): there should be a sensor that detects it and audio description should be then stopped automatically.

Both professional audio describers and expert end users are more interested in using this technology in museums or planetariums (e.g. in form of simulations of travels or diving in an ocean) and not at home. The reasons behind it are the following: (1) watching films is mostly a social activity and spending time in glasses keeps us removed in our own worlds – it is impossible to share the experience of watching film with others (2) this technology is more adequate for watching short-duration video materials (e.g. a city tour) as it could be too cognitively tiring to watch full-length film (over-stimulation).

End users are interested in using this technology for educational materials (that are also meant to be entertaining) or for training purposes (they said that *it feels as if such a content is real* which is the potential of this technology). End users are interested in watching materials no longer than 5-15 minutes.

Physical effort: participants consider that they would not always like to turn their heads while consuming content, especially if they used this technology at home. They are much more interested in turning themselves around and turning their heads around while consuming 360 video contents in museums or other institutions.

3.1.4. Second iteration: AD in Catalonia

The pre-pilot was conducted on 20 April 2018. UAB was the partner responsible for the workshop. The whole report is available here:
https://drive.google.com/open?id=1lfADeYeQJH2P_fAJMbyYKZWi69w6MhRB

Figure 11 Focus group at UAB

Participants profile

- Number of advanced end users: 4
- Demographics for users.
    1) Sex: 3 “male”, 1 “female”, 0 “other”, 0 “prefer not to reply”.

2) Age: 33, 25, 40, 27
3) Main language of the participants: 1 “Spanish”, 3 “Catalan”.
4) Level of finished studies: 0 “no studies”, 0 “primary education”, 1 “secondary education”, 0 “further education”, 3 “university”.
5) “I define myself as a…” 1 “blind person”, 3 “low vision person”, 0 “deaf person”, 0 “hearing impaired person”, 0 “deaf-blind person”.
6) Age in which your disability began: 1 “From birth”, 2 “0-4”, 0 “5-12”, 1 “13-20”, 0 “21-40”, 0 “41-60”, 0 “More than 60”.
7) What technology do you use on a daily basis? You can select more than one. 3 “TV”, 0 “PC”, 4 “Laptop”, 4 “Mobile phone”, 2 “Tablet”.
8) Do you have any device to access virtual reality content? 0 “Yes”, 4 “No”, 0 “I don’t know or I don’t want to reply”.
9) Which of the following is your preferred device for watching online video content (ie., Youtube, Vimeo, Netflix, Amazon Prime, broadcast catch up service, etc.? You can select more than one: 0 “PC”, 2 “Laptop”, 1 “Smartphone”, 1 “Tablet”, 0 “I don’t watch online video content”, 0 “Others”
10) Which of the following do you use on your connected devices to access the above content? 1 “Magnification (ie. Zoomtext)”, 1 “Screen readers (ie, JAWS, VoiceOver TalkBack)”, 1 “both”, 1 “none”.
11) Which of the following controls would you like to use with your screenreader/magnification tool when watching content online? 3 “Browse content library”, 3 “Identify content”, 3 “Functions as play, stop, pause, forward, rewind”, 3 “Switch AD/AS on and off”.

Summary of participants’ profile

4 users took part in the focus group (3 partially-sighted persons, one blind person). They were 1 female and 3 males, with ages ranging 25-40. 3 participants had university studies, one secondary education. None of them reported having a device to access virtual reality content. Laptop (4) and mobile (4) phone was the most used technology by the participants on the daily basis, followed by TV (3) and tablet (2). For one participant, visual impairment started from birth, for two participants between 0-4 and for another one between 13-20. One participant reported using magnification to access content, another participant reported using screen reader, third one reported using both magnification and screen reader and the fourth one reported using none of the mentioned options. Three participants identified function ‘browse content library’ as needed to watch content online. Also, three participants out of four considered the following functions as needed: “identify content”, “functions as play, stop, pause, forward, rewind”, “switch AD/AS on and off”.

Workshop discussion: AD placed on the action (privilege of sound)

One participant noticed that sound is related to the action and that it moves with the scene; sometimes it is hearable from one side and sometimes from the other.

All participants agreed that there is a relation between the origin of the sound and the place where action takes place.

Two participants considered that in AD more details would be needed, as it is sometimes scarce.

One participant said that the effort of watching content with AD that does not contain many details would be largely related to the level of remaining sight. This participant added that following the plot in this version could be problematic for completely blind persons.

Other participant assessed as positive the fact that not everything was conveyed in AD, as it gives time to 'rest.'

Workshop discussion: AD anchored to head position (voice of God)

The fact that participants could listen to the sound at the same level in both headphones was assessed as less confusing.

One participant considered this type of sound as better than the previous one, as sound that comes from one direction or another distracts from the video.

Two participants said that AD was more detailed, which made plot more understandable. One participant described the AD as 'informative', as it gave much information.

One participant pointed that this AD version was more understandable and that it provided more details: not only about the action, but also more subjective details, such as facial expressions. This participant deemed describing facial expression important. For this participant some information was still missing, but AD script was assessed as better than in the previous version.

Workshop discussion: AD anchored to soundscape (first person, past tense)

Participants assessed the AD script in this video as more complete. As one participant said, "It is very stimulating. It makes you feel much more immersed and inside the story."

One participant assessed the sound as less intrusive than the sound in the first video (the first sound option distracted this participant from what the participant was seeing). This participant added that this sound helps focus more on the story.

One participant said that AD in the first person let him understand the meaning of the video more and feel more immersed. This participant could not see the connection between the different actions in the previous versions.

Thanks to this version, two participants realized that this story was about the girl, as in this version "the girl is in the centre of the story." One participant added that before it was not clear.

Workshop conclusions

For most of the users, the preferred options were "voice of God" (second option) and "AD anchored to soundscape" (third option). Some users think, tough, that if they were asked about preferences in relation to the immersive sound options and not just audio description, the preferred options are "AD placed on the action" (first option) and "AD anchored to soundscape" (third option).

Regarding the script

All users agree that audio description that is more detailed and in first person makes them feel more immersed and understand the story better. They also believe that the more details, the better, as long as AD is well integrated in the video. Some users agree that the intonation of the audio description must be neutral, and that the user should deduce the emotions from the video. Other user thinks that, when the AD is narrated in first person, then it would be interesting to consider using a more interpreted voice-over.

Regarding the sound

The type of sound and the script are much related. The users think that we must take advantage of the possibilities of using immersive sound. Most users think that immersive sound has to be implemented only in the video soundtrack, and not in the AD, which should be more neutral (like second and third options), because otherwise it is distracting. Some other user thinks that it would be interesting to use immersive sound also for the audio description, because it can be used like a guide to know where the user should look at. In general, the different types of audio description and sound will depend on the type of video and content. Regarding the usage of glasses, depending on the remaining sight of the users, some users think that it is interesting to use the glasses and others prefer not to use them.

3.1.5. Second iteration: AD in the UK

The pre-pilot was conducted on 08 May 2018. RNIB was the partner responsible for the workshop. The report is available here:
https://drive.google.com/open?id=1ZWg2QyNfROiapvD6u4aJ_uTYlZtGnUnm

Participants profile

- Number of advanced end users: 6
- Demographics for users.
    1) Sex: 4 “male”, 2 “female”, 0 “other”, 0 “prefer not to reply”.
    2) Age: 53, 51, 42, 37, 34, 48
    3) Main language of the participants: English
    4) Level of finished studies: 0 “no studies”, 0 “primary education”, 2 “secondary education”, 2 “further education”, 2 “university”.
    5) “I define myself as a…” 4 “blind person”, 2 “low vision person”, 0 “deaf person”, 0 “hearing impaired person”, 0 “deaf-blind person”.
    6) Age in which your disability began: 5 “From birth”, 0 “0-4”, 0 “5-12”, 0 “13-20”, 1 “21-40”, 0 “41-60”, 0 “More than 60”.
    7) What technology do you use on a daily basis? You can select more than one. 6 “TV”, 0 “PC”, 6 “Laptop”, 6 “Mobile phone”, 3 “Tablet”.
    8) Do you have any device to access virtual reality content? 0 “Yes”, 6 “No”, 0 “I don’t know or I don’t want to reply”.
    9) Which of the following is your preferred device for watching online video content (ie., Youtube, Vimeo, Netflix, Amazon Prime, broadcast catch up service, etc.? You can select more than one: 0 “PC”, 0 “Laptop”, 6 “Smartphone”, 3 “Tablet”, 0 “I don’t watch online video content”, 0 “Others”
    10) Which of the following do you use on your connected devices to access the above content? 2 “Magnification (ie. Zoomtext)”, 4 “Screen readers (ie, JAWS, VoiceOver TalkBack)”, 0 “both”, 0 “none”.
    11) Which of the following controls would you like to use with your screenreader/magnification tool when watching content online? 6 “Browse content library”, 6 “Identify content”, 6 “Functions as play, stop, pause, forward, rewind”, 6 “Switch AD/AS on and off”.

Summary of participants’ profile

6 users participated in the focus group. Two of these six were partially-sighted and four identified as blind. Two were female and four male with ages ranging between 25 - 55 years. Two participants attended university, two were educated to secondary level and two had

undertaken further education. None of them reported having a device to access virtual reality content. Laptop, mobile phone, TV and Tablet were the most frequently used devices by the participants. Whilst all said that they used TV, phone and laptop on a daily basis, only three said that they used tablet as frequently as the rest of their other devices. With the exception of two participants, all others have had sight loss from birth. Two participants reported using magnification to access content and the rest four were screen reader users. All participants needed the independent access to following functions in order to watch content online: "browse content library", "identify content", "functions as play, stop, pause, forward, rewind", "switch AD/AS on and off".

Workshop discussion

AD placed on the action (privilege of sound)

4 participants agreed that there was a relation between the origin of the sound and the place where action takes place.

One participant asked if the directionality in sound was anything like watching a 3D film visually which indicates that directionality lent depth to the scene.

Three participants commented that the orientation helped put sounds in context with the additional audio description. i.e., Parking Lot.

One participant suggested that an audio introduction be introduced for further clarity of the characters and what was happening in the video – a preamble of sorts. It must be noted that this participant is a regular theatre goer and therefore uses AI frequently prior to the performances.

Observation: one participant (totally blind) seemed to be following the changes in audio placement by moving his head to track the action

AD anchored to head position (voice of God)

None of the participants chose to comment a great deal on this option. All agreed that this was the traditional description which was delivered on film and TV therefore it the usual format.

3 participants after listening to this option said that they did not want too much description as they wanted to listen to music. They did not wish to be overwhelmed - too much description over a music video meant you can't hear the music anymore. Other disagreed stating that if it was traditional AD that was being discussed then it depended on the context - if film is dubbed description is essential (Passion of the Christ - Hebrew), if every facial movement is described then it must be significant, essential to the context. If it is a Music Video - perhaps watch once with AD then again without.

4 participants agreed that the mix was good as it did not drown any sound effects.

All participants agreed that there was nothing special on this description – it was 'sufficient', it was the one they listened to TV and film all the time, it did not feel immersive but it was alright. It was like watching any other content on TV and cinema. It must be noted here that all our participants are regular users of description, so much so that most agreed that they now refuse to waste time on content that is not delivered AD.

AD anchored to soundscape (first person, past tense)

This was clearly the preferred choice - but all participants agreed that it wouldn't work for all genres, for example, soaps. However for the content used in this test, it worked very well. It felt like part of the video.

Participants agreed that the AD script worked, the story was immediately clear. They were more aware of what was going on in the story.

One participant commented on the voice of the narrator (audio describer) and delivery – it matched the spirit of the character in the video – which made it even more immersive – felt as if the narrator was part of the video.

One participant said that she felt quite emotional watching the video with this description which didn't occur while watching the other options.

Three participants realised after watching this version that the central character in this video is the girl. One participant added that before it was not clear.

Workshop conclusions: approved by focus group

For most of the users, the preferred options were "AD placed on the action (privilege of sound)" and "AD anchored to soundscape". It must be noted here that all participants were experienced users of audio description and are used to using the additional soundtrack across TV, film, theatre and other media. As 360 degree/ VR was described to the group as a gamified – slightly gimmicky experience, participants expected the AD to be appropriately striking for that experience. Stand alone, Voice of God was considered sufficient and at par with what is usually delivered on TV and in cinemas but when offered in the context of a 360 degree/ VR experience, the group wanted to feel the immersion that sighted viewers do visually.

One participant suggested a combination of this script with privilege of sound i.e., the description changes as one focused on different things. So the audio has different streams and movement of the head triggered different streams. Other participants were doubtful that it would work.

Script, voice, intonation

Majority of the participants did not want too much description but a good balance so they could enjoy the music but also understand the story – the key elements. In their opinion, the first person narration worked for several reasons including – the narration working as a tour guide - taking viewers through the experience but also, the voice chosen for the description, the intonation synced well with the actions described. Overall it felt as part of the experience leading to greater level of engagement. Viewers felt more emotionally invested in the story.

Spatial audio and first person narrative

Participants once again brought up the barber shop experience and the level of 'involvement' that it triggered. 360 degree or VR must feel different and not like a usual 2D presentation – so it isn't only about the level of understanding or clarity of what is happening in the video but also involvement in the video i.e., level of familiarity with the characters and significant events – feeling part of the story.

Orientation is a crucial part of 'involvement', knowing where the action is happening – are you in the middle of it or on the side and if it is the side, which side? Traditional 360 experiences for sighted audiences rely heavily on viewers responding to sounds as it generates higher level of immersion as one if involved in the video. Therefore the same rules apply to people with sight loss which may be achieved by spatial audio.

HMD

None of the participants wanted to wear HMD, most found it unnecessary and tiring.

3.2. Subtitling and Sign Language

Two focus group on subtitling and sign language were performed as part of the first iteration. They correspond to the two broadcasters CCMA and RBB who will later perform the pilots. In the second iteration different actions took place at UAB, RBB and CCMA, with diverging user profiles and services involved.

3.2.1. First iteration: subtitling and sign language in Germany

The focus group took place at RBB premises on 28.11.2017 and subtitling and sign language interpreting were discussed. The report is available here:

https://docs.google.com/document/d/17086euRxv2xPa0ujKpvRRoQODIVI7RsstZWHmaHjSQI/edit

Participants profile

- Number of professional users and profile: 4 (2 subtitle editors, 2 experts for 360° videos).
- Number of home users: 5
- Demographics for home users:
  1) Sex: 2 “male”, 3 “female”, 0 “other”, 0 “prefer not to reply”.
  2) Age: 37, 2x40, 52, 62
  3) Main language of the participants: 4 “German sign language”, 1 German
  4) Level of finished studies: we omitted this question
  5) “I define myself as a…”: 0 “blind person”, 0 “low vision person”, 2 “deaf person”, 3 “hearing impaired person”, 0 “deaf-blind person”.
  5.1 level of hearing impairment according to WHO: 5 “profound hearing loss (over 81 db)”
  6) Age in which your disability began: 1 “From birth”, 3 “0-4”, 1 “5-12”, 0 “13-20”, 0 “21-40”, 0 “41-60”, 0 “More than 60”.
  7) What technology do you use on a daily basis? You can select more than one. 5 “TV”, 1 “PC”, 5 “Laptop”, 5 “Mobile phone”, 2 “Tablet”.
  8) Do you have any device to access virtual reality content? 0 “Yes”, 5 “No”, = “I don’t know or I don’t want to reply”.

Summary of participants’ profile

Nine users took part in the focus group (5 home users, 2 subtitle editors and 1 expert for 360° content). They were 4 males and 5 females. The home user ages are ranging from 37 to 62. None of them reported having a device to access virtual reality content. Laptop, TV and mobile were the most-used technology by the participants on a daily basis (5), followed by tablet (2) and PC (1). The home users were hearing impaired participants. According to the WHO grades of hearing impairment, all participants have a profound hearing loss (hearing impairment started from birth to 5-12).

Conclusions from professional users

Conclusions are related to either subtitles or the signer as follow:

Subtitles

In the separate session with the professional users we jointly elaborated the subtitle production workflow and then defined any necessary enhancements for positioning in 360° content.

The current subtitle workflow depicts the necessary workflow steps for the production part.

1) Content is available (video, manuscripts and/or subtitles)
2) Import/open content
3) In the subtitle software
    - Check integrity
    - Format text in “frames”
    - Define time code base
    - Set time code for frames
    - Check
    - Replay (high quality preview, less authoring)
    - Acceptance
4) Export/save/send

The workflow has to be enhanced with additional tools to enable the following additional features:

- The existence of additional spatial information doubles the effort required to place subtitles within media items
- Time-based PLUS angle-based navigation with the help of shortcuts, scroll wheel and input field
- Time-based PLUS angle-based subtitle definition
- Time-based PLUS angle-based editing
- Time-based PLUS angle-based preview
- Time-based PLUS angle-based replay
- Preview in low-res flat angle view
- Replay and acceptance in high-res flat angle, flat unfolded and HMD view

Signer

Here, the workflow is obviously quite similar to that of the production of subtitles. Sign language video is pre-produced and needs to be added to the omnidirectional main video, depending on positions in the 360° field:

- Content is available (signer video)
- Import/open content
- Cut signer video if necessary
- Time-based PLUS angle-based signer definition
- Name, information of the speaker
- Time-based PLUS angle-based editing and positioning of textual and/or graphical notices
- Time-based PLUS angle-based preview
- Time-based PLUS angle-based replay
- Preview in low-res flat angle view
- Replay and acceptance in high-res flat angle, flat unfolded and HMD view

Conclusions from advanced home users

Conclusions have been gathered in two groups as follows:

Figure 12 Explaining immersive subtitles in RBB

Concerning the interaction: accessing the services

They would like to have a similar approach to the one current used in RBB's web player for video on demand and the TV player used for catch-up services. These offer a dedicated button with the German abbreviation "UT" (for subtitles) allowing the user to switch the subtitles on or off and to access the settings. The adjustment wheel icon for accessing all available settings is also very common and was mentioned. Summarized they would like to see a solution based on existing settings similar to those used for the catch-up TV service (position, background, size). The usage of the user interface in a HMD was identified as a possible challenge.

Concerning the services

The users saw a 360° video of RBB's news magazine Abendschau showing a "behind the scenes" story of the show's production. The video was shown in a desktop-based web browser and in a HMD (Oculus rift, VR glasses with smartphone).

One issue that was not obvious before the focus group discussion is that people with hearing loss have balance disorders and the usage of a HMD can potentially cause motion sickness. The users were sitting during the service consumption and our observation was that the advanced home users only made tentative head movements.

One scene from this video was illustrated with sketches on the whiteboard. The blue frame symbolised the users' field of view.

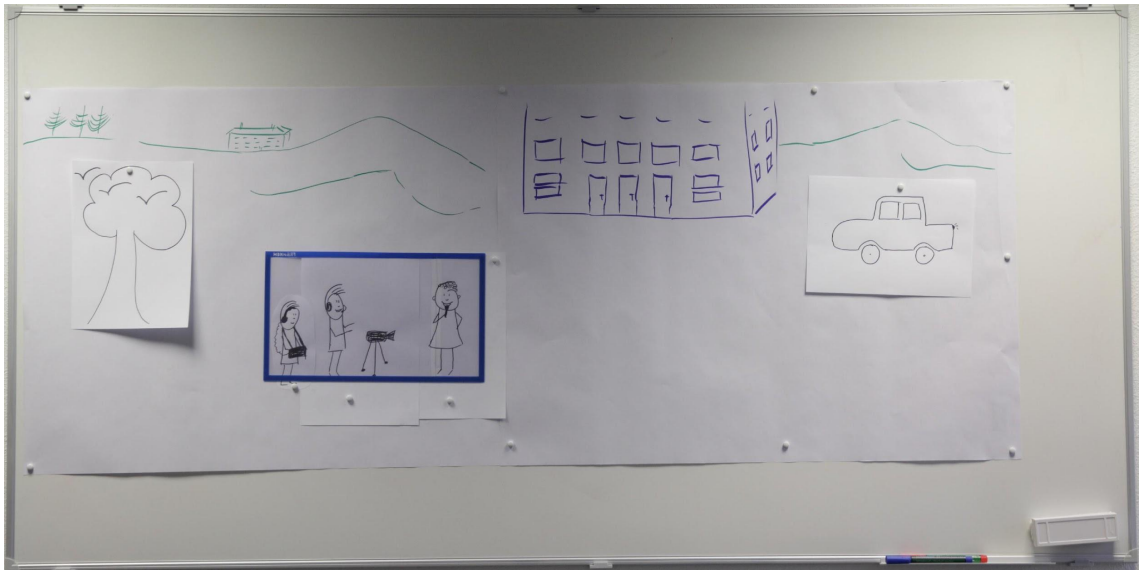

Figure 13 Basic scene

1st Use case: off-speaker, subtitles

The speech bubbles symbolised that somebody is speaking. In the first case, illustrated in Figure 14, an off-screen speaker explains something and the user is looking around in the content. The question was: Where should the subtitles been positioned? The red stripe symbolised the subtitles.

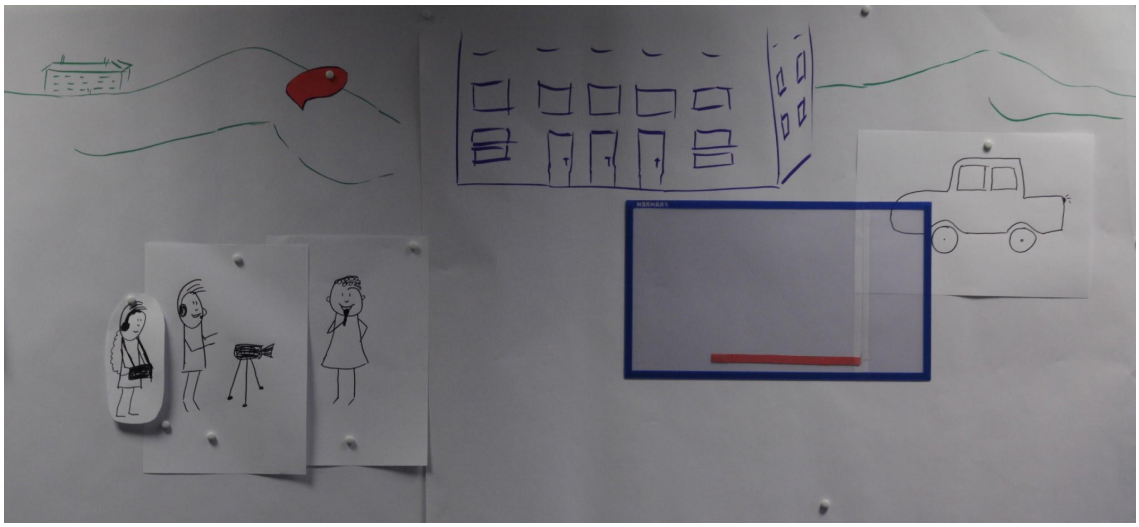

Figure 14 Off-screen speaker

Conclusion

The subtitles should be in a fixed position in the user's field of view according to the current standards, two-lined and each speaker gets its own colour. It is important for the users to get the same information as hearing people.

2nd Use case – one speaker, subtitles

In this example a presenter is speaking and different acoustic information such as singing birds in the tree, a car door slamming and a camera team preparing for a live recording is visible and audible.

The questions to the users were: Do you want to have a description of this audible background noises in addition to the subtitles for the speaker? If yes, how should they be displayed?

Figure 15 One speaker

Conclusion

In contrast to the above statement that all information i.e. speech and background noises should be available for people with hearing loss, one user was not sure if these noises are important enough to be mentioned.

One of our RBB team members asked again concretely: How should a user get the information that something important was happening outside of the field of view?

The user maintained his opinion and wanted to explore the content by himself. At the end of the discussion all agreed that it could be important in fictional films or documentaries to inform the user about background noises for a better understanding of the dramaturgy. The manuscript authors or subtitle editors responsible should decide which information is important for the advanced home user.

The conclusion was that information about background noises should be clearly distinguishable from subtitles. The users didn't express a clear preference if this information should be textual or graphical. They would like to test both uses cases in a pilot and also have an option to switch the information on or off.

In general the 360° content consumed with a HMD is more difficult for people with hearing loss as they don't get any acoustic feedback if something dramatic happens. An example was an approaching herd of horses. They won't hear them and will be surprised when they suddenly appear in their field of view. Vibration feedback is one of the familiar feedback mechanisms. Testers suggested delivering such feedback via a combination of head-mounted displays and controllers.

3rd Use case – two speakers, subtitles

In the 3rd use case one presenter is speaking and appears in the field of view. He stops talking and goes into the building. Simultaneously another presenter starts to speak outside of the field of view.

The question to the users was: How does the user know who to attribute the new subtitles to and where the presenter is in the 360° space?

Conclusion

The first idea of the users was to add an arrow to the subtitles to indicate the position of the speaker, depending on the actual viewing angle. The arrow should disappear as soon as the user reaches the respective field of view, and speaker and subtitles aligned. One of the subtitle editors proposed an alternative solution. The positioning of subtitles on the left or right edge of the current field of view indicates the direction in which the speaker is talking and the subtitles will go to the middle as soon as the user reaches the respective field of view.

4th use case – two speakers, signer

The 4th use case has the same starting situation as the 3rd use case. The difference is the type of access service.

The questions to the users were: where should the signer be positioned? How does the user know where the second presenter is?

Conclusion

The signer should be positioned as usual top right of the speaker or if no speaker is visible top right of the field of view. The signer should always be visible in the field of view. The users see difficulties if two speakers are talking in parallel, especially when both or only one of them is in the field of view. Usually, the signer indicates with a gesture the direction of the speaker. This can be done in a studio situation for 2D video. In 360° video content the signer does not know where the speakers are positioned in relation to the user's field of view.

Different approaches were discussed:

1) An arrow indicates the position of the speaker directly under the signer window and/or the name or description of the speaker is displayed.
2) The field of view will be changed by the video player ("forced perspective") so that the user will see the speaker when a conversation between two persons starts and can align the translation of the signer to the speaker. Afterwards the user decides the field of view self-sufficient.
3) The signer image has a fixed location and the user decides where he/she wants to look.

Figure 16 Focus group – RBB

Additional questions:

Personalisation options

1) Subtitles - the same settings as in catch-up service:
   - switch on or off subtitles
   - position (top, bottom)
   - size (three different: small, medium and large)
   - background (semi-transparent box, outline, frame around TV picture)
   - switch on or off notices
2) Signer:
   - position
   - size
   - switch on or of notices
   - switch on or off forced perspective on the speaker

Parallel usage in a group of people – synchronisation between devices

- the parallel usage in a group is a scenario that is conceivable but more in front of the TV
- the tablet as a second device was not preferred as the display is too small for the consumption of subtitles and signer
- the head mounted displays were not preferred – they prevent the communication in sign language and a usage alone was more preferred
- a HoloLens approach was discussed, where users can see each other and at the same time have a sign language translator positioned next to the TV-based end device.

3.2.2. First iteration: subtitling and sign language in Catalonia

The whole document with notes from the focus group can be found here: https://drive.google.com/file/d/1mVo3wkJ8TSVS6cougZB5zopJdrCeBI1s/view?usp=sharing

A press release of the Focus Group can be consulted here: http://www.ccma.cat/premsa/la-ccma-participa-en-el-projecte-imac/nota-de-premsa/2824511/ And tweets to the press release:
https://twitter.com/CCMA_cat/status/936624649286438912 (English)
https://twitter.com/CCMA_cat/status/936617166408953856 (Catalan)

Figure 17 Participants in CCMA focus group

Figure 18 Participants in CCMA focus group and Sign Language interpreter

The focus group took place at CCMA (Barcelona) on 28.11.17 and the access services discussed were subtitling and sign language interpreting.

Participants profile

- Number of home users: 10 (4 Sign Language users, 6 oralists).
- Number of professional users and profile: 6 (2 subtitlers, 3 technical experts, 1 user association member –not deaf–).
- Demographics for users:
    1) Sex: 8 “male”, 8 “female”, 0 “other”, 0 “prefer not to reply”.

2) Age: 25, 34, 38, 38, 41, 46, 47, 47, 49, 53, 53, 58, 61, 62, 65, 66.
3) Main language of the participants: 1 Spanish, 1 Sp/Cat, 10 Catalan, 3 Catalan Sign Language, 1 Spanish Sign Language.
4) Level of finished studies: 0 “no studies”, 0 “primary education”, 3 “secondary education”, 4 “further education”, 8 “university”. One person did not answer.
5) (Only for home users): “I define myself as a…” 0 “blind person”, 0 “low vision person”, 8 “deaf person”, 2 “hearing impaired person”, 0 “deaf-blind person”.
6) (Only for home users): Age in which your disability began: 4 “From birth”, 5 “0-4”, 0 “5-12”, 0 “13-20”, 0 “21-40”, 1 “41-60”, 0 “More than 60”.
7) What technology do you use on a daily basis? You can select more than one. 14 “TV”, 10 “PC”, 12 “Laptop”, 16 “Mobile phone”, 8 “Tablet”.
8) Do you have any device to access virtual reality content? 3 “Yes” (VCR?, Glasses, PC), 13 “No”, = “I don’t know or I don’t want to reply”.
9) Which of the following is your preferred device for watching online video content (i.e., Youtube, Vimeo, Netflix, Amazon Prime, broadcast catch up service etc.)? You can select more than one. 7 “PC”, 7 “Laptop”, 5 “Smartphone”, 3 “Tablet”, 0 “I don’t watch online video content”, 3 “Others” (TV)

Summary of participants’ profile

Sixteen users took part in the focus group (10 home users –4 Sign Language users, 6 oralists–, 2 subtitlers, 3 technical experts and 1 user association member –not deaf–). They were 8 males and 8 females, four participants with ages ranging 21-40, eight participants with ages ranging 41-60 and finally four participants with more than 60. Three participants had secondary education studies, four participants had further education studies, eight had university studies and one person did not reply to this question. Three of them reported having a device to access virtual reality content (VCR, glasses and PC, respectively). Mobile phone was the most used technology by the participants on a daily basis (16), followed by TV (14), laptop (12), PC (10) and tablet (8). The home users were deaf (8) and hearing impaired people (2), most of them having acquired the disability from birth (4) or with ages ranging 0-4 (5), and one person with ages ranging 41-60. The preferred devices for watching online video content was PC (7) and laptop (7), followed by smartphone (5), tablet (3) and TV (3).

Workshop conclusions: approved by focus group

A first session was carried out with all the focus group participants. The excellent turnout of home users in the focus group resulted in a slower development. It was therefore agreed to focus on questions of interest for both home and professional users, and leave the specific questions addressed to professional users for a second meeting only for professional users. This is why the information is reported in the changed order. First, we report about the results of the first focus group, in which both home and professional users took part, and then we report on conclusions from the professional users.

Conclusions from advanced home users

Conclusions have been drawn in three large groups as follows:

Concerning the interaction: accessing the services

Regarding interaction with the interface, both professional and advanced home users consider that it would be positive to be able to customize the interface once and then that the interface would remember those parameters.

Users suggest that this customization should be transferred from one device to the other (importing profile) and they request the possibility to create more than one profile. They also

consider the possibility of transferring a profile from your device to another external device (for example, at your friend's home).

Regarding interaction with access services, users positively value the possibility of alternative interactions (for example, voice commands), although they do not find it necessary in their case and they indicate that the implementation costs should be taken into account. However, they think that, if this is to be developed for other profile types anyway, it could be an additional resource.

Regarding companion screens, users like the possibility of using the smartphone to move the screen in a tactile way (like a "mouse") and to customize their preferences. A user even suggested the possibility to include a sensor in the finger that allows users to see their own fingers on the image. There are different opinions regarding the need of reproducing the content in the mobile too, since the smartphone is often used as an element to access additional contents.

When accessing audiovisual content with other people, users do not want the subtitles to be consumed on a different screen (for example, a smartphone), but on the same screen as the other users.

Concerning the services

Users suggest that subtitles should always appear in a fixed position in relation to the user's field of view.

Users suggest that subtitles in immersive media should be based on approved subtitling rules (for example, UNE rule in Spain) and, if necessary, improvements might be implemented to adapt existing rules to the new needs posed by immersive environments.

Users state that it is necessary to keep colour coding to identify characters.

Users require that basic subtitling elements that have been already approved in the regulations (for example, how to indicate music) need to be kept. However, they accept that new technologies may bring new possibilities.

Users require that subtitles include all the information, both the information present on the screen and also the information off the screen (ON and OFF).

Users suggest that it is necessary to include non-speech information (that is, sounds, extra linguistic information, etc.) and directions.

Users state that it is difficult to know where to look for the character who is speaking. The subtitle must indicate where you need to move your head (four directions). It is suggested that a wind rose or a compass is drawn to indicate where the sounds come from.

It seems that users prefer that icons or similar elements are always located in the same position. Some users prefer at the top, others at the bottom closed to the subtitle (dialogue), and others would like to move them. All in all, it seems that customization is the solution.

Users suggest the possibility to use a closed list of icons in order to illustrate non-speech information. For example, a lightning to indicate the sound of a storm.

There are different opinions regarding how to include non-speech information in the subtitles: with icons or text. In that sense, customization should be prioritized.

Users positively value the possibility of personalization, that is, having different layers that one can activate or not, depending on their needs. Like a "menu of options". However, there are some elements that do not need customization, such as the position of the subtitles – always at the bottom – except from specific cases, such as football matches, etc.

It becomes clear that there are different needs among users and, consequently, subtitles must be adapted to different profiles. For example, there could be different levels of speed (faster/slower).

Additionally, sign language must also be considered. Sign language would appear simultaneously to the person speaking on the screen and it would satisfy the needs from other users.

Users consider that summarized or simplified subtitles that do not reproduce exactly what is being said word by word do not generally help deaf people, because this type of subtitles make it more difficult to follow the audiovisual content. However, they admit that simplified subtitles may be useful for users with other type of needs (for example, people who need easy-to-read texts). In that sense, simplified/summarized subtitles can be an alternative.

Regarding sign language, users require that the sign language interpreter is always located at the same fixed position in relation to the user's field of view and with a background. They also prefer that each user has the possibility to customize the position of the sign language interpreter.

Users raise doubts regarding the success of these new technologies, and they believe that we need to be ready, but they are afraid that it happens the same as happened with other services, such as 3D cinema.

Conclusions from professional users

These are the answers or proposals expressed by professional subtitling producers: Vertical positioning of subtitles is not a main need in 360 subtitles, but it must be interesting to use different vertical positioning to separate different subtitles or non-speech information. Anyway home users must be able to decide or setup where they want to have the subtitle text.

As agreed with end users, there is not demand for supporting 'moving positions' (e.g. a subtitle following a person in 360º), and agree with expert users about the need to keep a subtitle fixed position in relation to the user's field of view.

During the production of subtitles, producers prefer to have an on-screen display (player) showing one dynamic angle of the 360 view, so they can choose which angle to see through cursors or mouse movements. They don't need to have a 2D distorted panoramic view showing full 360º.

For testing purposes of contents, producers think it should be done with HMD and on-screen display, so they can test both results, using HMD or directly from a display (PC screen or smartphone).

A 360 web subtitling editor could be very similar to, for example, the Anglatècnic tool they use nowadays, but it should add the 360 displaying and the possibility to add 'emoticons' and text messages to show 'sound actions' that take place in parallel to dialogue subtitles.

The Web Editor tool must offer original 360 immersive audio, because it is important to work correctly with subtitles and notifications showing where the sound comes from. The tool would add a wind rose item that will help the home user to locate the position of different sounds or dialogues.

About the possibility to have a Web Editor tool for Sign Language, it is observed that Sign Language would be produced in a similar way as it is done nowadays for classic 2D audiovisual

contents. The position for the Sign Language Picture in Picture box would be configured by the home user from player interface, and would be a fixed position in the visual area.

3.2.3. Second iteration: subtitling in Catalonia with hearing users

Pre-pilot tests took place on 10/05/2018 and 19/05/2018 led by UAB. The whole report, with user comments, is available here:

https://drive.google.com/open?id=10bH84IVORmgvVHc6MDdpbCHJThiXh_Vf

Participants profile

- Number of end users: *6.*
- Demographics for users.
    1) Age: 14, 16, 27, 28, 53, 58.
    2) Level of finished studies: 0 “no studies”, 2 “primary education”, 1 “secondary education”, 1 “further education”, 2 “university”.
    3) What technology do you use on a daily basis? You can select more than one. 5 “TV”, 2 “PC”, 2 “Laptop”, 6 “Mobile phone”, 2 “Tablet”, 0 “HMD”, 1 “Game console”, 0 “Other”.
    4) How often do you watch virtual reality content (for instance, 360º videos)?

	Never	Occasionally	At least once a month	At least once a week	Every day
In smartphone	4	2			
On a tablet	6				
On a PC	6				
In smartphone plugged to HMD	5	1			
In HMD	6				

Table 3: Results on usage of virtual reality content (UAB)

5) If you have never used virtual reality content such as 360º videos or only occasionally, please indicate why. Multiple answers are possible. 1 “Because I am not interested.”, 1 “Because it is not accessible.”, 4 “Because I have not had the chance to use it.”, 0 “Other reasons.”
6) Please state your level of agreement with the following statement: “I am interested in virtual reality content (such as 360º videos).” 1 “I strongly agree”, 2 “I agree”, 3 “Neither agree nor disagree”, 0 “Disagree”, 0 “Strongly disagree”

7) Do you have any device to access virtual reality content? 2 “Yes” (*and please indicate which ones they name: Google Cardboard and smartphone*), 4 “No”, 0 “I don’t know or I don’t want to reply”.
8) Do you like watching the following types of content on television or online?

	I like it very much	I like it	Neither like it nor dislike it	I don't like it	I don't like it at all
News		5	1		
Fiction (series, films)	4	2			
Talk shows	2	1	2	1	
Documentaries	1	4		1	
Sports	1	1		1	3
Cartoons			6		

Table 4: Preferences according to genre (UAB)

9) When subtitling is available, do you activate it for the following type of content?

	Always	Sometimes	Rarely	Never
News			1	5
Fiction (series, films)	1	1		4
Talk shows			1	5
Documentaries	1			5
Sports				6
Cartoons	1	1		4

Table 5: Subtitling usage according to content (UAB)

10) If it is available and you do not activate it, please select the reasons why. 0 "Because the interface is not accessible.", 2 "Because I don't want subtitling in all the content, only in certain types of content.", 4 "For other reasons. Please explain why: Because I want to listen to the contents in Spanish or Valencian; Because I don't need them, I watch contents in Spanish; I don't have enough time to read them; Because I understand the source language."
11) How many hours a day do you watch subtitled content? 3 "None", 2 "Less than 1 hour", 1 "1-2 hours", 0 "2-3 hours", 0 "3-4 hours", 0 "4 hours or more"
12) What do you use subtitles for? 1 "They help me understand", 0 "They are my only way to have access to the dialogue", 2 "I use them for language learning", 5 "Other. Please explain: I don't use subtitles; I don't use subtitles because I watch contents in Spanish; I don't use subtitles; They help me understand in case the language spoken is not Spanish/Catalan; To help me understand content in other languages."

Results

The tables below summarise the results following the template report in Annex XI. The first row indicates the feature put to test, namely six levels of comfort field of view, and three guiding mechanism (position, arrow, and compass). For the comfort field of view a second row defining each of the levels is added. The first column refers to the participant number (P1= participant 1). The values given by each participant correspond to the methodology described in Annex VIII: each participant had to write on a 1 to 5 scale their assessment of each comfort viewing field as well as their assessment of the three implemented systems to guide the user to the speaker location.

Comfort	**Level 1**	**2**	**3**	**4**	**5**	**6**
Field of view Font size in pixel	**30% 48**	**40% 42**	**50% 36**	**60% 30**	**70% 24**	**80% 18**
P1	3	3	2	1	1	2
P2	2	2	3	3	4	4
P3	1	1	3	3	3	3
P4	1	2	3	3	4	1
P5	1	1	2	3	3	4
P6	1	1	2	3	4	5
Mean	1.5	1.7	2.5	2.7	3.2	3,2

Table 6: Results ST: comfort field of view (UAB)

Guiding	**Position**	**Arrow**	**Compass**
P1	1	4	2
P2	1	5	3
P3	1	3	4
P4	1	3	3
P5	1	2	3
P6	1	4	1
Mean	1	3,5	2,7

Table 7: Results ST: guiding mechanism (UAB)

The whole document with notes and comments from users can be found here: https://drive.google.com/open?id=10bH84IVORmgvVHc6MDdpbCHJThiXh_Vf The aim of the test was to obtain qualitative data through a reduced sample of users, therefore statistical tests were not performed. Qualitative feedback from users regarding preferences is summarised next.

Conclusions

Comfort field of view (video 1)

The result of this test was: level 5 and level 6 got the highest rating, followed by level 4 and level 3. Level 1 and 2 were too small and difficult to read.

Some relevant considerations:

- Users generally preferred subtitles at the bottom of the field of view, since they claimed to be used to it. They did not like when subtitles covered part of the image, it was annoying.

- For most users, the bigger font was easier to read, but we need to make sure that the subtitles do not cover the image and that they are not too far from the image (depth issues). Personalisation for font size will be required.
- Some users reported double vision, due to the implementation of the subtitles (too close to the eyes, need to be closer to the image). Implementation needs to be improved.
- A user also reported colour blindness issues. Personalisation for font colour should be implemented.

Guiding to the speaker (video 2)
According to the average rating of the users the following preferences of the different approaches apply: 1) arrow 2) compass and 3) sided text.

Some relevant considerations:
- Vertical axis (up and down) is missing from all options, need to be implemented.
- We need to further test how to implement directions when two speakers (or more) are talking at the same time. This was confusing for the users.
- Most users preferred the arrow, but they want it to be clearer: bigger and maybe with a colour different from the subtitle.
- Two users claimed that all systems were not clear and confusing. One user suggested to use, apart from the arrows, an indicator close to the speaker (for example, a red dot) to be sure about who is talking at each time.

3.2.4. Second iteration: subtitling in Catalonia

Pre-pilot testing took place at CCMA premises between May 3rd-14th, and subtitling was the service discussed with user. The report is available here: https://drive.google.com/open?id=1vWCP0pfiMtRsAW0_KBby9fT_RDCYHnGl

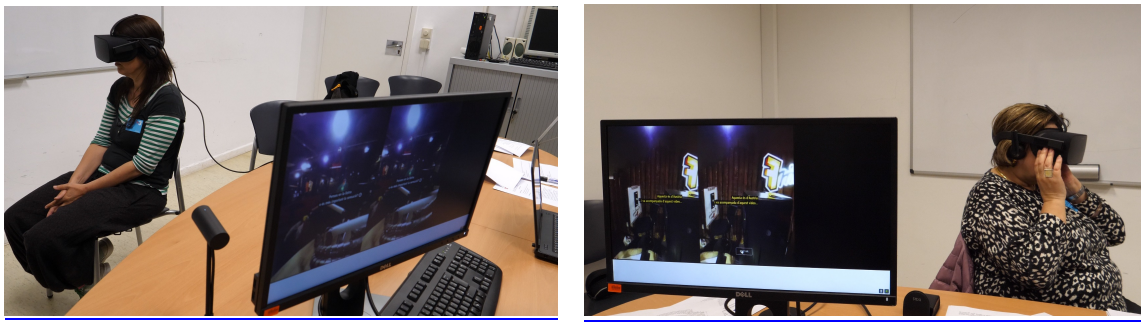

Figure 19 Participants in CCMA pre-pilot

Participants profile
- Number of end users: 5.
- Demographics for users

1.	Sex	0 male 5 female 0 other 0 prefer not to reply
2.	Age	46 – 50 – 56 – 62 – 63 years old

3.	Main language of the participants:	5 Catalan
4.	Level of studies	0 No studies 0 Primary education 1 Secondary education 1 Further education 3 University
5.	I define myself as a…	0 blind person 0 low vision person 2 deaf person 0 hearing impaired person 0 deaf-blind person 3 no hearing impaired – no low vision
6.	Age in which your disability began	1 From birth 1 0-4 0 5-12 0 13-20 0 21-40 0 41-60 0 More than 60
7.	What technology do you use on a daily basis? You can select more than one.	5 TV 2 PC 3 Laptop 3 Mobile Phone 3 Tablet
8.	Do you have any device to access virtual reality content?	5 No 0 I don't know or I don't want to reply.
9.	Which of the following is your preferred device for watching online video content (i.e., Youtube, Vimeo, Netflix, Amazon Prime, broadcast catch up service etc.)?	2 PC 2 Laptop 2 Mobile Phone 2 Tablet 0 I don't watch online video content 0 Others

Table 8: Demographics for users (CCMA)

Results

The tables below summarise the results following the template report in Annex XI. The first row indicates the feature put to test, namely six levels of comfort field of view, and three guiding mechanism (position, arrow, and compass). For the comfort field of view a second row defining each of the levels is added. The first column refers to the participant number (P1= participant 1). The values given by each participant correspond to the methodology described in Annex VIII: each participant had to write on a 1 to 5 scale their assessment of each comfort viewing field as well as their assessment of the three implemented systems to guide the user to the speaker location.

Comfort	**Level 1**	**2**	**3**	**4**	**5**	**6**
Field of view **Font size in pixel**	**30%** **48**	**40%** **42**	**50%** **36**	**60%** **30**	**70%** **24**	**80%** **18**

P1	4	3	5	5	4	3
P2	3	3	5	5	3	5
P3	3	3	4	3	4	2
P4	3	3	3	4	3	2
P5	2	3	3	5	4	3
Mean	3	3	4	4.4	3.6	3

Table 9: Results ST: comfort field of view (CCMA)

Guiding	Position	Arrow	Compass
P1	X	2	4
P2	X	4	3
P3	1	4	2
P4	3	5	5
P5	2	2	4
Mean	2	3.4	3.6

Table 10: Results ST: guiding mechanism (CCMA)

The aim of the action was to obtain qualitative data through a reduced sample of users, therefore statistical tests were not performed. Qualitative feedback from users is summarised next.

Conclusions

Comfort field of view (video 1)

The result of this test was very clear: level 4 got the highest rating, followed by level 3 and level 5.

The following improvements were requested by the testers:

- Users generally preferred subtitles at the bottom of the field of view. They did not like when subtitles covered part of the image (this is crucial for lip-reading).
- One user expressed disagreement regarding the space between two subtitle lines, and insisted that for two lines subtitling the upper and lower line should go together, in a compact black block.

Guiding to the speaker (video 2)

According to the average rating of the users the following preferences of the different approaches apply: 1) compass 2) arrow and 3) sided text.

The following improvements were requested by the testers:

- Icons should be improved and should have a different colour to differentiate them from the text.
- The icons should be close to the subtitle but outside, not integrated in the lines of the subtitle. Direction indicators should be independent from the subtitle.

3.2.5. Second iteration: subtitling in Germany

This pre-pilot action took place at RBB between 11 and 13 April 2018, and the service discussed was subtitling. A whole report is available here:

https://docs.google.com/document/d/1PagQb9xk8rdoLypfXupCGJN7oeV7w5C6yYsxUzcUjVg/edit

Participants profile

- Number of end users: *5.*
- Demographics for users

1.	Sex	2 male 3 female 0 other 0 prefer not to reply
2.	Age	40 – 60 years old
3.	Main language of the participants:	German German sign language
4.	I define myself as a…"	0 blind person 1 low vision person 3 deaf person 2 hearing impaired person 0 deaf-blind person
5.	Age in which your disability began	3 From birth 1 0-4 1 5-12 0 13-20 0 21-40 0 41-60 0 More than 60
6.	What technology do you use on a daily basis? You can select more than one.	5 TV 3 PC 4 Laptop 5 Mobile Phone 2 Tablet
7.	Do you have any device to access virtual reality content?	5 No 0 I don't know or I don't want to reply.

Table 11: Demographics for users (RBB)

Results

The tables below summarise the results following the template report in Annex XI. The first row indicates the feature put to test, namely six levels of comfort field of view, and three guiding mechanism (position, arrow, and compass). For the comfort field of view a second row defining each of the levels is added. The first column refers to the participant number (P1= participant 1). The values given by each participant correspond to the methodology described in Annex VIII: each participant had to write on a 1 to 5 scale their assessment of each comfort viewing field as well as their assessment of the three implemented systems to guide the user to the speaker location.

Comfort	**Level 1**	**2**	**3**	**4**	**5**	**6**
Field of view Font size in pixel	**30% 48**	**40% 42**	**50% 36**	**60% 30**	**70% 24**	**80% 18**
P1	2	4	4	5	5	4
P2	1	2	5	4	5	3
P3	4	6	5	5	5	4
P4	1	4	5	5	4	4
P5	5	3	4	4	3	2
Mean	2.6	3.6	4.6	4.6	4.4	3.4

Table 12: Results ST: comfort field of view (RBB)

Guiding	**Position**	**Arrow**	**Compass**
P1	1	1	1
P2	4	5	2
P3	1	5	5
P4	1	4	2
P5	1	2	5
Mean	1.6	4	3

Table 13: Results ST: guiding mechanism (RBB)

The aim of the action was to obtain qualitative data through a reduced sample of users, therefore statistical tests were not performed. Qualitative feedback from users is summarised next.

Conclusions

Comfort field of view (video 1)

The result of this test was very clear: level 3 and 4 got the same rating followed by level 5.

Guiding to the speaker (video 2)

According to the average rating of the users the following preferences of the different approaches apply: 1) arrow 2) compass and 3) sided text (position).

The following improvements were requested by the testers:

- The colours of the different speakers should be introduced once for all visible speakers when he/she starts to speak. In the test content the subtitles of the main speaker (radio moderator) were white and of the second speaker yellow. This was not introduced when the conversation starts.
- The compass symbol could be larger and also in the same colour as the speaker to whom it guides.
- The guiding should stop when the face of the speaker is completely in the field of view.

3.2.6. Second iteration: sign language in Germany

Pre-pilots took place at RBB 11-13 April 2018. The whole report is available here: https://drive.google.com/open?id=1PagQb9xk8rdoLypfXupCGJN7oeV7w5C6yYsxUzcUjVg

Participants profile

- Number of end users: 5.
- Demographics for users

1.	Sex	2 male 3 female 0 other 0 prefer not to reply
2.	Age	40 – 60 years old
3.	Main language of the participants:	German German sign language
4.	I define myself as a…"	0 blind person 1 low vision person 3 deaf person 2 hearing impaired person 0 deaf-blind person
5.	Age in which your disability began	3 From birth 1 0-4 1 5-12 0 13-20 0 21-40 0 41-60 0 More than 60
6.	What technology do you use on a daily basis? You can select more than one.	5 TV 3 PC 4 Laptop 5 Mobile Phone 2 Tablet
7.	Do you have any device to access virtual reality content?	5 No 0 I don't know or I don't want to reply.

Table 14: Demographics for users (RBB)

Results

The tables below summarise the results following the template in Annex XII. The first row indicates the feature put to test, namely six levels of comfort field of view, and two guiding mechanism (positioning left or right / arrow). The first column refers to the participant number

(P1= participant 1). The values given by each participant correspond to the methodology described in Annex IX: each participant had to write on a 1 to 5 scale their assessment of each comfort viewing field as well as their assessment of the two implemented systems to guide the user to the speaker location and of the usefulness of the forced perspective.

Comfort	**Level 1**	**2**	**3**	**4**	**5**	**6**
Field of view **Font size in pixel**	**30%** **48**	**40%** **42**	**50%** **36**	**60%** **30**	**70%** **24**	**80%** **18**
P1	1	4	4	4	2	2
P2	3	3	3	4	3	2
P3	2	3.5	5	5	2	2.5
P4	1	5	3	3	3	1
P5	2	2	2	2	4	2
Mean	1.8	3.5	3.4	3.6	2.8	1.9

Table 15: Results SL: comfort field of view (RBB)

Guiding	**Position**	**Arrow**
P1	1	1
P2	4	3
P3	4	3
P4	4	4
P5	3	5
Mean (with all testers)	3.2	3.2
Mean (exclusion of rating of P2)	3	3.3

Table 16: Results SL: guiding mechanism (RBB)

Forced perspective	
P1	4
P2	2
P3	5
P4	3
P5	5
Mean	3.8

Table 17: Results SL: forced perspective (RBB)

The aim of the action was to obtain qualitative data through a reduced sample of users, therefore statistical tests were not performed. Qualitative feedback from users is summarised next.

Conclusions

Comfort field of view (video 1)

The result of this test was very clear: level 2 to level 4 were the best rated ones in the following order: level 4, level 2 and level 3.

Guiding to the speaker (video 2)

When taking into account the ratings of all testers, both approaches got the same average rating. However, tester number two (P2) did not understand the 360° view, so we excluded the

rating from the average calculation. The result was a slightly higher average rating for the positioning on the left and right edge. The evaluation of the remarks from the testers did not help to make out a substantially clearer preference.

The following improvements were requested by the testers:

- The guiding should not be active for short sentences or single words. This applies for both approaches.
- One tester asked for the option to change the position of the sign language interpreter window to the left edge of the FOV.

Forced perspective (video 3)

The idea of the forced perspective was to help to identify the speaker when he/she starts to talk by changing the field of view automatically to the speaker. This was implemented only once for one speaker at the beginning of the video.

All testers liked the idea but asked for a slower movement.

4. CONCLUSIONS

Following ImAc's user centered methodology, users have been involved from the very beginning of the project, providing feedback through a series of qualitative actions.

After identifying the workflow and classifying the users into two main categories (professionals and advanced home users) (see Figure 1), a first iteration took place and focus groups were conducted by RNIB and UAB on audio description and audio subtitling, by RBB and CCMA on sign language, and by RBB, CCMA and UAB on subtitling.

This first iteration was conducted shortly after the start of the project and did not use any specific prototypes developed as part of ImAc. Rather, a more general approach was taken, asking users about their needs, requirements and opinions. The focus was on identifying the challenges of access services in 360° content and discussing possible solutions.

Professional users considered which new tools would be necessary to produce access services for 360° videos and additionally pointed out that both a 2D and a 360° preview would be necessary producing the content. While sign language producers considered that sign language interpretation for 360° videos would be produced in a very similar way as for 2D content, professional audio describers saw a challenge in the new format. In particular, they suggested the possibility of simultaneous audio descriptions. Professional subtitlers furthermore reflected on ways to guide users to the current speaker and mentioned an arrow, compass/wind rose or positioning the subtitles near the speaker in the current field of view as solutions.

Home users of the visual access services stressed that both subtitles and sign language videos should always be visible in the field of view, while the specific position should be customizable. It was furthermore important to them to have a proper speaker identification: sign language users suggested the use of arrows or "forced perspective" (automatic change of the field of view towards the speaker), while subtitle users mentioned the use of different colours. The latter user group generally expressed that subtitles should be prepared according to current standards and adapted to the challenges of immersive environments if necessary. They also pointed out that information on important sounds was necessary and suggested the use of icons or vibration feedback as alternatives to the traditional textual information. Home users of the audio access services highlighted the benefit of using immersive audio for both the main content as well as the access services and the need for voice interaction with the user interface. Furthermore, the challenges of differentiating between main and secondary action were discussed. It was important to the users that secondary narratives should only be audio described if actively triggered. Some blind users indicated that they would prefer to consume the immersive content without HMD but with headphones only.

Following the early focus groups, users' input was used to actually implement some prototypes and carry out a second round of home user testing (pre-pilot 1 actions), which served as a tool to fine-tune the results of the first iteration and prepare the pilots in WP5. RNIB and UAB carried out focus groups on audio description, CCMA and RBB performed interviews on subtitling, and RBB interviewed users of sign language interpreting services. In all of these user interactions, the participants were shown different solutions for the challenges identified in the first iteration of tests and they were asked about their preferences. For the visual access services, the tests focused on preferred sizes of the safe area to display the services and on preferred ways of guiding the users to the current speaker. In the tests regarding the audio access services different types of audio description were presented to the users, taking advantage of immersive audio technologies.

Regarding the safe area for the display of visual access services, subtitle and sign language users preferred similar sizes, even though a slightly larger area was favoured by hearing subtitle users. Participants also proposed specific improvements on personalization of subtitles, location of icons, and on how to indicate direction and speakers. As methods for guiding users to the current speaker the arrow, compass and “forced perspective” were favoured quite evenly among the users. When comparing traditional stereo audio description to two methods using immersive audio technologies, the latter were mostly preferred. These types of audio description enabled a greater immersion of the participants and were seen as beneficial.

Overall, a total of 77 users have contributed to define user scenarios, needs and requirements, distributed as follows: 21 users for audio description and audio subtitling in the first iteration and 10 users in the second iteration (total = 31 users); 25 users for the first iteration of subtitling and sign language and 21 for the second iteration (total= 46 users).

All in all, the end user contribution is central in the ImAc project both before, during and after technological development and implementation. This document has described the user-related actions performed in the first year of the project, as part of Task 2.1. The results of these actions were the basis for the first two iterations of user requirements (see deliverable D2.2) and had a direct impact on the actual pilots developed as part of WP5.

5. REFERENCES

[1] Kitzinger, J. (1995). Qualitative research: Introducing focus groups. *British Medical Journal*, *311*, 299-302.

[2] Morgan, D. L. (1997). *Focus groups as qualitative research* (2nd ed.). Thousand Oaks, CA: Sage.

[3] Barret, J & Kirk, S. (2000) Running focus groups with elderly and disabled elderly participants. *Applied ergonomics*, 31, 621-629.

[4] Seymour, J., Ingleton, C., Payne, S., & Beddow, V. (2003). Specialist palliative care: Patients' experiences. *Journal of Advanced Nursing*, *44*, 24-33.

[5] Harris, J., Foster, M., Jackson, K., & Morgan, H. (2005). *Outcomes for disabled service users.* Retrieved March 22, 2006, from http://www.york.ac.uk/inst/spru/pubs/ccatreps.htm#2005

[6] Kroll, T., Barbour, R. and J. Harris (2005) Using Focus Groups in Disability Research. *Qualitative Health Research* Volume 17 Number 5 May 2007 690-698 © 2007 Sage Publications 10.1177/1049732307301488

[7] Sheikh, A., Brown, A., Watson, Z., & Evans, M. (2017). Directing attention in 360-degree video, *IBC 2016*, Amsterdam, 9-13 September.

[8] Dooley, K. (2017). Storytelling with virtual reality in 360-degrees: a new screen grammar. *Studies in Australasian Cinema 11(3)*, 161-171.

[9] Mateer, J. (2017). Directing for Cinematic Virtual Reality: How the Traditional Film Director's Craft Applies to Immersive Environments and Notions of Presence. *Journal of Media Practice 18(1)*, 14–25.

[10] Gödde M., Gabler F., Siegmund D., & Braun, A. (2018). Cinematic Narration in VR – Rethinking Film Conventions for 360 Degrees. In J. Chen, & G. Fragomeni (eds), *Virtual, Augmented and Mixed Reality: Applications in Health, Cultural Heritage, and Industry*. Cham: Springer.

[11] European Broadcasting Europe (2016) *Access Services Paneuropean Survey* https://www.ebu.ch/files/live/sites/ebu/files/Publications/Presentations/EBU%20Access%20Services%20Survey%202016.pdf Access in January 2019

[12] Maszerowska, A., Matamala, A., & Orero, P. (eds) (2014). *Audio description. New perspectives illustrated*. Amsterdam: Benjamins.

[13] Matamala, A., & Orero, P. (eds) (2016). *Researching audio description. New approaches*. London: Palgrave Macmillan.

[14] Orero, P. (2007). Audiosubtitling: A Possible Solution for Opera Accessibility in Catalonia, *TradTerm* 13, 135-149.

[15] Verboom, M., Crombie, D., Dijk, E., & Theunisz, M. (2002). Spoken subtitles: Making subtitled TV programmes accessible. In K. Miesenberger, J. Klaus, & W. Zagler (eds), *Computers Helping People with Special Needs* (295-302). 8th International Conference, ICCHP 2002, July 15-20, Linz, Austria.

[16] Barbour, Rolaine (2008). *Doing Focus Groups*. London: SAGE.

ANNEX I – T2.1. USER CENTERED DESIGN: TABLE

ImAc Immersive Accessibility		**WP2** **T2.1. User centered design** **Stage 1: table**
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 761974		

T2.1. USER CENTERED DESIGN: TABLE

System side	Technical Component	User: professional or advanced home user	User scenario: indicate what the previous user will experience and how	input from
Design and architecture				
Content Management		professional	User can • save a processed ImAc file (ST or sign language video) via a save-file GUI • open an ImAc file from repository via open-file GUI	RBB

			• select an existing ImAc file via mouse click • manually assign an ImAc file to a video asset	
	1. Web site (for the users) 2.- Web services (for external systems) 3.- Communication protocols (for automations)	professional 1.- ImAc platform manager 2.- ImAc content producer 3.- ImAc content verifier 4.- ImAc content distributor 5.- Broadcaster continuity 6.- Maintenance user 7.- Broadcaster	1.- The ImAc platform manager will give access to the users and set the platform parameters. 2.- The ImAc content producer will upload the ImAc contents to the web site and assign metadata. 3.- The ImAc content verifier will check if a programme has ImAc content and this is right (well synchronized with the video…), for that the corresponding ImAc production/edition tool will be opened automatically from the ImAc preview website. It may also assign programme ID to the verified ImAc contents. 4.- The ImAc content distributor will search for contents and export to different formats. 5.- The Broadcaster continuity will access the system when automation has failed and to trigger Play, Stop, etc. They will also access to the content when there is a synchronisation problem, in this case it will not be necessary the use of the editor but apply an offset to the ImAc content from the Content Management interface. 6.- The maintenance user will access the system to check the platform events in case of an incident. 7.- The broadcaster will generate reports between dates and content type.	ANGLA
	Web GUI & API Interface	professional	User will use Web GUI Interface to manually control the ImAc platform, and API to automatically access & control the ImAc platform from broadcaster OTT related workflows: • Upload/Download video, audio, subtitle, sign language, AD contents. • Monitor ImAc contents • Download content to broadcaster premises	CCMA

Content packaging and distribution		professional	User can: ● trigger/monitor automatic, open and closed (as nowadays subtitles) packaging and distribution of ST and SL enhanced omnidirectional media	RBB
		professional	User can: ● Ingest an ImAc setup described by: ○ 1) output from production (to retrieve content and authoring infos) and ○ 2) user experience (player). ● Retrieve the encoded and packaged ImAc result (to be consumed by a player).	MSE
	Cloud coding & packaging	professional	User can use cloud tools to: ● Live content coding ● VoD content coding ● Monitor content coding status ● Add ImAc ST to a pre-coded content. ● Add ImAc Audio to a pre-coded content.	CCMA
Interface: the signalisation of accessibility services in the content stream.			Accessibility services will be automatically signalled when they are present	RBB
	New signalisation info in Dash manifest	professional	User will have ImAc accessible contents automatically shown when playing video/audio thanks to this in-manifest signalisation recognised by the player. This signalisation should become a standard, and must be ignored by non-compatible or legacy players.	CCMA
Interface: handling presentation settings from different		professional	Depending on the player and its operating system a user can ● pre-define the presentation settings (fixed for the end user or offering him a selection of different options) for every targeted content stream from a given number of presentation variants	RBB

layers (like content stream, operating system, player),				
Interface: providing access to accessibility services			Automatically done	RBB
	1. Screen reader integration (JAWS NVDA, Voiceover, Talkback) 2. Speech Recognition API 3. Compliance with the requirements in WCAG 2.0 (video playerS)	Advanced home user	1. Audio feedback: Depending on the platform, users with sight loss will be able to use audible feedback from JAWS/ NVDA/ Voiceover/ Talkback to navigate the player and control functions such as play, forward, rewind, stop, pause, volume, skip etc. 2. Voice commands: A voice controlled ecosystem will allow users with sight loss (also with mobility issues) to navigate the player by voice and choose any of the controls mentioned above 3. Screen magnification users will be able to navigate the player using magnification software such as Zoomtext.	RNIB
	HbbTV improved interface accessibility	Advanced home user	End user will have an accessible interface to access to content: ● Voice guided for people with vision difficulties or blind. ● Bigger buttons & bigger font letters for people with vision difficulties.	CCMA
Player for		Advanced	User can:	RBB

mobile phone, TV, head mounted display		home user	• start, pause, resume and stop playback of ImAc enhanced omnidirectional media • watch ImAc enhanced media in a plain video view or in a connected HMD	
	ImAc player	Advanced home user	The user will enjoy the experience in the tv, tablet or head mounted display or any combination of these devices. The experience will be synchronised across devices. The player will be available as a web application, so the user won't need to download anything. The tv will start the show, the complementary 360 video is distributed synchronised to the main show, the other devices will show new Audio description and subtitle services adapted to each user impairment The user will access to the contents published in the server and will enjoy an experience adapted to the device where he/she is consuming the content. The user will adapt the device that will store his/her preferences.	**i2CAT**
	Multiplatform player	Advanced home user	Automatic adaptation to different interfaces and platforms. Automatic detection and presentation of accessibility contents.	**CCMA**
Player for preview in editor tools		professional Users of the production tools and editors (for subtitling, audio description and sign language). These users are identified below in "Editor tools	The users of the web tools and editors need to use the player to position and preview the result when they create the individual items (subtitle, audio description segment or sign language segment). They will also see the final result over the video using the player.	**ANGLA**

		(WP4)"		
Integration and testing	Integration module system installed in our premises	professional	(would be related with content management, packaging and distribution) Professional user would use this module to integrate IMAC cloud platform with internal broadcaster systems & workflows (CMS, MAM, Accessibility content management systems…)	**CCMA**
Service side	**Technical Component**	**User: professional or home user**	**User scenario: indicate what the previous user will experience and how**	**input from**
Accessibility interface **Subtitles**		Advanced user	User can ● switch off and on ST presentation ● watch ST in omnidirectional media ● choose from different presentation styles ● choose from different ST feedback styles	**RBB, CCMA**
Accessibility interface **Audio Description**		Advanced user	Users can ● Hear different audio description depending on the current angle visualization.	**CCMA**
		Advanced home user	User preferences: - Users with sight loss will have the ability to control the audio description track in relation to the main media i.e., volume of the track and type – object based, surround, stereo etc. - User will have access a customised experience - player to have the ability to remember user preferences - -	**RNIB**

			- - i.e., user with sight loss should have audio description switched-on as default.	
Accessibility interface **Sign Language**		Advanced home user	User can ● switch off and on SL presentation ● watch SL in omnidirectional media ● choose from different SL feedback styles ● choose from different SL feedback styles	**RBB, CCMA**
Accessibility interface **Audio Subtitles**		Advanced home user	User preferences: - Users with sight loss will have the ability to control the audio subtitle track in relation to the main media i.e., volume of the track and type – object based, surround, stereo etc. - User will have access a customised experience - player to have the ability to remember user preferences i.e., user with sight loss should have audio subtitles switched-on as default.	**RNIB**
Editor Tools	**Technical Component**	**User: professional or advanced home user**	**User scenario: indicate what the previous user will experience and how**	**input from**
Audio production tools	Web audio description tool (stage 1) Professional audio description editor (stage 2)	professional 1.- ImAc content producer 2.- ImAc content verifier 3.- Broadcaster	1.- The ImAc content producer will use the web tool (in stage 1) or the professional editor (in stage 2) to make the audio descriptions, that is to create each audio description segment with text, TCs, audio, segment metadata (attenuation, position, etc.) and preview it over the video. It will have graphics to help the user (vumeter, waveform, thermometers, time left during the recording, etc.), video controls (frame jump, slow speed, etc.) and edition facilities (key shortcuts, segment operations such as insert/delete/test, text edition such as cut/paste and search/replace, file operations such as import/export audio track, etc.). It will use the same tool or editor to verify the complete result over the video. The web tool or the editor will also let the user set the parameters (audio parameters, speed, time restrictions, windows setup, default values, etc.).	**ANGLA**

		continuity	2.- The ImAc verifier will use the web tool or the editor to verify the integrity of the ImAc content. For that it will load the ImAc content file and the video file in the web tool or editor and test it over the video to check the audio-segment synchronisation. 3.- The Broadcaster continuity user will use the web tool or the editor to correct some continuity incidents such as TC offset.	
	Object-based audio editor	professional	**Stand-alone software components that can be used on different operating systems** · The user will be able to connect the editor and renderer software to his DAW (Digital Audio Workstation) via MIDI TC or another protocol · The user can compose object-based audio scenes and monitor the rendering in real-time for various output formats (e.g. 2.0, 5.1, 22.2, headphones, …) · The user will be able to add certain accessibility related features, e.g. o Audio Description objects o Interaction or adjustment ranges for speech/dialogue related objects · The editor will export an ADM (Audio Definition Model) file which can be used for the distribution of the content NOTE: The integration of a 360° video player within the editor should be possible, but depends on the interest of other partners to conduct a pilot or trial for enhanced audio description with 360° content.	**IRT**
	Web GUI Immersive Audio description production tool	professional	User will produce immersive audio description through a Web GUI from where it's possible to monitor a low quality flat video and generate audio description sequences linked to different angle visualizations.	**CCMA**

Audio reception tools	**Object-based renderer for Web Browsers**	**Professional /advanced home user**	**A framework written in JavaScript which can be used for rendering of object-based audio content in Web Browsers** · The professional user (Web developer, content provider, etc.) can use the framework to receive and render object-based audio content in ADM format. It may be integrated in any HTML page. · Depending on the authoring metadata within the ADM file or stream and the graphical user interface, the end user has certain possibilities to personalize and / or interact with the program: o Adjust the level of speech/dialogue within the allowed range o Enable / disable additional Audio Description tracks o Adjust the level of the Audio Description o Navigate within the audio scene and adapt head movements (especially useful for headphone listening) o Render the program for various output formats (e.g. 2.0 and headphones) NOTE: The combination of the web renderer with a 360° video player is basically possible, but depends on the interest of other partners to conduct a pilot or trial for enhanced audio description with 360° content.	**IRT**
Subtitling Tools		professional	User can ● preview the authoring outcome in low res (plane or 360°) ● produce ST by inserting text with the keyboard ● (re-)structure text input into ST frames (e.g. two lines of 36 characters each) ● (re-)assign ST frames to positions in time, height and viewing angle ● (re-)define presentation styles of ST frames ● (re-)define which presentation style will be offered to the end user ● (re-)define the ST feedback (fixed absolute position | fixed viewing position | confinedly	**RBB**

			following viewing angle | added graphical hints where ST are located)	
	Web Editor Tool	professional	The exact functionalities of the editor depend on what will be defined in the user requirements in ImAc. The following features are planned: ● The user will open the editor in a web browser. ● The user will be able to edit basic subtitle data: subtitle text, timing, foreground and background colour. ● The editor will limit the text length per line and the number of lines, such that the created subtitles are compatible to TV distribution (Teletext). ● The editor will provide some quality checks to guide the user (e.g. ensure that subtitles will be visible long enough to read them) ● The editor will show the video in 2D (i.e. image is probably distorted). Later (for the second pilot) the editor should be able to view video and subtitles in a 360°environment (e.g. via a HMD or a tablet). ● The user will be able to edit position of the subtitle. Details need to be defined. To be clarified in ImAc: ● What degree of freedom should be supported (e.g. should vertical positioning of subtitles be supported?) ● It is not planned to support “moving positions” (e.g. a subtitle following a person in the video) – Is there a demand for it? Probably, not all features will be implemented for the first pilot. Notes: I would suggest asking professional subtitle editors how they could imagine to present subtitles in 360° or VR to the user. VR also adds the question where a subtitle is positioned in 3D. These are the more interesting questions we need to answer in phase 1, I think. The requirements for the editor software (e.g. regarding UI) are probably similar to those in standard editing environments for most parts. The main question regarding the editor would be, if an editor prefers to work in a 360° environment (HMD?), or if he would edit in a usual 2D interface	**IRT**

			and would just check the results in 360°/VR.	
	Web GUI Immersive subtitle production Tool	Professional 1 - ImAc content producer	User will produce immersive subtitles through a Web GUI from where it's possible to monitor a low quality flat video and reference the subtitle text to some position or angle in the panoramic view. Subtitles will be marked as Closed Captions when needed.	**CCMA**
	Web subtitling tool (stage 1) Professional Subtitling editor (stage 2)	professional 1.- ImAc content producer 2.- ImAc content verifier 3.- Broadcaster continuity	1.- The ImAc content producer will use the web tool (in stage 1) or the professional editor (in stage 2) to make the subtitling, that is to create each subtitle with text, TCs, position, alignment, programme character (with its colours, font, etc.) and preview it over the video. It will have graphics to help the user (waveform, thermometers, etc.), video controls (frame jump, slow speed, etc.) and edition facilities (key shortcuts, subtitle operations such as insert/delete/jump, text edition such as cut/paste and search/replace, file operations such as import/export, etc.). It will use the same tool or editor to verify the complete result over the video. The web tool or the editor will also let the user set the programme characters (number, font, colour, size, etc.) and other parameters (letters per line, reading speed, time restrictions, windows setup, default values, etc.). 2.- The ImAc verifier will use the web tool or the editor to verify the integrity of the ImAc content. For that it will load the ImAc content file and the video file in the web tool or editor and test it over the video to check the audio-subtitle synchronisation. The editor can be automatically opened from the Content Management with its corresponding ImAc content and the changes will be recorded directly in the Content Management. 3.- The Broadcaster continuity user will use the web tool or the editor to correct some continuity incidents such as TC offset. The editor will be automatically opened from the Content Management with its corresponding ImAc content and the changes will be recorded directly in the Content Management.	**ANGLA**
Sign Language Editor	Web Sign language tool (stage 1)	1.- ImAc content producer	1.- The ImAc content producer will use the web tool (in stage 1) or the professional editor (in stage 2) to make the sign language, that is to create each sign language segment with text, TCs, video, segment metadata (size, position, etc.) and preview it over the video. It will have graphics to help the user (waveform, thermometers, time left during the recording, etc.), video controls	**ANGLA**

	Professional sign language editor (stage 2)	2.- ImAc content verifier 3.- Broadcaster continuity	(frame jump, slow speed, etc.) and edition facilities (key shortcuts, segment operations such as insert/delete/test, text edition such as cut/paste and search/replace, file operations such as import/export video, etc.). It will use the same tool or editor to verify the complete result over the video. The web tool or the editor will also let the user set the parameters (video parameters, speed, time restrictions, windows setup, default values, etc.). 2.- The ImAc verifier will use the web tool or the editor to verify the integrity of the ImAc content. For that it will load the ImAc content file and the video file in the web tool or editor and test it over the video to check the synchronisation. 3.- The broadaster continuity user will use the web tool or the editor to correct some continuity incidents such as TC offset.	
		professional	User can ● preview the authoring outcome in low res (plane or 360°) ● add sign language video items (SL) to omnidirectional media ● split SL files into clips ● (re-)assign SL clips to positions in time, height and viewing angle ● (re-)define presentation styles of added SL clips (lighting, contrast etc.) ● (re-)define the SL feedback (fixed absolute position | fixed viewing position | confinedly following viewing angle | added graphical hints where SL are located)	**RBB**
Pilots	**Technical Component**	**User: professional or advanced home user**	**User scenario: indicate what the previous user will experience and how**	**input from**
Content production	Web tools of WP4 (stage 1) Professional editors of WP4 (stage 2)	1.- ImAc content producer 2.- ImAc content verifier	These user scenarios are described above in “Editor tools (WP4)” for each type of web tool or editor.	**ANGLA**

		professional	Users can produce ImAc content ● ST enhanced media ● SL enhanced media	**RBB, CCMA**
Pilots **Content consumption**		Advanced home user	Users can consume ImAc content ● ST enhanced media ● SL enhanced media ● Audio described media (in different audio formats) +audio subtitles [for the cross national pilot] ● Media with 3D audio and no audio description but with audio subtitles [for the cross national pilot]	**RBB, CCMA** **RNIB**

ANNEX II – T2.1. PROFESSIONAL USERS

ImAc Immersive Accessibility		**WP2** **T2.1. User centered design** **Stage 1: Professional user side**
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 761974		

System side	Technical Component	User	User scenario: indicate what the previous user will experience and how	input from
Content Management			User will use Web GUI Interface to manually control the ImAc platform, and API to automatically access & control the ImAc platform from broadcaster OTT related workflows: • Access the platform • Upload/Download video, audio, subtitle, sign language, AD contents. • Monitor ImAc contents • Download content to broadcaster premises • Export into different formats a processed ImAc file (ST, sign language video or AD) via a save-file GUI • Select an existing ImAc file via mouse click • Manually assign an ImAc file to a video asset • Check integrity of ImAc file • Check synchronisation between main and ImAc content, alarm if necessary • Report status of available data and its content type	RBB, CCMA, ANGLA

			Content Management - Input for focus groups: ❏ Is there a specific content management system for your work today? ❏ What file formats do you currently work with (for video, ST, AD, SL) ❏ Do you like open-file-menus? ❏ Do you like drag and drop? ❏ Do you check integrity today, if yes how? ❏ What feedback do want from the content manager? number of emissions, … ❏ ...	
Content packaging and distribution			User can: ● Ingest an ImAc setup from the Content Management described by: ○ 1) output from production (to retrieve content and authoring infos) and ○ 2) user experience (player). ● Trigger/monitor automatic, open and closed (as nowadays subtitles) packaging of ST, SL, AD enhanced omnidirectional media ● Trigger/monitor automatic, open and closed (as nowadays subtitles) distribution of ST, SL, AD enhanced omnidirectional media ● Monitor status of content packaging and distribution ● Retrieve the encoded and packaged ImAc result (to be consumed/previewed by a player). **Content packaging and distribution - Input for focus groups:** ❏ How do you monitor the processing of your results today? How does it look like? ❏ Do you need a preview here? ❏ …	RBB, MSE, ANGLA
Interface: the signalisation of accessibility services in the			Accessibility services will be automatically signalled when they are present. Users can: ● Monitor signalisation of ImAc content ● Configure signalisation of ImAc content	RBB, CCMA

content stream.			• Preview signalised ImAc content **Signalisation of accessibility services in the content stream - Input for focus groups:** ❑ Do you like drag and drop? ❑ Do you check integrity today, if yes how? ❑ How would you signalize accessibility in broadcast contents? ❑ How would you signalize accessibility in broadband contents?	
Interface: handling presentation settings from different layers (like content stream, operating system, player),			Depending on the player and its operating system a user can • pre-define the presentation settings (fixed for the end user or offering him a selection of different options) for every targeted content stream from a given number of presentation variants **Interface: handling presentation settings from different layers (like content stream, operating system, player) - Input for focus groups:** ❑ Which presentation settings do you see? Which options should the user have to personalize the accessibility services? ❑ Which presentation variants for subtitles, audio description and signer do you want to offer? ❑ Do you want to offer a selection of variants to choose from? ❑ For ST and SL: ❑ What do you think about viewing angle feedback mechanisms (for instance "ST follows view", "ST is fixed on objects")? ❑ What do you think about event notifications ("there is something happening/somebody speaking behind you")? ❑ ...	RBB
Integration and testing			(Is related with content management, packaging and distribution) Professional user would use this module to: • integrate IMAC cloud platform with internal broadcaster systems & workflows	CCMA

			(CMS, MAM, Accessibility content management systems…) **Integration and testing - Input for focus groups:** ❑ Which systems of the broadcasters environment should be interact with the ImAc platform? ❑ ...	
Editor Tools	**Technical Component**	**User**	**User scenario: indicate what the previous user will experience and how**	**input from**
Player for preview in editor tools			The users of the web tools and editors need to: ● Use the player to position and preview the result (angle, frame jump, slow speed, etc.) when they create the individual items (subtitle, audio description segment or sign language segment) ● See the final result over the video using the player. **Player for preview of content - Input for focus groups:** ❑ Is it sufficient to have a spherical preview in a plain web player? ? And for monitoring the final result? ❑ Is it necessary to have a preview in a head mounted display and/or smartphone with VR glasses? And for monitoring the final result? ❑ ...	ANGLA
Audio production tools			The ImAc content producer will use the web tool (in stage 1) or the professional editor (in stage 2) to: ● Perform file operations such as import/export audio track ● Preview and control the main video (angle, frame jump, slow speed, etc.) ● Create plain (object-based see below) audio description segment with text, TCs, audio, segment metadata (attenuation, position, etc.) ● Edit AD data (key shortcuts, segment operations such as insert/delete/test, text	ANGLA, CCMA, IRT

			edition such as cut/paste and search/replace) ● Compose object-based audio scenes and monitor the rendering in real-time for various output formats (e.g. 2.0, 5.1, 22.2, headphones, …) ● Export an ADM (Audio Definition Model) file which can be used for the distribution of the content ● Add certain accessibility related features, e.g. ○ Audio Description objects ○ Interaction or adjustment ranges for speech/dialogue related objects ● Preview the result with the video ● Verify the quality with the help of graphics to help the user (vumeter, waveform, thermometers, time left during the recording, etc.), also let the user set the parameters (audio parameters, speed, time restrictions, windows setup, default values, etc.). ● Verify synchronisation with main video ● Adapt synchronisation **Audio production tools - Input for focus groups:** ❑ … ❑ ...	

Audio reception tools			A framework written in JavaScript which can be used for rendering of object-based audio content in Web Browsers. The professional user (Web developer, content provider, etc.) can use the framework to: ● receive and render object-based audio content in ADM format. It may be integrated in any HTML page. ● personalize and / or interact with the program: ○ Adjust the level of speech/dialogue within the allowed range ○ Enable / disable additional Audio Description tracks ○ Adjust the level of the Audio Description ○ Navigate within the audio scene and adapt head movements (especially useful for headphone listening) ○ Render the program for various output formats (e.g. 2.0 and headphones) **Audio reception tools - Input for focus groups:**	**IRT**
Subtitling Tools			User can ● Perform file operations such as import/export ST files ● Preview and control the main video (angle, frame jump, slow speed, etc.) ● Preview the authoring outcome in low res video (plane or 360°) ● Produce ST by inserting text with the keyboard ● (re-)structure text input into ST frames (e.g. two lines of 36 characters each) ● (re-)assign ST frames to positions in time, height and viewing angle ● (re-)define presentation styles of ST frames ● (re-)define which presentation style will be offered to the end user ● (re-)define the ST feedback (fixed absolute position | fixed viewing position | confinedly following viewing angle | added graphical hints where ST are located) ● Verify the result regarding synchronisation and defined presentation parameters **Subtitling tools - Input for focus groups:**	**RBB, IRT, CCMA, ANGLA**

			❑ What end user requirements for the presentation? ❑ What degree of freedom should be supported (e.g. should vertical positioning of subtitles be supported?) ❑ It is not planned to support "moving positions" (e.g. a subtitle following a person in the video) – Is there a demand for it? ❑ Notes: I would suggest asking professional subtitle editors how they could imagine to present subtitles in 360° or VR to the user. VR also adds the question where a subtitle is positioned in 3D. These are the more interesting questions we need to answer in phase 1, I think. The requirements for the editor software (e.g. regarding UI) are probably similar to those in standard editing environments for most parts. The main question regarding the editor would be, if an editor prefers to work in a 360° environment (HMD?), or if he would edit in a usual 2D interface and would just check the results in 360°/VR. ❑ During the production of subtitles, do you prefer a 360 view presented in one 2D panoramic view (although distorted) or show only the 2D view in one direction (not distorted view) and use the arrow keys to move around? ❑ ...	
Sign Language Editor			User can ● Perform file operations such as import/export SL files ● Preview and control the main video (angle, frame jump, slow speed, etc.) ● Add sign language video items (SL) to the omnidirectional video ● Split SL files into clips ● (re-)assign SL clips to positions in time, height and viewing angle ● (re-)define presentation styles of added SL clips (lighting, contrast etc.) ● (re-)define the SL feedback (fixed absolute position | fixed viewing position | confinedly following viewing angle | added graphical hints where SL are located) **Sign Language Editor Tool - Input for focus groups:** ❑ Same questions as in ST tool ❑ ...	**RBB,** **ANGLA**

Pilots	Technical Component	User	User scenario: indicate what the previous user will experience and how	input from
Content production			These user scenarios are described above in "Editor tools (WP4)" for each type of web tool or editor. Users can produce ImAc content ● ST enhanced media ● SL enhanced media	**ANGLA,** **RBB,CCMA**

ANNEX III – T2.1. ADVANCED HOME USERS

ImAc Immersive Accessibility		**WP2** **T2.1. User centered design** **Stage 1: Advanced home users**
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 761974		

	Technical Component	User scenario: indicate what the previous user will experience and how	input from
Interface: providing access to accessibility services	1. Screen reader integration (JAWS NVDA, Voiceover, Talkback) 2. Speech Recognition API 3. Compliance with the requirements in WCAG 2.0 (video players)	User uses existing tools which supports him controlling the player and the accessibility interfaces 1. Audio feedback: Depending on the platform, users with sight loss will be able to use audible feedback from JAWS/ NVDA/ Voiceover/ Talkback to navigate the player and control functions such as play, forward, rewind, stop, pause, volume, skip etc. 2. Voice commands: A voice controlled ecosystem will allow users with sight loss (also with mobility issues) to navigate the player by voice and choose any of the controls mentioned above 3. Screen magnification users will be able to navigate the player using magnification software such as Zoomtext. Bigger buttons & bigger font letters for people with vision difficulties.	RNIB, CCMA

		Input for focus groups - questions concerning the access to the accessibility interfaces (subtitles and signer are missing here): ❑	
Multiplatform Player for desktop, mobile phone (cardboard supported, gyroscope sensor based) , TV, head mounted display (e.g. Oculus, Playstation VR, Vive) WP3: Immersive platform, T3.5 player	Web application	• User can start, pause, resume and stop playback of ImAc enhanced omnidirectional media in a plain video view (mobile phone, TV) or in a connected HMD or as a combination of these devices (TV + smartphone + HMD, TV + smartphone, TV + HMD, TV, HMD, smartphone, smartphone + HMD) ○ Automatic adaptation to different interfaces and platforms. ○ Automatic detection and presentation of accessibility contents. • The experience will be synchronized across devices. ○ complementary 360 video is distributed synchronized to the main show, the other devices will show new audio description and subtitle services adapted to each user impairment • The user will adapt the device that will store his/her preferences for the accessibility interface and the player. Synchronisation between devices can be asked in the focus group according to specific use cases - input needed here **Input for focus groups - define use cases for synchronisation between different devices:** ❑ Do you want to consume together with a group of people 360° content? The 360° videos is playing on the TV. The user can navigate in the video by using the arrow keys on the remote control AND/OR is it possible to use a second device like tablet or smartphone to control the viewing direction? This case will imply the synchronisation between the devices. ❑ You need to consume the 360° content with accessibility services. You will use a	RBB, i2CAT, CCMA

		second device like a smartphone with glasses or a head mounted display. Do you see the necessity to consume the content in your device synchronously to the content on the TV or with other devices?	
	Technical Component	**User scenario: indicate what the previous user will experience and how**	**input from**
Accessibility interface: Subtitles **WP3: Immersive platform, T3.5**		User can ● switch off and on ST presentation ● watch ST in omnidirectional media ● choose from different presentation styles ● choose from different ST feedback styles **Input for focus groups:** ❑ Possible scenarios to present subtitles ❑ Subtitles are presented always in front of you (slightly below eye line) and follow the head position, independently to the content ❑ The 360° space is divided into several parts and subtitles are available for each section on a fixed position (slightly below eye line) ❑ Subtitles are presented always in front of you (slightly below eye line) and follow the head position, independently to the content, additionally ❑ The user gets notices about the position of speaker - clarify in the focus group the presentation (another font?, graphics?, background?) ❑ Should it be possible to switch off the notices? ❑ The subtitles are presented in the spherical location where the speaker is ❑ Which settings want the user have to personalize the subtitles? Size, font, background, position (slightly below eye line, besides speaker...?)	**RBB, CCMA**
Accessibility		Users can	**CCMA, RNIB,**

interface:
 	
Audio Description / Audio Subtitles
 	
WP3: Immersive platform, T3.5
		• Hear different audio description/audio subtitles depending on the current angle visualization. (different use cases are defined below, users of the focus groups will give us feedback on that)
		
		User preferences:
		
		• Users with sight loss will have the ability to control the audio description/audio subtitle track in relation to the main media
		  ○ volume of the track and type – object based, surround, stereo
		• Depending on the authoring metadata within the ADM file or stream and the graphical user interface, the end user has certain possibilities to personalize and / or interact with the program:
		  ○ Adjust the level of speech/dialogue within the allowed range
		  ○ Enable / disable additional Audio Description/Audio subtitle tracks
		  ○ Adjust the level of the Audio Description/audio subtitle
		  ○ Navigate within the audio scene and adapt head movements (especially useful for headphone listening)
		  ○ Render the program for various output formats (e.g. 2.0 and headphones)
		• User will have access a customised experience - player to have the ability to remember user preferences i.e., user with sight loss should have audio description/audio subtitles switched-on as default.
		
		Input for focus groups: use cases for AD
		
		• The first is around the placement of audio description of which I can see three main strategies:
		  ○ A, AD position linked to action being described,
		  ○ B, AD position fixed in relation to scene (as if they are sitting on the sofa next to you as you turn your head),
		  ○ C, AD position fixed in relation to users head (like a devil sitting on your shoulder as you turn your body to follow the action).
		
		• B and C then invite a further question on where the audio description should be placed. Do users prefer it off to one side, in front of them, behind them or
			IRT

		potentially even centred on the user. A tool to allow users to place the AD and record their preference could give us some great data. ● B and C also raise the question of how to describe in AD where action is happening if the placement of the AD doesn't communicate that. Do you pair the audio description to a sound effect "[scraping sound placed within the scene] Daniel takes a knife from the block", can you get away with just the AD (or just the sound effect) or do you need to describe where the action takes place "a shadow moves at the back of the room"? ● This is potentially a far trickier piece of work because it requires a variety of content and different storytelling techniques. Since the intention is to build an audio description capable 360 video player then if the tool created in the project can be made available for further research then it would allow this question to be explored further.	
Accessibility interface: **Sign Language** **WP3: Immersive platform, T3.5**		User can ● switch off and on SL presentation ● watch SL in omnidirectional media ● choose from different SL feedback styles ● choose from different SL feedback styles **Input for focus groups:** ❑ Possible scenarios to present a signer ❑ Signer is presented beside the speaker and/or on a fixed position and follow the head position, independently to the content ❑ The 360° space is divided into several parts and signer are available for each section on a fixed position (beside speaker and/or on a defined place) ❑ Signer is presented in front of you beside and/or on a fixed position and follow the head position, independently to the content ❑ The user gets notices about the position of speaker - clarify in the focus group the presentation (another font?, graphics?,	**RBB, CCMA**

		background?, signer?) ❏ Should it be possible to switch off the notices? ❏ Which settings want the user have to personalize the signer? Size, background, position (fixed position, beside speaker...?)	
	Technical Component	**User scenario: indicate what the previous user will experience and how**	**input from**
Pilots, WP5 **Content consumption**		Users can consume ImAc content (rough summary, details are in the lines above concerning accessibility interfaces) ● ST enhanced media ● Signer enhanced media ● SL enhanced media ● Audio described media (in different audio formats) + audio subtitles [for the cross national pilot] ● Media with 3D audio and no audio description but with audio subtitles [for the cross national pilot]	**RBB, CCMA, RNIB**

ANNEX IV – GENERAL QUESTIONNAIRE IN ENGLISH

ImAc Immersive Accessibility		**T1.2. General questionnaire**
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 761974		

eneral questions

1. Sex:

a) Female
b) Male
c) Other
d) I prefer not to reply

2.Age: _____

3.Main language: _____

4.Level of finished studies

a) No studies
b) Primary education
c) Secondary education
d) Further education
e) University

5.I define myself as a…

a) Blind person
b) Low vision person
c) Deaf person
d) Hearing impaired person
e) Deaf-blind person

6.Age in which your disability began:

a) From birth

b) 0-4
c) 5-12
d) 13-20
e) 21-40
f) 41-60
g) more than 60

7.What technology do you use on a daily basis? You can select more than one.

a) TV
b) PC
c) Laptop
d) Mobile phone
e) Tablet

8.Do you have any device to access virtual reality content?

a) Yes (If yes, which one? _____)

b) No

c) I don't know or I don't want to reply

9.Which of the following is your preferred device for watching online video content (i.e., Youtube, Vimeo, Netflix, Amazon Prime, broadcast catch up service etc.)? You can select more than one.

a) PC
b) Laptop
c) Smartphone
d) Tablet
e) I don't watch online video content.
f) Others (if so, please specify:_____)

10.**(only for visually impaired users)** Which of the following do you use on your connected devices to access the above content?

a) Magnification (i.e. Zoomtext)
b) Screen readers (i.e., JAWS, VoiceOver, TalkBack)
c) Both
d) None

11.**(only for visually impaired users)** Which of the following controls would you like to use with your screenreader /magnification tool when watching content online?

a) Browse content library
b) Identify content
c) Functions such as play, stop, pause, forward, rewind
d) Switch AD/AS on and off

ANNEX V – FOCUS GROUP INSTRUCTIONS

ImAc Immersive Accessibility		**WP2** **T2.1. User centered design** **Stage 2: instructions**
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 761974		

T2.1. USER CENTERED DESIGN: INSTRUCTIONS ON HOW TO DEVELOP FOCUS GROUP

The aim of the focus group is to gather input from users (both end users and professionals) on the needs, expectations and requirements for access services in immersive accessibility.

Previous to the focus group

1. Focus groups will ask questions departing from Pro-user-T2.1. compiled table and End-user_T2.1._compiled table. If any partners wants to add questions, please add them on Monday 21st November at the latest.
2. To prepare for your focus group, read the recommendations available here and write down a similar procedure: https://drive.google.com/drive/folders/1JRbnZegeccVuRJsBrqD_E3s4wh5WckEZ
3. Number of participants: a minimum of 5. Number of facilitators: 2.
4. Have the documentation concerning consent form and information sheet available in your language (see step 1 below)

During the focus group, please follow these steps:

1. Welcome participants and ask them to read (or read them out) the consent form and the information sheet, available here:

   https://drive.google.com/open?id=1G-qpAqo3tdFE34voFPY-chOPDlWMjkAq

   Version in English: T2.1._IMAC INFORMATION SHEET and consent_final.docx

   Version in Catalan: T2.1._IMAC INFORMATION SHEET and consent_final_ CAT.docx

   Version in Spanish: T2.1._IMAC INFORMATION SHEET and consent_final_ES.docx

2. Ask participants to sign consent forms and information sheets.

   (Please keep the signed consent forms and hand in to Pilar in next meeting in Barcelona February 2018)

3. Ask participants to fill in the demographic information. Please notice that two questions do no apply to professional users without disabilities.

   https://drive.google.com/open?id=1G-qpAqo3tdFE34voFPY-chOPDlWMjkAq

   Version in English: T2.1. IMAC_GENERALQUESTIONNAIRE_FINAL.docx.

   Version in Catalan: T2.1. IMAC_GENERALQUESTIONNAIRE_FINAL_CAT.docx.

   Version in Spanish: T2.1. IMAC_GENERALQUESTIONNAIRE_FINAL_SPA.docx.

4. Start focus group, following suggested structure and taking into account the list of questions provided. Perhaps you can have a photograph taken, or interview the participants for one of the ImAc films we have to provide.

5. Summarise the results of the report and read them aloud to the participants, who have to approve them.

After the focus group

6. Send your focus group report approx. 01/12/2017. It must follow the template available here:

   https://drive.google.com/open?id=1TUjkFCnK3Dcx1x1HP2zoEB5UOJHs6lSL

   (T2.1.focusgroupreporting_v1, under FOCUS_GROUP_ALL folder)

7. Provide original consent forms to UAB (next February face-to-face meeting or send by snail mail).

ANNEX VI – FOCUS GROUP REPORTING TEMPLATE

ImAc Immersive Accessibility		**T2.1.** **Focus group reporting template** **17.11.17, v1**
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 761974		

Focus group template

1. Focus group general information

- Partner responsible for the workshop: *please indicate your acronym.*
- Place and date: *please indicate place and date.*
- Access service(s) discussed: *please indicate if SDH, SL, AD or AST.*

2. Participants profile

- Number of home users: *please indicate only number of home users (persons with disabilities).*
- Number of professional users and profile: *please indicate only number of professional users and profile (subtitlers, audio describers, sign language interpreters).*
- Demographics for users (*please notice that questions 5 and 6 may not be applicable): (please add your numbers where an X is found).*
    1. Sex: X "male", X "female", X "other", X "prefer not to reply".
    2. Age: *please specify the different ages*
    3. Main language of the participants: *please specify the different languages*
    4. Level of finished studies: X "no studies", X "primary education", X "secondary education", X "further education", X "university".
    5. "I define myself as a…" X "blind person", X "low vision person", X "deaf person", X "hearing impaired person", X "deaf-blind person".
    6. Age in which your disability began: X "From birth", X "0-4", X "5-12", X "13-20", X "21-40", X "41-60", X "More than 60".
    7. What technology do you use on a daily basis? You can select more than one. X "TV", X "PC", X "Laptop", X "Mobile phone", X "Tablet".
    8. Do you have any device to access virtual reality content? X "Yes" (*and please indicate which ones they name: XXXX*), X "No", X "I don't know or I don't want to reply".
    9. Please add results for questions 10 and 11 if you had blind and visually impaired users in your testing.

3. Workshop conclusions

3.1. Conclusions from professional users

Please summarise your results here, indicating user needs, expectations and recommendations, taking the compiled table as a reference.

3.2. Conclusions from expert end users

a) Concerning the interaction: accessing the services

Please summarise your results here, indicating user needs, expectations and recommendations, taking the compiled table as a reference.

b) Concerning the services

Please summarise your results here, clearly identifying the access services you refer to, and identifying user needs, expectations and recommendations, taking the compiled table as a reference.

ANNEX VII – T2.1. ETHICAL FORMS

ImAc Immersive Accessibility	GA: 761974

INFORMATION SHEET

Project: ImAc (Immersive Accessibility)

Main researcher: Sergi Fernández (i2Cat)

Ethical adviser: Pilar Orero

The aim of the tests is to get feed-back on how access services can be implemented in immersive media. This will allow us to identify the needs of diverse audiences and research how the quality of experience and the quality of the service can be improved.

During the test, which can take various forms (experiment with questionnaire, focus groups, interviews, etc.), you will be asked to provide some demographic data. Then, you will be asked to watch an input, perform a task or give your opinion on various aspects. If needed, objective data will be recorded during the session. The researcher will give you more details of the specific test assigned to you and the data collection methods. Please ask as many questions as needed to clarify the procedure.

Virtual reality may produce some sort of discomfort such as virtual reality sickness when visualizing virtual reality contents, information will be provided and appropriate measures will be taken to guarantee the participants' safety and well-being. Immersive environments are not recommended for individuals with claustrophobia, heart conditions, back conditions, a history of seizures, epilepsy, and/or sensitivity to flashing lights. Also participants thought to be unstable or under the influence of drugs or alcohol will not be admitted.

In the case that some physiological or eye-tracking apparatus are used to gather data, you will not experience any discomfort, since the apparatus used are the latest generation and are not invasive.

Now please read the consent form.

CONSENT FORM (written version)

Project: ImAC (Immersive Accessibility)

Your participation in the tests is absolutely voluntary.

You can discontinue your involvement in the study at any time without prior justification. This shall have no repercussions or negative consequences of any sort.

In the case that some physiological or eye-tracking apparatus are used to gather data, you will not experience any discomfort, since the apparatus used are the latest generation and are not invasive.

Virtual reality may produce some sort of discomfort such as virtual reality sickness when visualizing virtual reality contents, information will be provided and appropriate measures will be taken to guarantee the participants' safety and well-being. Immersive environments are not recommended for individuals with claustrophobia, heart conditions, back conditions, a history of seizures, epilepsy, and/or sensitivity to flashing lights. Also participants thought to be unstable or under the influence of drugs or alcohol will not be admitted.

The information you provide will be used in the project but it will remain anonymous.

ImAc is a European project led by Sergi Fernández, from the company i2Cat. The ethical adviser responsible of ethical procedures is Pilar Orero. You can contact Pilar Orero at pilar.orero@uab.cat and ask for more information about the project and the project results.

The researcher administering the test is ((NAME and SURNAME)).

If you are willing to participate, please confirm the following statements by signing at the end of this document.

- I have read and understood the information given for this research or have had the information read to me,
- I have had the opportunity to ask questions about the research.
- I consent to take part in the research sessions.

_____

Name of the participant Date Signature

_____

Name of the researcher Date Signature

Signed by Pilar Orero (UAB IP ImAc)

ANNEX VIII - INSTRUCTIONS FOR SUBTITLING PRE-PILOTS

T2.1. USER CENTRIC DESIGN: INSTRUCTIONS ON HOW TO DEVELOP PRE-PILOTS (SUBTITLING)

BEFORE THE PRE-PILOTS

1. **Methodological approach**: although the 1st iteration was based on focus group, the 2nd iteration for subtitling will be based on one-to-one semi-structured interviews with some stimuli followed with questions to trigger discussion
2. **Users**: same home users who participated in focus groups (1st iteration) will be invited to participate in the interviews (2nd iteration). If they cannot attend, other users will be invited. Please notice that the recommended number of participants is 5, with a maximum of 6.
3. **Facilitators**: it is recommended that 2 facilitators are present: 1 interviewer (to give instructions, ask questions, assist) and one to take notes. The interviewer must remain neutral (words, gestures, tone of voice, movements) so as to not bias the respondent's answer.
4. **Room**: make sure the atmosphere is comfortable for the participants, with an adequate room arrangement.
5. **Materials to be ready**:
    a. **Stimuli for subtitling**:
        i. **Comfort viewing field**. 3-minute video split in 6 parts, each part with one of the six comfort viewing field levels applied. 30 seconds per level, showing subtitles top centered, bottom centered, bottom left. All video parts have added at the start a still image (approx. 5 seconds) and a short description about what is tested (example: comfort field of view level 1). At the end of each condition, the following still image will appear with the following text: "How comfortable was viewing the subtitles? Do not take into consideration the device used but the viewing experience. Please rate it on a 1 to 5 scale (1= lowest value, 5= highest value)".
        ii. **Guiding user to speaker**. 3 clips, 30 seconds each, segmented with a still image (approx. 5 seconds) and a short description. Options: position left/right, arrow, compass. The same video snippet will be used for a better comparison between the different options. At the end of each condition, a still image with the following text: "How useful was [arrow/compass/position left and right] to guide you to the speaker? Please rate it on a 1 to 5 scale (1= lowest value, 5 = highest value)".
    b. **Consent form and information sheet**. Available in Google Drive WP2/T2.1/ T2.1. pre-pilot tests/Methodology/Ethical forms. Please notice that this form is slightly different from the one used for the focus groups. German version to be updated.
    c. **Demographic questionnaires**. Available in Google Drive WP2/T2.1/ T2.1. pre-pilot tests/Methodology/Questionnaires. Please notice that this form is the same as the one used for the first round of focus groups.
    d. **Devices**: HMD.

6. **Aim**: ask about user preferences, recommendations, and needs concerning comfort viewing field and guiding system to the speaker.
7. **Previous testing**: please make sure all the equipment works before conducting the interviews.
8. **Expected length**: 45 minutes per participant.

DURING THE INTERVIEW

1. **Welcome** participant and briefly explain ImAc and the aim of the interview (5 minutes)

2. **Signature**. Consent form and information sheet signature (5 min). Please keep the signed consent forms and/or recordings and hand in to Pilar in next meeting in Munich May 2018.

3. **Questionnaires**. Fill in demographic information: one general questionnaire per participant (5 min).

4. **Start tasks and discussion**, following the structure below. (20 min)

The facilitator explains task 1 and aim of it.

Task 1. Participants are shown the stimuli for task 1 and assess it on a 5-point Likert scale. At the end, qualitative information is requested from participants, using the following suggested questions:

- After having seen all the examples, would you change your rating?
- Which one is your preferred one? Why?

Task 2. Participants are shown the stimuli for task 2 and assess it on a 5-point Likert scale. At the end, qualitative information is requested from participants, using the following suggested questions:

- After having seen all the examples, would you change your rating?
- Which one is your preferred one? Why?

5. **Final discussion** (9 minutes)

Please ask user to share their opinions, needs, recommendations, preferences on **all** options shown, and ask them to suggest improvement to current implementation. To lead the discussion, you can ask general preference questions, to be adapted depending on the interview development:

- What are your opinions about these subtitling possibilities?
- What do you like about each of these subtitling possibilities?
- What would you change? How can they be improved? Do you have any recommendations?
- What option feels more immersive?
- What option is more comfortable to follow?
- What option is more comfortable to read?
- Are the options intuitive?

6. Farewell and thanks to participants (1 minutes).

AFTER THE INTERVIEWS

Report: upload your report one week after the interviews under Google Drive under WP2/T2.1/ T2.1. pre-pilot tests/Methodology and let Pilar Orero/Anna Matamala know. It must follow the template available in the same folder ("T2.1._FG_It2_report_SUB").

Consent forms: provide original consent forms to UAB (next May face-to-face meeting or send by snail mail).

ANNEX IX - INSTRUCTIONS FOR SL PRE-PILOTS

T2.1. USER CENTRIC DESIGN: INSTRUCTIONS ON HOW TO DEVELOP PRE-PILOTS (SIGN LANGUAGE)

BEFORE THE PRE-PILOTS

1. **Methodological approach**: although the 1st iteration was based on focus group, the 2nd iteration for SL will be based on one-to-one semi-structured interviews with some stimuli followed with questions to trigger discussion
2. **Users**: same home users who participated in focus groups (1st iteration) will be invited to participate in the interviews (2nd iteration). If they cannot attend, other users will be invited. Please notice that the recommended number of participants is 5, with a maximum of 6.
3. **Facilitators**: it is recommended that 2 facilitators are present: 1 interviewer (to give instructions, ask questions, assist) and one to take notes. The interviewer must remain neutral (words, gestures, tone of voice, movements) so as to not bias the respondent's answer. Please take into account that the presence of a sign language interpreter may be needed during the test.
4. **Room**: make sure the atmosphere is comfortable for the participants, with an adequate room arrangement.
5. **Materials to be ready**:
    a. **Stimuli**:
        i. **Comfort viewing field**. 3-minute video split in 6 parts, each part with one of the six comfort viewing field levels applied and a still image with a short description of what is about to be tested. 30 seconds per level. At the end of each level, the following still image will appear with the following text: "How comfortable was viewing the sign language interpreting? Do not take into consideration the device used but the viewing experience. Please rate it on a 1 to 5 scale (1= lowest value, 5= highest value)".
        ii. **Guiding user to speaker**. 3 clips, 30 seconds each, segmented with a still image (approx. 5 seconds) and a short description. Options: position left/right, arrow. The same video snippet will be used for a better comparison between the different options. At the end of each condition, a still image with the following text: "How useful was [arrow/positioning left or right] to guide you to the speaker? Please rate it on a 1 to 5 scale (1= lowest value, 5 = highest value)".
        iii. **Forced perspective**. 1 minute with a still image (approx. 5 seconds) and short description. At the end, a still image with the following text: "How useful was changing the field of view to see the speaker when s/he starts to speak? Please rate it on a 1 to 5 scale (1= lowest value, 5 = highest value)".
    b. **Consent form and information sheet**. Available in Google Drive WP2/T2.1/ T2.1. pre-pilot tests/Methodology/Ethical forms. Please notice that this form is slightly different from the one used for the focus groups. German version to be updated.
    c. **Demographic questionnaires**. Available in Google Drive WP2/T2.1/ T2.1. pre-pilot tests/Methodology/Questionnaires. Please notice that this form is the same as the one used for the first round of focus groups.

   d. **Devices**: HMD.
6. **Aim**: ask about user preferences, recommendations, and needs concerning comfort viewing field and guiding system to the speaker.
7. **Previous testing**: please make sure all the equipment works before conducting the interviews.
8. **Expected length**: 55 minutes per participant.

DURING THE INTERVIEW

1. **Welcome** participant and briefly explain ImAc and the aim of the interview (5 minutes)

2. **Signature**. Consent form and information sheet signature (5 min). Please keep the signed consent forms and/or recordings and hand in to Pilar in next meeting in Munich May 2018.

3. **Questionnaires**. Fill in demographic information: one general questionnaire per participant (5 min).

4. **Start tasks and discussion**, following the structure below. (25 min)

The facilitator explains task 1 and aim of it.

Task 1. Participants are shown the stimuli for task 1 and assess it on a 5-point Likert scale. At the end, qualitative information is requested from participants, using the following suggested questions:

- After having seen all the examples, would you change your rating?
- Which one is your preferred one? Why?

Task 2. Participants are shown the stimuli for task 2 and assess it on a 5-point Likert scale. At the end, qualitative information is requested from participants, using the following suggested questions:

- After having seen all the examples, would you change your rating?
- Which one is your preferred one? Why?

Task 3. Participants are shown the stimuli for task 2 and assess it on a 5-point Likert scale. At the end, qualitative information is requested from participants, using the following suggested questions:

- What did you like about this system of showing the sign language interpreter?
- What did you not like?

5. **Final discussion** (9 minutes)

Please ask user to share their opinions, needs, recommendations, preferences on **all** options shown, and ask them to suggest improvement to current implementation. To lead the discussion, you can ask general preference questions, to be adapted depending on the interview development:

- What are your opinions about these SL possibilities?
- What do you like about each of these SL possibilities?
- What would you change? How can they be improved? Do you have any recommendations?
- What option feels more immersive?
- What option is more comfortable to follow?
- What option is more comfortable to view?
- Are the options intuitive?

6. Farewell and thanks to participants (1 minutes).

AFTER THE INTERVIEWS

Report: upload your report one week after the interviews under Google Drive under WP2/T2.1/ T2.1. pre-pilot tests/Methodology and let Pilar Orero/Anna Matamala know. It must follow the template available in the same folder ("T2.1._FG_It2_report_SL").

Consent forms: provide original consent forms to UAB (next May face-to-face meeting or send by snail mail).

ANNEX X - INSTRUCTIONS FOR AD PRE-PILOTS

T2.1. USER CENTRIC DESIGN: INSTRUCTIONS ON HOW TO DEVELOP FOCUS GROUPS (PRE-PILOTING TESTS)- AUDIO DESCRIPTION

BEFORE THE FOCUS GROUP

1. **Users**: same users who participated in the iteration 1 of the focus groups will be invited to participate. If any of them cannot make it, other users will be invited. Please notice that the recommended number of participants is 5.
2. **Facilitators**: it is recommended that at least 2 facilitators take place in the focus group. One will lead the focus group, the other one will take notes and write down the conclusions.
3. **Room**: make sure the atmosphere is comfortable for the participants, with an adequate room arrangement.
4. **Materials to be ready before the focus group**:
    a. **Stimuli for AD**: three versions of [private_content] have been audio described by RNIB in English, and translated into Catalan by UAB. IRT is taking care of technical implementation.
    b. **Consent form and information sheet**. Available in Google drive under WP2/T2.1/ T2.1. pre-pilot tests/Methodology/Ethical forms. Please notice that this form is slightly different from the one used for the first round of focus groups.
    c. **Demographic questionnaires**. Available in Google drive under WP2/T2.1/ T2.1. pre-pilot tests/Methodology/Questionnaires. Please notice that this form is the same as the one used for the first round of focus groups.
    d. **Devices**: 5 HMD (Google cardboard) + 5 smartphones able to display Youtube + 5 headphones.
5. **Aim**: the focus group aims to ask about user preferences, recommendations, and needs concerning sound in VR AD.
6. **Previous testing**: please make sure all the equipment works before conducting the focus group.
7. **Expected length**: 90 minutes.

DURING THE FOCUS GROUP

1. **Welcome** participants and briefly explain ImAc and the aim of focus groups (10 min)

2. **Signature**. Consent form and information sheet signature, after reading it aloud (10 min). Please keep the signed consent forms and/or recordings and hand in to Pilar in next meeting in Munich May 2018.

3. **Questionnaires**. Fill in demographic information: one general questionnaire per participant (10 min).

4. **Start tasks and discussion**, following the structure below. (45 min altogether)

The facilitator explains three sound possibilities.

> *"We are going to show a 360º musical video with a song in English and with AD (in English/Catalan) in HMD. We have created three different versions, with different types*

of sound, and we would like to have your opinion. You will watch three different versions individually. After each version, we will have a short discussion. At the end, we will again discuss it globally."

Task 1. Please listen to the first version. It is what we call "AD placed on the action" (privilege of sound) (5 minutes). Each participant listens to it on HMD. Discussion 1. (10 minutes): ask users about opinion, preferences, and recommendations (see questions below).

Task 2. Please listen to the second version. It is what we call "AD anchored to head position" (voice of God) (5 minutes). Each participant listens to it on HMD. Discussion 2. (10 minutes): ask users about opinion, preferences, and recommendations (see questions below).

Task 3. Please listen to the third version. It is what we call "AD anchored to soundscape" (first person, past tense) (5 minutes). Each participant listens to it on HMD. Discussion 3. (10 minutes). Ask users about opinion, preferences, and recommendations (see questions below).

Task 4. Final discussion about all modes (15 minutes)

Please ask users to share their opinions on **all** sound possibilities and suggest improvement to current implementation. To lead the discussion, you can ask general preference questions:

- 'What are your opinions about these three sound possibilities?'
- 'What do you like about each of these sound possibilities?'
- 'What is the main potential of each of these three sound possibilities?'
- 'Do you have any suggestions and recommendations about the spatial sound in AD?'
- 'How the spatial sound in AD could be improved?'
- 'What version provides a bigger sense of being immersed in the action? Why'

Make sure you ask about the relationship between script and sound:

- 'Would you agree that sound and script are very much related'

5. Approval of conclusions (10 minutes)

Summarise the results of the report and read them aloud to the participants, who have to approve them. These will be the approved conclusions of the focus group and will need to be included in the report.

6. Farewell and thanks to participants (5 minutes).

AFTER THE FOCUS GROUP

Report: upload your focus group report one week after the focus group under Google Drive under WP2/T2.1/ T2.1. pre-pilot tests/Methodology and let Pilar Orero/Anna Matamala know. It must follow the template available in the same folder ("T2.1._FG_It2_report_AD").

Consent forms: provide original consent forms to UAB (next May face-to-face meeting or send by snail mail).

ANNEX XI - REPORTING TEMPLATE FOR SUBTITLING PRE-PILOTS

1. General information

- Partner responsible for the workshop: *please indicate your acronym.*
- Place and date: *please indicate place and date.*
- Access service(s) discussed: *subtitling.*

2. Participants profile

- Number of end users: *please indicate only number of end users (persons with disabilities).*
- Demographics for users *(please add your numbers where an X is found).*
    12) Sex: X "male", X "female", X "other", X "prefer not to reply".
    13) Age: *please specify the different ages*
    14) Main language of the participants: *please specify the different languages*
    15) Level of finished studies: X "no studies", X "primary education", X "secondary education", X "further education", X "university".
    16) "I define myself as a…" X "blind person", X "low vision person", X "deaf person", X "hearing impaired person", X "deaf-blind person".
    17) Age in which your disability began: X "From birth", X "0-4", X "5-12", X "13-20", X "21-40", X "41-60", X "More than 60".
    18) What technology do you use on a daily basis? You can select more than one. X "TV", X "PC", X "Laptop", X "Mobile phone", X "Tablet".
    19) Do you have any device to access virtual reality content? X "Yes" (*and please indicate which ones they name: XXXX*), X "No", X "I don't know or I don't want to reply".
    20) Which of the following is your preferred device for watching online video content (ie. Youtube, Vimeo, Netflix, Amazon Prime, broadcast catch up service, etc.? You can select more than one: X "PC", X "Laptop", X "Smartphone", X "Tablet", X "I don't watch online video content", X "Others" (If so, please specify)

3. User ratings

Please fill in the table below with user ratings. P1 means participant 1.

Comfort	Level 1	2	3	4	5	6
P1						
P2						
P3						
P4						
P5						
P6						
Mean						

Guiding	Position	Arrow	Compass
P1			
P2			
P3			
P4			
P5			
P6			
Mean			

4. User comments on each presentation mode

Please list user's comments below per element

- **Comfort field of view**

P1:
P2:
P3:
P4:
P5:
P6:

- **Guiding to the speaker**

P1:
P2:
P3:
P4:
P5:
P6:

5. Interviews' conclusions

Please summarise the conclusion of all interviews here.

ANNEX XII - REPORTING TEMPLATE FOR SL PRE-PILOTS

1. General information

- Partner responsible for the interviews: *please indicate your acronym.*
- Place and date: *please indicate place and date.*
- Access service(s) discussed: *sign language.*

2. Participants profile

- Number of end users: *please indicate only number of end users (persons with disabilities).*
- Demographics for users *(please add your numbers where an X is found).*
    21) Sex: X "male", X "female", X "other", X "prefer not to reply".
    22) Age: *please specify the different ages*
    23) Main language of the participants: *please specify the different languages*
    24) Level of finished studies: X "no studies", X "primary education", X "secondary education", X "further education", X "university".
    25) "I define myself as a…" X "blind person", X "low vision person", X "deaf person", X "hearing impaired person", X "deaf-blind person".
    26) Age in which your disability began: X "From birth", X "0-4", X "5-12", X "13-20", X "21-40", X "41-60", X "More than 60".
    27) What technology do you use on a daily basis? You can select more than one. X "TV", X "PC", X "Laptop", X "Mobile phone", X "Tablet".
    28) Do you have any device to access virtual reality content? X "Yes" (*and please indicate which ones they name: XXXX*), X "No", X "I don't know or I don't want to reply".
    29) Which of the following is your preferred device for watching online video content (ie. Youtube, Vimeo, Netflix, Amazon Prime, broadcast catch up service, etc.? You can select more than one: X "PC", X "Laptop", X "Smartphone", X "Tablet", X "I don't watch online video content", X "Others" (If so, please specify)

6. User ratings

Please fill in the table below with user ratings. P1 means participant 1.

Comfort	Level 1	2	3	4	5	6
P1						
P2						
P3						
P4						
P5						
P6						
Mean						

Guiding	Position	Arrow
P1		
P2		
P3		
P4		
P5		
P6		
Mean		

Forced perspective	
P1	
P2	
P3	
P4	
P5	
P6	
Mean	

7. User comments on each presentation mode

Please list user's comments below per element.

- **Comfort field of view**

P1:
P2:
P3:
P4:
P5:
P6:

- **Guiding to the speaker**

P1:
P2:
P3:
P4:
P5:
P6:

- **Forced perspective**

P1:
P2:
P3:
P4:
P5:
P6:

8. Interviews' conclusions

Please summarise the conclusion of all interviews here.

ANNEX XIII - REPORTING TEMPLATE FOR AD PRE-PILOTS

1. Focus group general information

- Partner responsible for the workshop: *please indicate your acronym.*
- Place and date: *please indicate place and date.*
- Access service(s) discussed: *audio description.*

2. Participants profile

- Number of expert end users: *please indicate only number of end users (persons with disabilities).*
- Demographics for users *(please add your numbers where an X is found).*
    - Sex: X "male", X "female", X "other", X "prefer not to reply".
    - Age: *please specify the different ages*
    - Main language of the participants: *please specify the different languages*
    - Level of finished studies: X "no studies", X "primary education", X "secondary education", X "further education", X "university".
    - "I define myself as a…" X "blind person", X "low vision person", X "deaf person", X "hearing impaired person", X "deaf-blind person".
    - Age in which your disability began: X "From birth", X "0-4", X "5-12", X "13-20", X "21-40", X "41-60", X "More than 60".
    - What technology do you use on a daily basis? You can select more than one. X "TV", X "PC", X "Laptop", X "Mobile phone", X "Tablet".
    - Do you have any device to access virtual reality content? X "Yes" (*and please indicate which ones they name: XXXX*), X "No", X "I don't know or I don't want to reply".
    - Which of the following is your preferred device for watching online video content (ie. Youtube, Vimeo, Netflix, Amazon Prime, broadcast catch up service, etc.? You can select more than one: X "PC", X "Laptop", X "Smartphone", X "Tablet", X "I don't watch online video content", X "Others" (If so, please specify)
    - Which of the following do you use on your connected devices to access the above content? X "Magnification (ie. Zoomtext)", X "Screen readers (ie, JAWS, VoiceOver TalkBack)", X "both", X "none".
    - Which of the following controls would you like to use with your screenreader/magnification tool when watching content online? X "Browse content library", X "Identify content", X "Functions as play, stop, pause, forward, rewind", X "Switch AD/AS on and off".

3. Workshop discussion

Please summarise your results here, indicating user preferences regarding the different options that were provided for each access service.

3.1. AD centered in the scene (voice of God)

3.2. AD placed on or in the direction of the action (AD on action)

3.3. AD anchored in the scene (Friend on a sofa)

4. Workshop conclusions

Please write down the approved workshop conclusions here.

ANNEX XIV – ADAPTED QUESTIONNAIRE

1. **Age:** _____

2. **Level of finished studies**
   a) No studies
   b) Primary education
   c) Secondary education
   d) Further education
   e) University

3. **What devices do you use on a daily basis? Multiple replies are possible.**
   a) TV
   b) PC
   c) Laptop
   d) Mobile phone
   e) Tablet
   f) HMD
   g) Game console
   h) Other:_____

4. **How often do you watch virtual reality content (for instance, 360º videos)?**

	Never	Occasionally	At least once a month	At least once a week	Every day
In smartphone					
On a tablet					
On a PC					
In smartphone plugged to HMD					
In HMD					

5. **If you have never used virtual reality content such as 360º videos or only occasionally, please indicate why. Multiple answers are possible.**
   a) Because I am not interested.
   b) Because it is not accessible.
   c) Because I have not had the chance to use it.
   d) Other reasons. Please explain: _____

6. **Please state your level of agreement with the following statement: "I am interested in virtual reality content (such as 360º videos)."**
   a) I strongly agree
   b) I agree
   c) Neither agree nor disagree
   d) Disagree
   e) Strongly disagree

7. **Do you own any device to access virtual reality content?**
   a) Yes (If yes, which one? _____)
   b) No
   c) I don't know or I don't want to reply

8. **Do you like watching the following types of content on television or online?**

	I like it very much	I like it	Neither like it nor dislike it	I don't like it	I don't like it at all
News					
Fiction (series, films)					
Talk shows					
Documentaries					
Sports					
Cartoons					

9. **When subtitling is available, do you activate it for the following type of content?**

	Always	Sometimes	Rarely	Never
News				
Fiction (series, films)				
Talk shows				
Documentaries				
Sports				
Cartoons				

10. **If it is available and you do not activate it, please select the reasons why**
    a) Because the interface is not accessible.
    b) Because I don't want subtitling in all the content, only in certain types of content.
    c) For other reasons. Please explain why: _____.

11. How many hours a day do you watch subtitled content?

a) None
b) Less than 1 hour
c) 1-2 hours
d) 2-3 hours
e) 3-4 hours
f) 4 hours or more

12. What do you use subtitles for?

a) They help me understand
b) They are my only way to have access to the dialogue
c) I use them for language learning
d) Other. Please explain: _____