# D6.2

| | |
|---|---|
| **Version** | 1.10 |
| **Date** | 2015/10/19 |
| **Author** | UAB |
| **Dissemination status** | PU |
| **Document reference** | D6.2 |

# D6.2 – Pilot-D Solution Integration and Trials

| | |
|---|---|
| **Grant Agreement nº:** | 621014 |
| **Project acronym:** | HBB4ALL |
| **Project title:** | Hybrid Broadcast Broadband for All |
| **Funding Scheme:** | CIP-ICT-PSP |
| **Annex I reference version:** | 2013/10/22 |
| **Project Duration:** | 2013/12/01 – 2016/11/30 (36 months) |
| **Coordinator:** | Universitat Autònoma de Barcelona (UAB) |
| **Beneficiaries:** | Rundfunk Berlin-Brandenburg (RBB)<br>Institut fuer Rundfunktechnik GmbH (IRT)<br>Rádio e Televisão de Portugal SA (RTP)<br>Televisió de Catalunya SA (TVC)<br>Schweizerische Teletext (TXT)<br>Vsonix Gmbh (VSX)<br>Fundación Centro de Tecnologías de Interacción Visual y Comunicaciones VICOMTECH (VIC)<br>Screen Subtitling Systems Ltd (SCREEN)<br>Holken Consultants & Partners (HC)<br>People's Playground BV (PPG)<br>Universidad Politécnica de Madrid (UPM) |

This project is supported by funding from the 'Competitiveness and innovation framework programme 20017-2013' of the European Union.

Project no. 621014
**HBB4ALL**
Hybrid Broadcast Broadband for All

CIP- Pilot actions
Competitiveness and innovation framework programme 2007-2013

# D6.2 – Pilot-D Solution Integration and Trials

Due date of deliverable: 2015/07/31
Actual submission date: 2015/10/19

Start date of project: 2013/12/01
Duration: 36 months

Lead beneficiary for this deliverable: Universitat Autònoma de Barcelona (UAB)

Revision: 1.1

| **Project funded by the European Commission within the Seventh Framework Programme** | | |
|---|---|---|
| **Dissemination Level** | | |
| **PU** | Public | **X** |
| **PP** | Restricted to other programme participants (including the Commission Services) | |
| **RE** | Restricted to a group specified by the consortium (including the Commission Services) | |
| **CO** | Confidential, only for members of the consortium (including the Commission Services) | |

## History Chart

| Issue | Date | Changed page(s) | Cause of change | Implemented by |
|---|---|---|---|---|
| 0.1 | 2015/06/24 | - | Document structure | UPM |
| 0.2 | 2015/07/03 | ALL | Integration of contributions | RBB, VIC, UPM |
| 0.3 | 2015/07/08 | ALL | Integration of UAB suggestions and comments | UAB, UPM |
| 0.4 | 2015/07/30 | ALL | Second integration of UAB | UAB, UPM |
| 0.8 | 2015/08/30 | ALL | Integration of contributions | UAB, UPM |
| 0.9 | | ALL | Integration of UAB | UAB, UPM |
| 1.0 | 2015/10/03 | ALL | Prefinal version | UAB |
| 1.1 | 2015/10/19 | ALL | References and format | UAB, UPM |

## Validation

| No. | Action | Beneficiary | Date |
|---|---|---|---|
| 1 | Prepared | UAB | 2015/09/28 |
| 2 | Approved | RTP | 2015/10/01 |
| 3 | Released | UAB | 2015/10/19 |

# Table of Contents

# Figures



# Tables

# Executive Summary

This document describes the progress of HBB4All Work Package 6 / Pilot D throughout the second year of the project execution. Pilot D is concerned with the implementation of sign language interpretation services for audiovisual content, taking advantage of the hybrid system that allows TV and broadband network simultaneous transmission.

The main challenge met in this Pilot is Personalisation. The partners in Work Package 6 have been working on service components that allow the final user to customise the rendering of the sign language interpretation on screen. Pilot-D explores how to design signing services that are customisable for deaf people that use signing language as a means of communication. Pilot-D investigates the technical pathways to offer the personalisation of the signing interpretation service according to the preferences and needs of the final user. In order to achieve this important goal, all services are tested with users by means of contacting deaf people local associations and professional SL interpreters. Thus WP6 partners will accommodate the services to provide a satisfactory experience for final users, which also fulfil the needs of all stakeholders.

As introduced in Deliverable 6.1, during this year two service components have been implemented and tested:

- The Signer HbbTV [3] application (by RBB, Germany), that provides an on-demand signing service;
- An IP/Web based signing pilot (by RTP, Portugal), that uses a website to provide a customisable window for the signing implementation.

Both services allow customisation of Size of SLI subscreen and Position. These functionalities are fully deployed for live contents, but some difficulties have been encountered and identified when applying the service to recorded contents.

Moreover, Pilot D had three additional goals, which have also been addressed in this year of solution integration and trials:

- A new signing workflow model for authoring and distribution of the signing services has been set up and tested.
    - The business approach, developed by Screen, takes into account the organizational processes and professional profiles involved. It has not changed from the previous version for Deliverable 6.1.
    - The technical approach, depicting the functional blocks for the signing service provision, is improved with respect the previous version to satisfy the requirements of subpilot implementations.

- An Avatar based signing component has been developed by Vicomtech. This component is a domain specific spoken language-LSE translation platform that uses a virtual interpreter. UAB has carried out a pretest with final users. According to the results obtained, some modifications have been proposed to improve the avatar's performance.

- The RBB team and the UAB team, with the goal to explore user preferences concerning service provision, have carried out complementary user tests and expert tests. RTP, RBB and UAB team are setting up new tests on the three services to carry out within the next Operational Phase.

In sum, the results obtained in Pilot D after the second year of the project are satisfactory. On the one hand, the services proposed (Signer HbbTV, IP/Web Based signing service) have been defined and the possible difficulties identified, to be solved in the operational phase. On the other hand, the additional goals have been fulfilled: The signing workflow model has been designed and tested, the Avatar based signing component has been implemented and pretested with users, who suggested few improvements, and user pretests and tests have been designed and prepared for next Operational Phase.

# 1. Introduction

In this deliverable we describe the accomplishments reached within Task T6.2 of the HBB4ALL project, named "Solution Integration and Trials". As described in D6.1[1], task 6.2 has consisted in the development of small heuristics trials and lab tests conducted by the different partners. These trials and lab tests involved volunteers**.**

**RTP** and **RBB** have carried on the integration of HbbTV/Web-based applications for hybrid signing services customisable for end users. The most relevant feature of these developments is that end users can choose some parameters of the final rendering such as the signer video size, position or background.

One of the main goals of Pilot D was to design a workflow model to specify the architecture of signing services. A main complexity in this design appears because a signing service is delivered as an open service, and cannot be merged as a component as subtitles can. **Screen** developed a business model diagram that shows the main business objects and relationships involved in the business of producing sign language services for television broadcasts. **UPM** designed a technical generic signing workflow that meets a variety of requirements, to be used by broadcasters for the provision of sign language interpretation.

The components that have been developed and tested in this second year of the HBB4ALL project are the Signer HbbTV application (by **RBB**, Germany), that provides an on-demand signing service; and the IP/Web based signing pilot (by **RTP**, Portugal), that uses a website to provide a customisable window for the signing implementation.

The Signer HbbTV application component offers a pre-mixed picture-in-picture (PiP) video stream consisting in a main TV video and a sign language video area. The application will offer a range of customisation to the final user like signer video size, position or background. The application concept is finished and the tests defined. Technical tests with automatic, command-line-based video mixing have started for the server-side video mixing approach. Experts will be demanded to evaluate the cost efficiency, the picture quality and the affected service performance. In November 2015, RBB will start its tests with a panel of users, using mainly a survey and questionnaires.
The IP/Web based signing pilot developed by **RTP** will offer a double screen signing service for the catch-up content, enabling fully customisation of both screens. **UPM** is providing technical support to develop this option. This component will be tested with users throughout next year.

As an additional goal, Pilot D has worked on an Avatar signing application. The application, developed by **Vicomtech**, make use of five different modules to achieve the virtual interpretation: (1) a text-to-Sign Language translation module, (2) a gesture capture system to create an internal (3) sign dictionary, (4) an animation engine and a (5) rendering module. The application works on the weather forecast domain, but can be adapted to work on other domains. **UAB** has run a pretest on this application and made some suggestions from the obtained results.

Last but not least, **UPM** has developed a complete HbbTV [3] play-out system that is available to all partners in WP6 to run tests. The component developed by **UPM** is able 1) to generate the specific HbbTV signalling, 2) to multiplex the HbbTV-related content with the audiovisual content and 3) to broadcast the resulting stream in lab conditions to test the HbbTV content in actual receivers.

The success of Task 6.2 in putting into place the technology, the infrastructure and the services in place by the partners is the essential pre-requisite for all large-scale trials in the ensuing Operation Phase.

In order to improve the clarity of the exposition, the structure of this Deliverable follows the pattern adopted by all WPs, as will be described next.

## 1.1. Purpose of the document

This report introduces the major achievements accomplished by the WP6/ Pilot D during the second year of the HBB4ALL project. The Chapter 1 of this deliverable is an executive summary. The present chapter introduces information about the structure, terms and definitions of the deliverable.

Next, Chapter 2 presents the workflow signing model deployed for both business and technical approaches. In Chapter 3, RBB informs about the development of the Signer HbbTV application, a component which enables a sign language video playback system on HbbTV devices that can be customised for the parameters "Signer size" and "Signer position". Vicomtech introduces in the second section of Chapter 3 their work on an avatar signing component, which consists in a domain-specific translation platform from text to signing language. Next, a Playout system for tests within the frame of HbbTV applications is presented by UPM. This component generates the specific HbbTV signalling, multiplexes the HbbTV-related content with the audiovisual content, and broadcasts the resulting stream for lab testing. In section 4, RTP describes the double-screen signing service currently provided in their website for live content.

In Chapter 4 the integration of sub-pilots is described. First section presents the architecture of the component developed within the German sub-pilot. The second section describes the component integration for the IP/Web-based signing pilot implemented by RTP and UPM.

Chapter 5 focuses on the service pilots, which are very relevant at this moment because the HBB4ALL Operational Phase is approaching. RBB and RTP/UPM explain how their services are going to be deployed and tested in two different subsections. This chapter also includes a complete description of the avatar signing service developed by VIC due to importance in the project, although it is not being deployed in a service pilot. At the ending section, an overview of common challenges is provided.

In Chapter 6 the UAB team presents the complementary user tests on preferences on different subscreen size parameters by deaf users and experts (interpreters). Finally, this document ends up with a Conclusions chapter.

## 1.2.Acronyms and abbreviations

In this document, when necessary, identified partners within the project are referred to using the abbreviated names initially defined within the Consortium Agreement for HBB4ALL and reproduced on the cover sheet of this document. Moreover, this deliverable uses certain acronyms that are explained in the following section.

## 1.3. Definitions and glossary

**Access Service** – the provision of additional services or enhancements that improve the accessibility of TV services for viewers with disabilities or special needs.

**Accessibility** – The degree to which a product, device, service, or environment is available to as many people as possible. Accessibility can be viewed as the "ability to access" and possible benefit of some system or entity. Accessibility is often used to focus on persons with disabilities or special needs and their right of access to entities, often through use of Assistive technology or Access Services.

**Business model** – describes the rationale of how an organization creates, delivers, and captures value. This may be viewed in a narrow sense (*economic value, what are the costs, and if there are revenue streams to pay for them*). Increasingly, a business model includes social or other forms of value.

**Catch-up TV** – A service that allows a viewer to see a TV program independent of when it was broadcast. This is usually a kind of on-demand service on the Internet.

**CDN** – Content Delivery Network. Content repositories especially built to deliver audiovisual content in an efficient manner.

**DOG - Digital Onscreen Graphic**

**DVB** – Digital Video Broadcasting. European organization that elaborates specifications for digital TV that are used all over the world.

**HbbTV** – Hybrid Broadcast Broadband TV. Standard technology for Connected TV that supports the HBB4ALL project. HbbTV is a major pan-European initiative building on work in the Open IPTV Forum aimed at harmonizing the broadcast and broadband delivery of entertainment to the end consumer through connected TVs and set-top boxes.

**ISL** – International Sign Language

**IPTV** - Internet Protocol Television

**LSE** – Spanish Sign Language. The avatar signing interpreter is based on this sign language and it appear widely referenced in this deliverable.

**MPEG-DASH** or **DASH** – Motion Picture Expert Group – Dynamic Adaptive Streaming over HTTP. This technology is supported by HbbTV from the version 1.5. It allows an adaptive streaming depending on the network capabilities.

**SLI** – Sign Language Interpretation

# 2. General workflow & analysis of system architecture

The specification of a new signing workflow model is one of the objectives of WP6, according to the HBB4ALL DoW. This model must satisfy the requirements of a variety of technological scenarios, including hybrid transmission.

The HBB4ALL signing workflow model is described in this section and it includes two different approaches:

- A business approach, which takes into account the organizational processes and professional profiles;
- A technical approach, which depicts the functional blocks for the signing service provision.

The business approach has not changed with regard to deliverable D6.1 version [1]. However, the technical approach included in this section integrates some improvements to completely satisfy the requirements of the final subpilots implementations.

## 2.1. Business approach for the HBB4ALL signing workflow

The production costs of visual signing are comparable with the cost of providing live captioning. However, currently the greatest commercial challenge to delivering programmes with visual signing on broadcast television networks is the distribution cost. Currently the provision of visual signing (on a broadcast network) for the entire schedule of programmes on a TV channel would incur the significant costs of an additional channel (that carried a version of the programme with open signing - a broadcaster mix). It is considerably more cost-effective to deliver closed signing (as a lower bandwidth overlay), perhaps via broadband, and at the viewer's choice, combine the signing with the broadcast programme in the viewer's receiver (e.g. using HbbTV).

The business models for access services involve two components: a business model for the service provision; and a business model for any device needed by the viewer to access the service. For signing services there is currently no specialised device – as signing services are delivered as open services, either as a separate channel or more commonly as specially scheduled and produced programmes. The main current business models for signing service provision are: production budget funding; private sponsorship; and public funding.

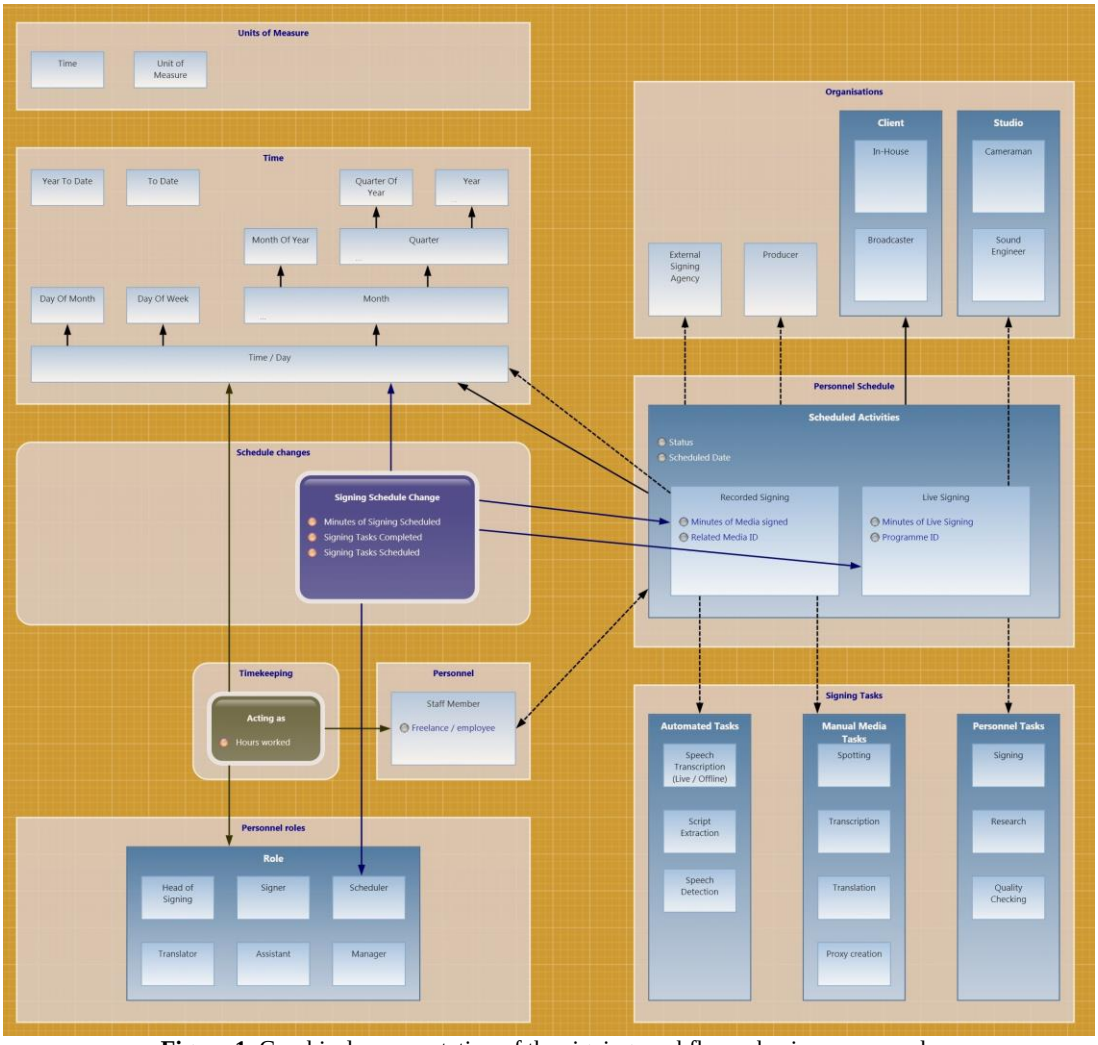


**Figure 1.** Graphical representation of the signing workflow – business approach.

The move to digital distribution and the move from broadcast to Internet-based distribution have major impacts for the provision of signing services. In addition, the rise in popularity of 'Over-the-top' delivery of television programmes (e.g. catch up and Video on Demand) has led to a growth in the number of receivers that are connected to both a broadcast signal and a broadband connection. In the future it is anticipated that a transition to closed signing services will leverage a 'universal design' approach (i.e. access services will take advantage of the capabilities that will exist in mainstream consumer electronics), thus stepping past the often encountered, "chicken and egg", situation; where provision does not occur because specialised access devices are not widely deployed with the target audience.

The business case for commercial broadcasters funding signing access services is not very compelling and the commercial case is not usually perceived as being strong where there are no regulatory requirements. However the use of hybrid delivery and mixing the signing with the programme in the viewer's device also has implications for regulators responsible for frequency and bandwidth allocation. Typically, current visual signing services are bandwidth-hungry, if signing is delivered as an additional channel rather than being

offered as an open service seen by all viewers. A closed signing service using hybrid delivery has far lower (or near zero) broadcast bandwidth requirements.

Utilising a hybrid distribution mechanism and capitalising on universal design principles in the viewer's device does at least remove two significant barriers to increased signing service provision: the cost of distribution; and the cost of deploying specialised receiver devices.

Screen have produced a formal 'business model diagram' and documentation for signing production based on information for literature reviews, informal interviews with signing professionals and previous business experience. The business model diagram developed shows the main business objects and relationships involved in the business of producing sign language services for television broadcasts. The model has been tested against the workflows of the WP6 partners and has been found to cover the principal aspects of signing for both live and offline production.

The main diagram in the model uses the DIW methodology developed by Kalido, a free tool that is available for the creation and viewing of this type of diagram. A document has been produced that describes the top-level diagram and the business objects and relationships that the diagram portrays. The diagram does not show all the details of a practical or real business, but establishes a core vocabulary to identify and describe the principal objects and relationships in signing production.

The core business objects in the model are the scheduled activities that essentially describe the activities of the company and the processes that are active at specific points in time. The model expresses certain 'business rules' about the business process (e.g. you cannot 'schedule' an activity without choosing a date). The transactions identified within the model represent events that change information within the model, a transaction connects the entities that are responsible for change and the entities that are changed within the model.

The model defines a group of business entities that are external to the signing production; Studio (sub-divided into two subclasses, Cameraman or Sound Engineer), Client, Producer and External Signing Agency[1].

- Producer represents an entity that influences or controls signing activity.
- Client is an entity that commissions and receives / uses the output.
- Studio is a business entity used for a 'live' signing service. The model shows there is a necessary interaction between the company and the Studio (e.g. for reception of audio and video feeds of a broadcast that is signed while it is transmitted live).

In addition, the model identifies the personnel and the roles undertaken by personnel during signing activities. The model makes a distinction between a role and a signing task, in that a role is the 'capability to perform' a specific signing task. This allows the model to support staff members who change role periodically (e.g. depending upon the signing task that needs doing). The model also specifically separates translation (from a foreign language) and signing (conversion to sign language), as only certain personnel may be capable of both language translation (for example for foreign language programming) and signing.

[1] An External Subtitling Agency models a business entity that may undertake signing activities, but is external to the modelled business. This models activities that may actually be performed by external resources.

The signing tasks identified by the model are split into two categories, those that are performed by staff members and those tasks that may be automatically performed by a system. The model also identifies which tasks might apply to recorded or live signing activities. Identified staff member tasks in the model include:

- Proxy Creation – making a low bitrate media asset to facilitate offline signing production.
- Quality Checking - checking offline signing against a pre-determined set of criteria.
- Research - collecting information related to a broadcast to facilitate signing.
- Signing - The act of signing the equivalent of heard audio.
- Spotting - offline determination of signing timing (i.e. when signing needs to occur).
- Transcription - offline conversion of speech in a broadcast into text.
- Translation - offline conversion of 'text' in one language into 'text' in another language.

The model anticipates that in real practise, signing typically requires a combination of these tasks performed by a single staff member, or a group, in sequence or in parallel. For example 'offline signing' may involve Proxy Creation, Research, Spotting, Transcription, *Translation*[2], Signing and Quality Checking.

Identified automated tasks in the model include:

- Script Extraction - automated extraction of dialog text from a script file.
- Speech Transcription - automated conversion of speech from the audio into text.
- Speech Detection - automated detection of speech within the audio to list speech events.

## 2.2. Technical approach for the HBB4ALL signing workflow

This technical generic signing workflow, shown in Figure 2, addresses the following objectives:

- It is flexible enough to be used in the variety of scenarios and particular implementations that have been considered in the project and specially the implementations that have been selected for the HBB4ALL operational phase, as described in this deliverable;
- It is consistent and compatible with the workflows formerly and currently deployed for the signing service in WP6 broadcasters (RTP and RBB);
- It is generic enough to be used by other broadcasters in the future for the provision of sign language interpretation.

[2] For a foreign language programme.

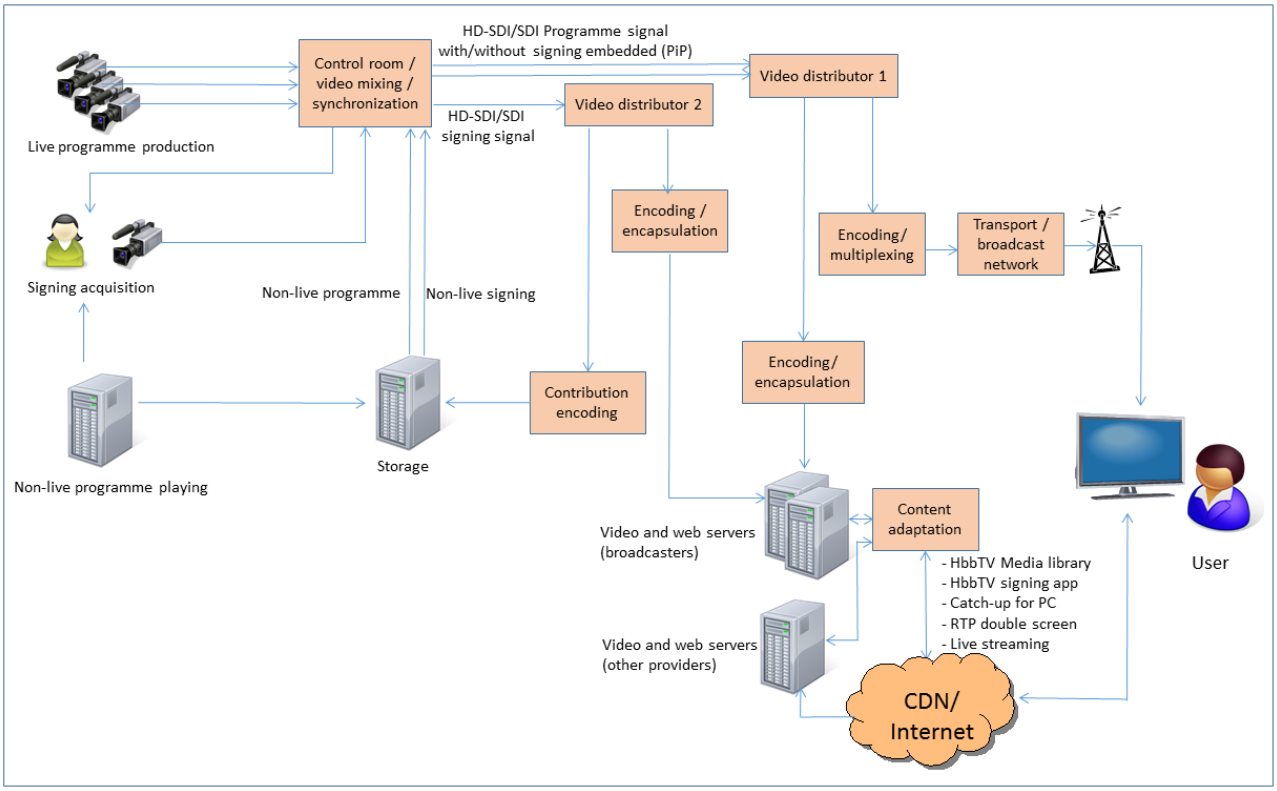


**Figure 2.** Signing workflow – technical approach.

This model tries to fulfil all the possible requirements that a broadcaster could need. It represents an abstract value chain from the production of the signing service to the end user. The requirements this model satisfies are:

- It can be used for signing in live and non-live programmes. As shown on the left side, the main video may come from a production environment or from a hard drive or a tape.
- It can be used for live or non-live signing. It must be taken into account that a possible scenario consists in live signing for non-live content. If the produced signing signal is not live-broadcasted, it is stored (storage equipment in the figure) and linked to the respective non-live content. Both contents are available to be inserted again in the distribution workflow, as shown by both parallel arrows in the figure (labelled as "non-live programme" and "non-live signing").
- The signing interpretation signal can be embedded or not embedded in the main video. If the signing interpretation signal is embedded, this processing would be made in the "video mixing" phase. In any case, the workflow preserves the signing signal and the programme signal as they are.
- It allows to record the signing interpretation and to link it to the main (programme) video (storage box).
- It allows the signing service to be delivered by means of the broadcast or broadband networks. On the right side, a hybrid terminal and an end user are depicted.
- It allows the main video to be delivered by means of both networks, too.
- It allows a variety of signing services to be delivered by means of the IP/broadband network:

- “Truly hybrid” HbbTV approach, involving an HbbTV application able to join two video signals coming from two different networks.
- HbbTV catch-up TV applications, including sign language interpretation. For this purpose, the signing content would be delivered by means of the "video and web servers" in the figure.
- HbbTV live streaming applications for signing, as developed in the German subpilot of WP6.
- Live streaming for PC and IT devices.
- PC applications that integrate both main video and signing interpretation (as the currently deployed RTP double-screen service).
- Catch-up TV application including signing for PC and IT devices, as designed in the Portuguese subpilot of WP6, even if the service is not provided by the broadcaster.
- Use of a second screen for the sign language. The second screen paradigm has arisen during the last years as a way to provide additional contents, including multimedia. Some proposals suggest the use of this second screen to present signing. For this reason, it has been included in the workflow as a possible implementation.

- The model represents the possible user interaction via the broadband channel (Internet), as a return channel. For this purpose, double arrows are depicted in the figure.
- It allows a third entity to provide the signing service via Internet. This is a powerful idea for future implementation. In HBB4ALL, UPM is going to provide the signing service for catch-up content in the Portuguese subpilot. For this aim, the signing and programme signals are recorded or linked in the provider servers, which are labelled in the figure as "Video and web servers (other providers)".

The "video distributor" boxes in the figure provide output signals from video inputs. The "video distributor 1" allows delivering the programme signals (including or not the signing window) through both networks. In the particular case of the German subpilot, this block provided two signals: the programme video without signing to be delivered by means of the broadcast network and the programme video with embedded signing (or a variety of mix options) to be delivered via Internet.

The "video distributor 2" block enables both the storage of a produced signing signal and the availability of the sign language service in the web servers or the broadcasters. This feature is currently used by RTP to provide the double-screen service in its website.

Some blocks are labelled as "encoding" in the diagram. The aim of this block is to encode the video streams to be stored or transmitted. The processing is different in each case (e.g., it depends of possible compression standards or the desired visual quality).

The first version of the workflow model fulfilled all the requirements identified in the HBB4ALL first year and it was included in deliverable D6.1 [1]. This second version is based on the previous one, which has been refined to take into account the design of final implementations for the Operational Phase.

# 3. Service Components

## 3.1. Signer HbbTV application

The Signer HbbTV application is being developed by RBB.

### *3.1.1. Overview*

The Signer HbbTV application component enables a sign language video playback system on HbbTV devices. Within an HbbTV application a pre-mixed picture-in-picture video stream will be offered, that consists of a main TV video and a sign language video area. The aim is to provide users from the target group with sign language video adaptation functionalities, enabled also for single-video decoding devices. The application will offer the user a number of parameters for adaptation, such as signer video size, position or background. All offered parameter combinations will have been pre-mixed and made available as dedicated variants of the said PiP video. The playback system receives its input from a video resource where the pre-mixed video variants are located. Users can launch the signer application from a generic HbbTV launcher bar, which is the standard entry point to HbbTV services used by all German broadcasters. MP4 and MPEG DASH video formats are considered for this prototypical implementation.

### *3.1.2. Status of completion*

RBB has finished the Signer HbbTV application concept. The application will load all available information on how the PiP video can seemingly be adapted from a dedicated configuration file. That file defines the adaptation categories and their included options, e.g.:

*"signer size":* *very large* | *large* | *medium* | *small*

*"signer position":* *left* | *middle* | *right*

The combination of the parameters and options are integrated into the file name declaration and thus define the video file name to be addressed inside the HbbTV application, e.g.:

*HAlign_left_SignerSize_large.mpd*

RBB has also completed the definition of the video post-production. RBB will aim for a "cutting room" approach, where the main video and the pre-recorded sign language interpretation video will be mixed into a PiP video. Each combination of the video parameters will add a new video variant, exactly matching the adaptation options provided in the HbbTV application. An additional, more experimental approach is the server-side video composition, which is described in section 6.2.

### *3.1.3. Relation to the sub-pilots*

The Signer HbbTV application component will be integrated in the German sub-pilot, as described in section 4.1.

## 3.2. Avatar based signing component

Vicomtech is developing this component service.

### *3.2.1. Overview*

The objective of this component is to provide a domain specific spoken language-LSE translation platform. Its functionality could be easily adapted to other domains in further developments. The application was built and tested on the weather domain. This domain was chosen for (1) using a relatively small and predictable vocabulary, (2) having just one speaker and (3) showing graphic help such as weather maps as a cue for potential mistranslation cases. The application introduces one novelty: the avatar processes the hand-gesture and the bodily expression separately according to the required emotion.

The application consists of five different modules: (1) a text-to-Sign Language translation module, (2) a gesture capture system to create an internal (3) sign dictionary, (4) an animation engine and a (5) rendering module.

#### *3.2.1.1. Text-to-Sign Language Translation Module*

The text-to-sign language translation module is based on rules that were designed taking into account a corpus from the chosen application domain. To do so, a code-system was created in order to represent LSE signs in written strings. Each concept that in LSE has a fixed sign has its corresponding tag in our written representation of LSE. The output of this module is a sequence of signs that strictly follow the LSE grammar that is used by the animation engine.

#### *3.2.1.2. Capture System Module*

In order to translate Spanish into LSE, we had to compile a LSE database. To do so we developed a capture system combining two different motion capture approaches. It uses non-invasive capturing methods and allows entering more sign entries easily. The system can be used by any person but only one person can use the system in each capture session.

In order to capture hands movements, two motion capture CyberGloveII gloves were used, one for each hand. These gloves allowed tracking precise movements of both hand and fingers. In addition, they can connect to the server via Bluetooth, which allows more comfortable and free movements when signing.

Body movements have also a great significance in LSE. In order to capture the movements of the whole body, the Organic Motion system was used. This system uses several 2D cameras to track movements. The images are processed to obtain control points that are triangulated to track the position of the person that is using the system. Thanks to this system, the person signing does not have to wear any kind of sensors, allowing total freedom of movement. The captured movements result more natural and realistic. The person signing had to wear the gloves while standing inside the Organic Motion System at the same time.

In order to join the animations captured with both systems it was necessary to join and process the captions before saving them as whole signs. Autodesk Motion Builder was used for that purpose. This software is useful to capture 3D models in real time and it allows creating, editing and reproducing complex animations.

Once the realistic animations were obtained, they were stored in a database to feed the application with vocabulary in Sign Language.

### 3.2.1.3. Sign dictionary

The sign or gesture dictionary contains the words used in the code given to each concept linked to the actual gesture that the avatar has to interpret. The gesture dictionary is composed by a finite number of lemmatized concepts gathered from the domain specific corpus. Furthermore, all synonyms are gathered within the same entry. The sign dictionary can contain three types of entries:

- ***One-to-One Concepts***: concepts that match a word-token in Spanish and that are expressed in one sign in LSE. Synonyms are listed under the same LSE sign.
- ***Grammatical or void Words***: these entries are listed in the dictionary as evidence of processing, but are linked to an empty concept. They do not trigger any kind of movement because in LSE they do not exist.
- ***Multi-word Concepts***: some concepts may map to more than one word-token in Spanish. These concepts are registered as one entry in the gesture dictionary and they map to just one concept in LSE.

The compiled sign dictionary contains 472 lemmas. These entries have proved to be enough to translate the domain-specific corpus used to extract the translation rules. All these concepts or lemmas were captured with Capture System Module and added to the sign dictionary so that they could be interpreted by the virtual interpreter.

### 3.2.1.4. Animation engine

The Animation Engine was developed with the aim of providing natural transitions between signs as well as modifying the execution of the signs depending on the emotion of the virtual interpreter. Emotion is essential in LSE. Each sign should be represented using not only the hands and the face, but at least, also the upper body of the interpreter. It is based on executing the corresponding sign and changing the speed of the animation depending on the emotion that the virtual interpreter has to reproduce according to the real input at that moment.

The Animation Engine runs as follows: the appearance of the virtual interpreter is loaded from the Virtual Character database. While the virtual interpreter does not receive any input it has a natural behaviour, involving blinking, looking sideways, changing the weight of the body between both feet, crossing arms, etc. When the Text to Sign Language module sends the translation to the Avatar Engine module, it stops the natural behaviour (except blinking) and starts the sequence of signs. If any emotion or mood cue is registered as input, the speed of the animation changes accordingly; for example, it slows down if sad, speeds up if angry. Additionally, the virtual interpreter's expression is also modified using morphing techniques.

The Animation Engine module was developed using Open Scene Graph. It applies any sign animations stored in Sign Language database captured with the Capture System Module previously. In order to concatenate several animations and to obtain realistic movements, a short transition between the original signs is introduced. Thus, the final result is the virtual interpreter signing with very realistic movements.

### 3.2.1.5. Rendering

The objective of this module is to visualize the virtual interpreter synchronized with the multimedia content. For the current prototype, this module inserts the avatar in the broadcasted TV show. For synchronizing the

virtual interpreter with the visual content, the system takes into account the time in which each sentence is pronounced in the audio.


**Figure 3.** Image of the avatar' video for evaluation.

### *3.2.2. Status of completion*

The avatar based signing component is a domain specific spoken language-LSE translation platform. Its functionality is based on the weather domain but it can be easily adapted to other domains. In order to do so, it is necessary:

- Revise and increase the designed rules
- Capture new vocabulary
- Increase the vocabulary

The animation engine and render modules are finished and are independent of the domain.

### *3.2.3. Relation to the sub-pilots*

This component will be used in the avatar signing application. The component has the all functionality of the application. As it was mentioned before it is a domain specific spoken language-LSE translation platform where the translation is performed by a virtual interpreter.

## 3.3. HbbTV play-out system for tests

### *3.3.1. Overview*

UPM has a wide experience in the development of HbbTV applications and play-out systems. As a result of this experience, UPM has developed a complete HbbTV play-out system that has been available in the WP6 for tests. The introduction of an HbbTV application in a digital television stream is based on two prerequisites:

The generation of specific signalling, according to HbbTV and DVB norms. In fact, the DVB norm [4] specified by HbbTV is conceived for application signalling and carriage in any connected TV system - not just HbbTV. This signalling includes the URL where the hybrid terminal will find the complete HbbTV content.

The generation of a DSM-CC object carousel to include application data in the broadcast transport stream.

The results of these processes are multiplexed with the rest of audio, video and data components. The hybrid terminal can, based on the information obtained via broadcast, retrieve the complete application data via the broadband network.

The component developed by UPM is able 1) to generate the specific HbbTV signalling, 2) to multiplex the HbbTV-related content with the audiovisual content and 3) to broadcast the resulting stream in lab conditions to test the HbbTV content in actual receivers.

### *3.3.2. Status of completion*

This component is completely developed. Figure 4 shows the main graphical interface of the component that enables the configuration of a complete MPEG-2 transport stream and the addition of HbbTV applications.

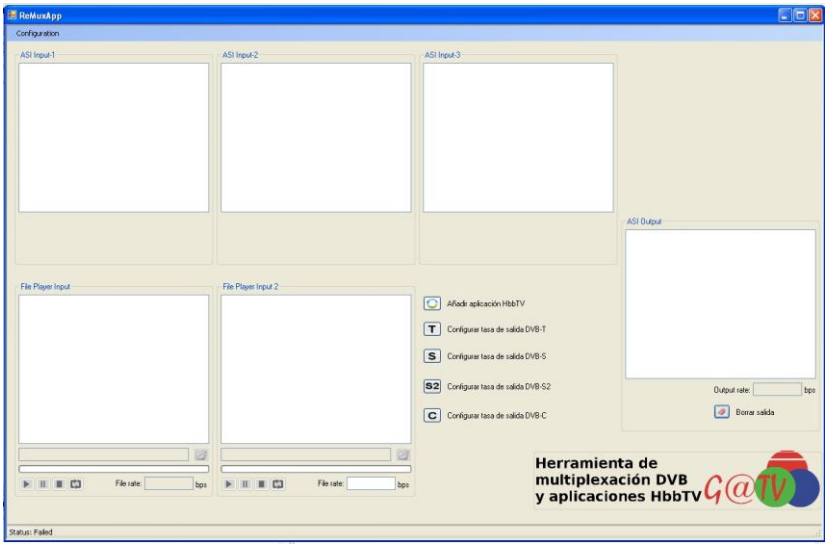


**Figure 4.** HbbTV play-out component developed by UPM

### 3.3.1. Relation to the sub-pilots

The playout system and the involved devices have been resources available for the rest of partners in the project in this period. They have been used to play HbbTV demo content in meetings between UAB and UPM to specify the audiovisual contents for user tests, as describe in section 6.

## 3.4. Double-screen signing interface

### 3.4.1. Overview

RTP is currently providing the double-screen signing service in its website but just for live content. The main objective of the IP/Web-based signing subpilot is the extension of these functionalities to recorded contents.

As shown in the Figure 5, the double-screen functionality provides a large window of the signing interpretation to optimize the intelligibility of the service. The signing window is also embedded in a smaller size on the screen for the Web live delivery (as shown in the left side of the figure) and the broadcast programme.


**Figure 5.** Current double screen signing service for live contents.

However, when the previously broadcasted programmes are available in the catch-up web application, this additional sign language window cannot be displayed, as shown in the Figure 6.


**Figure 6.** Sign language service is not available for the double screen functionality in catch-up content.

The objective of this component is to store the sign language interpretation signal and:

- To provide it for catch-up content, enabling the double screen functionality
- To independently manage the sign language signal, enabling customisation and new services like sign language interpretation in second screens (i.e., a personal additional hardware screen).

### *3.4.2. Status of completion*

This component is currently being developed. Although RTP was going to provide the storing capabilities to record the sign language video, finally the required equipment is not available. For this reason, UPM has worked on recording on the fly the signing stream for live content to include it in a web-based catch-up application. The programme content (catch-up) is available by means of an API created by RTP. This API also offers metadata about the content. The next step in the development will be the creation the web interface to integrate both signals in the double screen approach. This interface will be based on web technologies and hosted by UPM. This development will be finished in October 2015.

### *3.4.3. Relation to the sub-pilots*

This component will be deployed in the IP/web-based signing application (Portuguese sub-pilot).

# 4. Sub-Pilot integration

## 4.1. HbbTV-based signing pilot

Figure 7 describes the workflow that integrates the different components of the German sub-pilot.

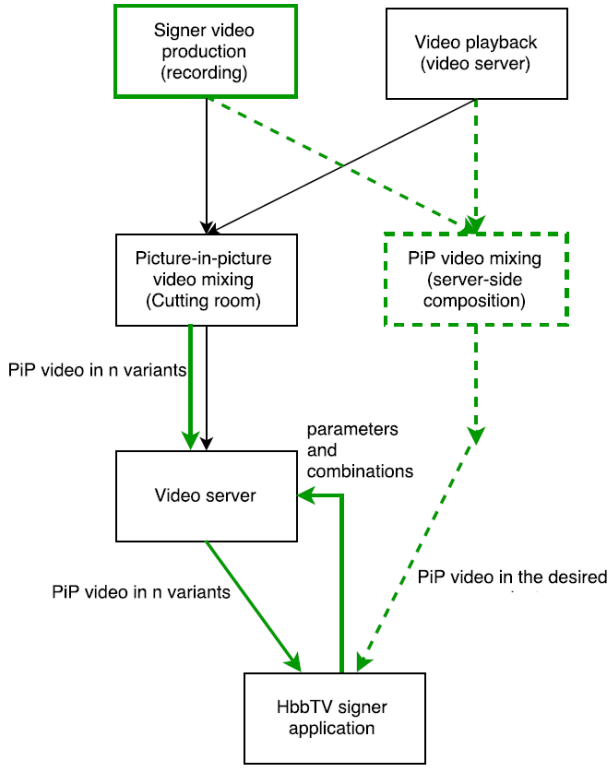


**Figure 7.** Complete workflow for the German sub-pilot.

### *4.1.1. Production*

For the German sub-pilot RBB has already identified the specific programme types that we find suitable for a sign language translation on demand: news, documentary and sports. RBB has created a list with the desired programmes, which will be the main input for the sign language translation and HbbTV service. In a TV studio RBB will record a sign language interpreter who is subtitling a played video from the list of desired programmes. For each of the programmes the recorded sign language video will then be mixed together with the main programme video as PiP video in a cutting room. According to the pre-defined PiP composition specifications a number of different video variants will be produced, different in signing video size, its position and background configuration.

RBB is considering adding a server-side video composition module to the workflow (see section 6.2) that allows for an automatic, command-line-based video mixing and would be triggered by the settings done in the HbbTV application. This module is in early stages and would simplify the overall workflow and the video mixing.

### 4.1.2. Publishing & Delivery

The mixed video variants for all the programme types will be transferred to a simple video server for storage and delivery through the HbbTV application.

The possible server-side composition would eliminate this step, as the PiP would be mixed and instantly transferred to the application in the cloud and on user demand.

### 4.1.3. Reception & User Application

The HbbTV application will allow for the presentation of the sign language service. The user will only see one video variant of the desired programmes. Dependent on the settings chosen by the user, a different video variant will be called through the application in the background, starting at exact the playing time, the former variant was left playing.

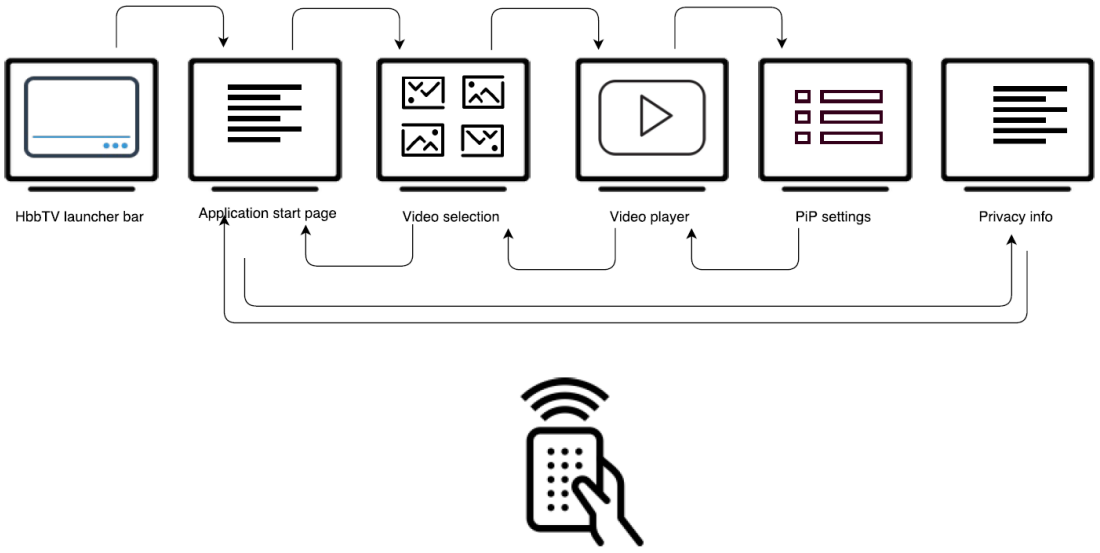


**Figure 8.** Application description.

The application framework itself can be seen as a GUI component that can be controlled with the help of configuration files. In the configuration files can be defined the title of the start screen, the number of programme videos and the desired parameters for the PiP composition. Only defined parameters will be present in the settings GUI for a (virtual) PiP adaptation by the user. This allows for a fine-grained realisation of the test plan, especially the incremental feature testing described in section 5.1.3. The chosen parameters and their values define also the file name of the video variant that matches the adaptation desire of the user.

## 4.2. IP/Web-based signing pilot

Figure 9 shows the complete workflow for providing the IP/Web-based signing pilot. It is a particular implementation of the generic workflow model described in section 2.2 and it takes into account the current signing workflow in RTP.

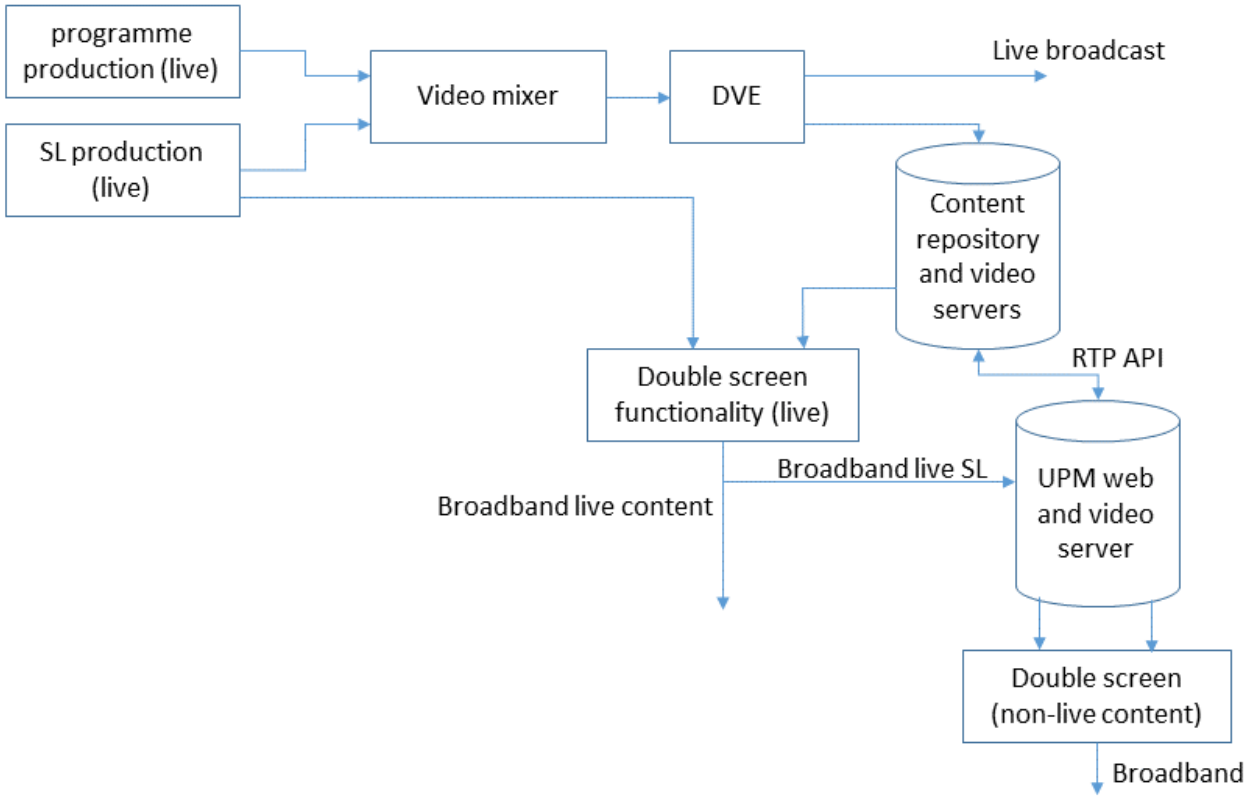


**Figure 9.** Complete workflow for the Portuguese sub-pilot.

### *4.2.1. Production*

Since RTP is currently producing the signing service for live content, the web-based signing component does not require additional production work.

On the other hand, UPM is in charge of storing the signing video clips. This does not affect the broadcaster production workflow but UPM is using additional equipment. The signing video acquisition takes advantage of the HLS stream provided by RTP for the double screen functionality in the web site.

### *4.2.2. Publishing & Delivery*

The signing service will be available by means of a web server provided by UPM. This server will store the signing video clips (acquired from the RTP live signing streaming for live content) and will link the RTP contents. For this purpose, RTP has provided a REST API, which will be used to automatically access the

multimedia contents via a web interface (as explained in the next subsection). Due to record capabilities, the service pilot will include a limited number of TV programmes.

### *4.2.3. Reception & User Application*

The reception / user application is a key part of the component since it consists in the interface to provide the service. The application will be based on web technologies to integrate the new double screen functionality in web browsers. Moreover, this activity has a wide focus and it is also considering a variety of devices to receive and play the signing service. Some tests have been carried out regarding this web-based interface, as shown in chapter 5.

# 5. Service pilots

This chapter describes the service pilots that WP6 is going to deploy in the HBB4ALL operational phase: HbbTV-bases signing pilot (RBB) and IP/Web-based signing pilot (RTP+UPM). Although the avatar signing application is not a service pilot, it is included in this chapter due to its importance in the project.

## 5.1. HbbTV-based signing pilot

### *5.1.1. Status of work*

According to set-up described in section 4.1, the HbbTV sign language application workflow is currently pre-defined. This includes the desired content categories to be sign language translated, the video production and mixing, the video provision and delivery, as well as the concept for the application, which will be forwarded to RBB's HbbTV developers. Additionally, technical tests with automatic, command-line-based video mixing have started for the server-side video mixing approach described in section 6.2.1.

### *5.1.2. Timeframe*

RBB plans to have the HbbTV development ready by M22 in September 2015. Also then RBB will have completed a first "dry run" of the production workflow, including recording, mixing and publication.

In M24, November 2015, RBB will start its dedicated tests with a panel of users. The users will be requested to test the service for a period of approximately 12 weeks. At the end of the pilot run time RBB will considering the addition of the server-based composition approach to the overall sign language workflow.

### *5.1.3. Test users and evaluation approach*

Early research on the ergonomics of accessibility services in digital television was conducted in the EU-funded project DTV4All. A clear result was that there is no single configuration that fits all viewers in all viewing contexts.

The German pilot can potentially target ~6 million households with HbbTV-enabled devices in Berlin-Brandenburg area and beyond in Germany for a quantitative analysis. For a qualitative validation RBB will again work closely with organisations for the deaf in Berlin and Brandenburg. Up to 30 users from the target group will be chosen for an evaluation of the pilot. It is required that all the test users must have an HbbTV-enabled device ready for the pilot phase. RBB aims at organising a closing workshop at the end of the Operational Phase with all the involved users.

While HbbTV services theoretically could provide free customisation of any aspect of the presentation of accessibility services, unlimited choice has a negative impact on usability. Also, certain limitations in implementation call for the identification of a relevant set of configurations that cover all user groups and the majority of viewing contexts, especially with regard to programme genres.

Therefore, RBB has prepared a three-fold qualitative evaluation approach, utilising a survey-assisted incremental feature testing complemented by a System Usability Scale (SUS) and a visual logging process at the end.

A. Incremental feature testing, survey-assisted. In a two weekly rhythm (tbd) the participants are offered choices on particular features, i.e. starting with signer video size, then signer position, then both. In a survey they will express their preferences in relation to different viewing contexts. This will be complemented by actual usage statistics of particular options.

B. The SUS questionnaire will be provided to the test users at the beginning and end of the pilot phase. It comprises of ten standard assertions to be rated, targeting the usability of the application

C. The participants will get a form with key screenshots of the application every week, where they can mark and comment any issues during the pilot period. As a large part of the target group communicates by fax, these forms can be sent for support requests during the pilot or compiled at the end of the pilot period. Based on the visual issue reports the RBB team will be able to later prepare new solution prototypes and paper prototype building blocks, addressing prevailing issues from the pilot.

In addition to the end user tests RBB will conduct a series of interviews with the editors and managers involved in online and the sign language production and provisioning process to gather feedback on aspects of the workflow and costs involved.

### *5.1.4. Challenges, problems and potential solutions*

Although the service is not yet running, RBB would expect only a limited number of possible technical problems. These relate primarily to the potential addition of the server-side composition and the inherent dependency of user-chosen options from the application and the forwarding of these values to the video-mixing server. This would influence the video mixing and therefore the overall availability of PiP sign-language videos.

Other foreseen problems are more of an organizational nature. Our aim is to recruit a demographically representative sample of testers. However as the testers need to have an HbbTV device this may prove difficult. If this is the case we will look into the option of providing testers with equipment and adjusting the sample group demographics.

A further potential risk is that the testers do not understand the application or the methodology. For both we are being advised by local user associations thus minimizing this risk.

## 5.2. IP/Web-based signing pilot

### *5.2.1. Status of work*

Currently UPM is working on the automatic recording of RTP contents from the live signing stream that RTP is producing for the double screen web functionality in live programmes. Moreover, UPM is working on the design of the web interface that will offer the double screen feature for catch-up content. For this purpose, the web interface will integrate two media players, which will be configurable to provide a more satisfying service. In this way, user will be able to choose the size of both player windows inside the web interface.

At this moment, UPM has performed some preliminary tests to check the setting options in the size of both windows, as shown in the figure. Figure 10 shows how the signing video could be larger than the programme one according to the user preferences. Below each player, a button enables the enlargement of the respective

video signal. In this way, the available area is distributed among both signals (content shown in Figure 10 does not belong to RTP, it is another piece of test content available in the project).

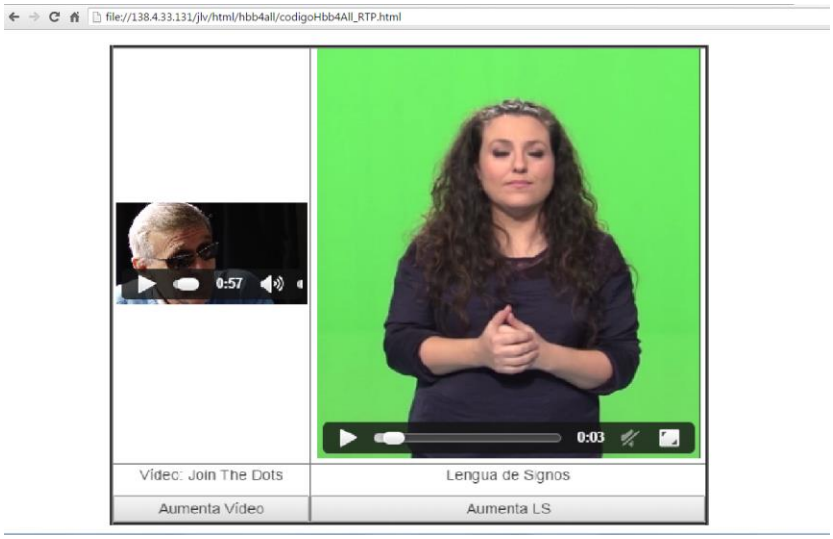


**Figure 10.** Preliminary interface implemented by UPM for the IP/Web-based signing pilot.

### *5.2.2. Timeframe*

This is the current timeframe for the deployment of the service pilot:

- August 2015: UPM ends the automatic recording of RTP signing stream
- September 2015 - October 2015: Web development to host both players
- November 2015: Service pilot available, for a certain sets of programmes

### *5.2.3. Test users and evaluation approach*

RTP is planning to contact the Portuguese users associations (FPAS – Federação Portuguesa de Associações de Surdos [Portuguese Deaf and Hard-of-hearing Association]) to set up a meeting to define a plan and carry out a model of evaluation based on testing the Portuguese sub-pilot in a multi-platform mode;

These tests should consider multi-platform devices, such as desktop, smartphone and tablet;

Since the Portuguese sub-pilot is based on web service, the main goal is to evaluate the usability of the pilot. The SUS , presented by UAB, will be translated into Portuguese and should be used as presented, without any major changes, because the user will be manipulating and will be interacting with their own device during the test.

### 5.2.4. Challenges, problems and potential solutions

The first problem that has been faced in this service pilot has been the availability of record and storage capabilities to preserve the sign language signal since RTP cannot use professional equipment for this purpose. To solve this problem, UPM will bring the required recording capabilities. This fact has enriched the generic workflow model to preview that an alternative organization (i.e., other than the broadcaster) can provide a signing service via broadband channel. The access to the signing full-resolution signal is done via Internet for live content. It must be taken into account that RTP is currently providing the double-screen service for live content. UPM takes advantage of the correspondent video stream to store the signing signal.

However this implementation causes a restriction: just a selection of signing programmes is affordable inside the HBB4ALL project. For this purpose, the selected programme is "Jornal da Tarde".

Regarding the main video signal, it is available by means of an API created by RTP to provide multimedia content and metadata in web applications. This REST API is also used in the Portuguese subtitling service pilot of WP3. On the other hand, a possible problem has been identified: RTP splits large programmes in several video clips for web distribution. This fact will be taken into account in the implementation of the double screen service for non-live content that will be deployed in this sub-pilot, in order to offer a seamless and synchronised experience for both video streams.

Since this double-screen signing service will be supported by web browsers for PC, the compatibility with different browsers could cause a problem. This issue has been identified. To solve it, the implementation will take into account the most extended browsers: Mozilla/Firefox, Chrome, Safari and MS Internet Explorer.

Finally, this signing service will not be integrated in RTP website. This causes a potential risk since few users might take advantage of the service. To solve it, RTP might link the service in the website. Moreover, RTP will take advantage of its contact with deaf people Portuguese associations to promote the use of the service.

## 5.3. Avatar signing application

### 5.3.1. Status of work

The avatar signing application is a domain specific spoken language-LSE translation platform. Its functionality is based on the weather domain but it can be easily adapted to other domains. The translation is performed by a virtual interpreter that is in charge of making the necessary signs.

The avatar signing application is finished in the weather domain. As it was mentioned it can be adapted to other domain but it is necessary to increase the rules as well as the vocabulary and dictionary.

#### 5.3.1.1. Test users and evaluation approach

The avatar signing application was tested in order to gather user feedback. Two native Spanish Sign Language (LSE) users tested the application. The first user is a deaf male from Madrid in his 40s. The second user is a hearing CODA (Child of Deaf Adults) from Granada in her 40s. Participants were tested individually in a quiet room at the School of SLI in Barcelona.

During the interviews the avatar clip was presented to the participants on an 11-inch computer screen. Participants were placed in front of the screen and could manipulate the clip, so that they could stop, replay

or forward it as they wished. The two interviews were held separately and there was no group discussion afterward.

The interviewer was a bimodal bilingual so that the interview with the deaf user was held in Catalan sign language and the interview with the CODA in Catalan. They were asked in order to gather different information, especially from different linguistic levels:

- About communication, discourse and syntax:

The signing avatar helps me understand the weather forecast. (Communicative level)

The signing avatar shows coherence and cohesion. (Discourse level)

The use of space, both syntactic and topographical, is mostly comprehensible. (Discourse and syntactic level)

The sentence structure the avatar signs is mostly comprehensible (Syntactic level)

- About the lexicon:

The avatar vocabulary is native LSE signs. (Lexical level)

The avatar weather-related vocabulary is native LSE signs. (Terminological level)

- About the phonetics and phonology:

The avatar formational parameters are native LSE

The avatar signing speed is phonologically correct.

- Which of the above linguistic features affect comprehension the most?

About the translation:

The avatar translation errors

About the avatar appearance:

Size, colours, position, …

### *5.3.2. Challenges, problems and potential solutions*

The two interviewed users agreed that the most important aspects to be improved are:

1) Facial expression at both lexical, sub lexical and syntactic levels;

2) The use of space both syntactically and topographically.

When asked for a general assessment about the avatar, the two participants provided completely different reactions. While the deaf user was positive and considered the avatar was providing some access to the weather forecast information, the CODA user said she could not get any information above the lexical level. The deaf user seemed quite impressed at the beginning but started to find out that the avatar showed clear limitations to give him full access while playing the full clip. On the other hand, the position of the CODA user was clearly against the avatar from the beginning of the interview. She started playing the clip with voice, but due to the lack of synchronization she was unable to follow the signed version of the avatar. After some seconds she decided to stop and replay the clip from the beginning with no sound but she said it was

still impossible for her to understand any of the sentences. After a couple of minutes she stopped the clip and refused to try to make sense of the avatar signing discourse and commented on more specific features and both lexical and sub lexical levels.

When considering more specific aspects, they both agreed that the signing avatar signed unnaturally because of the lack of facial expressions, including mouthing. They also agreed that some of the signs were not comprehensible but they were not sure if this was because a different regional dialect was used or because the avatar hand shapes were not fully recognisable.

A summary of results follows.

The lack of **facial expression** makes the avatar look unnatural. Facial expressions convey not only affective information but it also affects virtually every linguistic level. On the discourse level, its linguistic value is similar to that of intonation in spoken languages, so that its lack makes it almost impossible to segment sentences in the string of signs. It, thus, affects coherence. On the morphological level, facial expression conveys adverbial information (for example, VIENTO FUERTE (strong wind) the signs should take a certain frown and the lip protruded and blowing air). On the lexical level, certain signs have a specific facial expression which is specified sublexically and might have a phonological value, so that two signs can only differ in their facial expression and have completely different meaning. In the clip, the string FATAL TIEMPO (terrible weather) was barely understood because of this.

The lack of **use of space** affects comprehension and naturalness. In sign languages it is the most important cohesive device. The space is used syntactically to shows the verb agreement between subject and objects. It is also used topographically to show spatial information. This was considered to be very important in weather forecast information clips because it should be used to express the movement of rain or clouds, in the territory.

The basic **syntactic structure** SOV was considered basic but comprehensible by the deaf user and too basic and compromising comprehension by the CODA user. Again, they agreed that the lack of facial expression and pauses sometimes made it difficult to understand when a sentence started and finished.

The **weather vocabulary** used was recognisable and well-selected but both users commented that signs such as LLUVIA (rain), VIENTO (wind) o MAR (sea)/OLAS (waves) were sometimes considered incorrectly used because they were not used in the correct form (facial expression and sign speed) that expressed the degrees or strength providing the appropriate meaning.

The **common vocabulary** was recognised as native LSE, except for the number signs when expressing temperatures, which were neither from the Madrid nor the Granada regional dialects. Both users commented that they looked like Catalan signs, or that they might be Basque. Some signs were not recognised or comprehensible but the users were not sure if this was because a different regional dialect was used or because the avatar handshapes were not pronounced naturally enough.

On the **sub lexical level**, the users commented on three features:

1) Some of the **handshapes** were not fully-fledged, especially those that would require an internal movement of the finger joints (such as, PRIMAVERA (spring) or NUBE (cloud));

2) Some signs that would need to **contact** a face or body part were signed without the contact (for example, VER (see) or CALOR (hot));

3) And sub lexical **facial expressions** (as in FATAL TIEMPO (terrible weather)) and **mouthings**. Sometimes a sign needs to be accompanied by the mouthing of the word or a special mouth configuration in order to select the appropriate meaning. This is especially important with polysemic signs (for example, PRIMAVERA (spring), PEOR (worse)/ BASTANTE (quite), TARDE (late), MOTIVO (reason)/ CAUSA (cause) / PORQUE (why), EJEMPLO (example)/ SI (If) conditional,

…)

The avatar **general appearance** was considered correct in terms of size, colours and on-screen position. The only negative comment was about the arm and forearm positions that were unnatural in some cases and sometimes made it more difficult to identify a given sign. Unfortunately, no specific example was provided.

In conclusion, given all this data, it was considered that there are some possibilities of improving the avatar performance, especially regarding facial expression. They should be considered and tested properly in a more representative context. We can take into account the findings of Silke Gutermuth [5], that indicate that sign language users focus their gaze on the face of the interpreter (see Figure 11 for a a record of eye movements of a user on the SLI screen). If these findings are confirmed by our research, facial expression of the interpreter may be the biggest source of information for sign-language users and has to be accurate in the avatar performance.

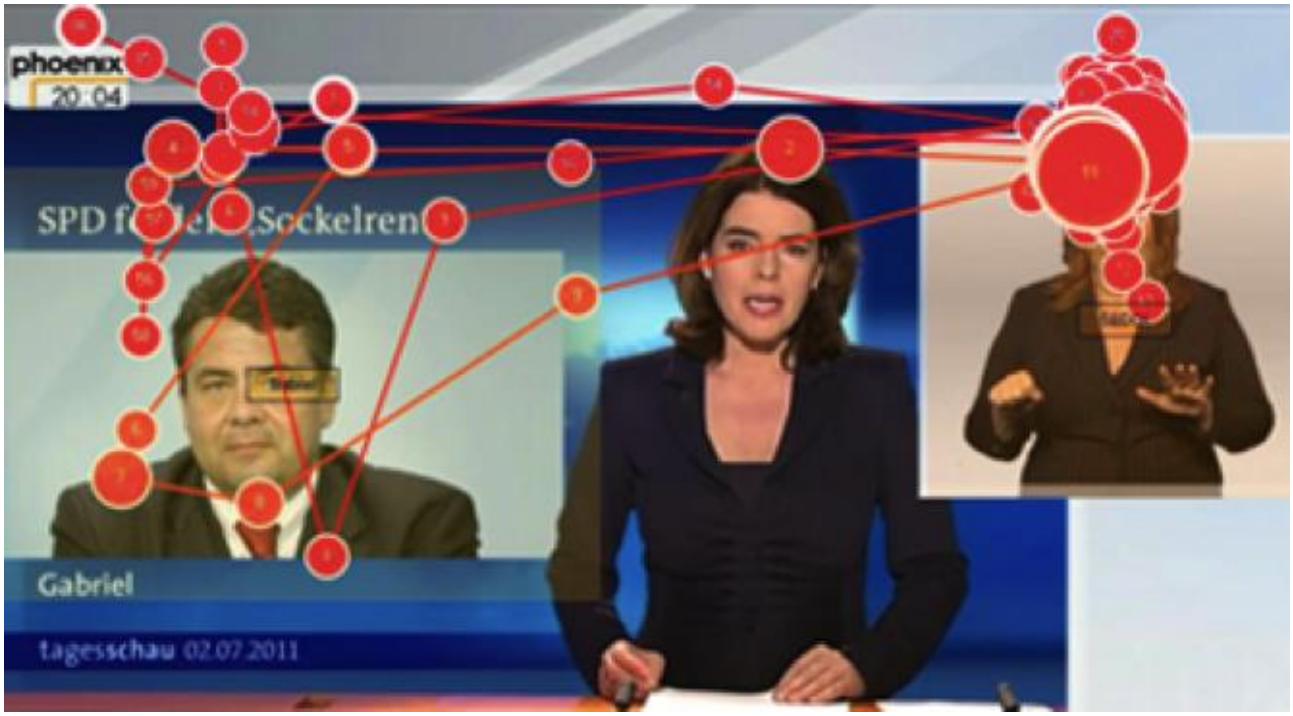


**Figure 11.** Eye-track recordings in sign-language users. Users focus mainly on the face of the interpreter, as shown by the size and number of red dots.

## 5.4. Overview of common challenges

Next table summarises the challenges of each subpilot and it includes the solutions that are proposed in each case to face the challenges and the possible problems. The avatar signing application is included in the table even if it is not a subpilot, due to its importance in the project.

| SUBPILOT | CHALLENGES | PROBLEMS/SOLUTIONS |
|---|---|---|
| HbbTV based signing pilot | Use HbbTV application to mix a Picture-in-Picture video stream with a main TV video and a sign language video area. It has to offer video adaptation functionalities and enabled for single-video decoding purposes. | Technical problems: the forwarding of user-chosen options to the video-mixing server could compromise availability of PiP sign-language videos.<br><br>Testing problems:<br><br>Testers may not have HbbTV equipment.<br><br>Solution: Provide the equipment and adjusting the sample group demographics.<br><br>Difficulties understanding the application of the methodology.<br><br>Solution: request local user association's advice. |
| Ip/Web-based signing pilot | Provide a double-screen functionality in live programmes using a web interface. User should be able to choose the size of both player windows. | Problem: Availability of record and storage capabilities to preserve the sign language signal is compromised because RTP cannot use professional equipment for this purpose.<br><br>Solution: UPM provides the equipment within the framework of Hbb4all.<br><br>Problem for non-live content: RTP splits large programmes in several video clips for web distribution.<br><br>Solution: take this feature into account in the implementation of the double-screen service.<br><br>Problem: possible incompatibility with some browsers.<br><br>Solution: implementation will take into account the most |

| | | |
|---|---|---|
| | | extended browsers.<br><br>Problem: the service will not be integrated in RTP web service, which will be then available to few people.<br><br>Solution: RTP links the service in the website, and promotes its use via the deaf people association in Portugal. |
| Avatar signing application | To provide a domain specific (weather) spoken language-LSE translation platform, by means of an Avatar.<br><br>The avatar should not only translate the information but also transitions between signs have to look natural, and overall performance should transmit emotion. | Problems: facial expression should be improved in order to achieve a better SLI. |

**Table 1.** Overview of common challenges.

Moreover, WP6 has identified the next common challenges:

- **Effective personalization**. Personalization is in fact a key issue in HBB4ALL in order to achieve completely satisfying service implementations. This issue has been tackled in a different manner in both service pilots. In the case of the HbbTV signing application, a truly hybrid approach had allowed a customisable implementation. However, the required capabilities (double video-decoder) are not present in conventional receivers. For this reason, the "all-in-one" solution was selected for the subpilot deployment. This solution is not flexible but RBB has identified the way to offer customisation options: the availability of multiple "all-in-one" video streams in the broadband link to be selected by the user. In the case of the IP/Web based signing pilot, customisation options are easier to deploy and preliminary tests in the user interface have shown how both web-embedded players can work simultaneously, while offering setting options (position, size).

- **User feedback compilation**. Since HBB4ALL consists of deployment of subpilots, the compilation of user feedback to check the features of the services is a key in the project. Pilot D has different strategies for this purpose. In the case of HbbTV signing application, RBB has designed a calendar of option deployment to know the user preferences. The service pilot will be available for a set of users. In the case of the IP/web signing subpilot integrating double screen feature, RTP will involve the participation of local deaf people associations. Finally, the complementary user tests carried out by UAB will provide more information about the service features.

# 6. Complementary user tests

The UAB team has carried out the following complementary user tests:

- Expert users test (included in D6.1 [1]): Sign language interpreters interviews
- Preliminary user test (included in D6.1 [1]): Focus group with signing Deaf users
- Preliminary user test (included in D6.2): Pretests on-screen size survey

The following complementary tests are prepared and ready to run next autumn:

- Tests for SLI on TV in lab conditions (from September): eye-tracking measurement, memory tests, comprehension test and a survey of preferences to assess the end-user experience.
- Test for the signing avatar (from November): user preferences and user experience with the signing avatar developed by Vicomtech.

## 6.1.Preliminary user tests

### *6.1.1. Focus groups with Deaf sign language users*

The first preliminary user tests included two focus groups with a total of 8 deaf sign language users. The goals, methods and outcomes of these tests were included in D6.1 [1].

After the interviews with the professional interpreters and the focus group with the deaf users, the list of formal parameters of SLI on-screen was narrowed down from a myriad of possibilities to the first two parameters to be assessed: size of the interpreter's sub-screen and position of the interpreter's sub-screen. However, before the experimental tests in lab conditions could be designed the variables for each of the parameters needed to be further limited.

The tests in lab conditions with on-screen sign language interpretation will be parallel to those carried out to test subtitles in WP3 of the project. The test will include eye-tracking measurement, memory tests, comprehension tests and a survey of preferences to assess the end-user experience. The procedure before the actual tests will include a small outline of the test and then filling out the consent form and the demographic survey. This will take approximately 20 minutes. After that the on-screen sign language test will begin. Individually, the participants will watch a video clip of about 3 minutes for each of the tested conditions. After each of the clips, the users will answer three different questionnaires to evaluate the user experience and preferences, the linguistic content memory and the visual content memory. We will need to allow about 15 minutes to conduct all the tests for each of the clip. Thus making of utmost importance to keep the number of parameters and variables to a maximum of 4 so that the full test does not exceed 1 and half hours per user to conduct. Following this decision it was agreed to test two variables for each of the parameters, size and position.

### *6.1.2. User survey with Deaf sign language users*

During the previous discussion with focus groups, Deaf consumers agreed that overlapping of on-screen elements should be completely avoided. Provided that subtitles (and sometimes the Digital On Screen Graphic) are normally displayed at the bottom of the screen, most users agreed that sign language sub-screen

could be placed in a central position. However, there was no agreement regarding the right/left location. Interestingly, when sign language and subtitling coexisted on screen, a variety of opinions raised in the focus group: one user said it was more comfortable to start looking the sign language on the left of the screen and then continue reading the subtitles from left to right. Another user argued exactly the opposite, that is, to start by reading the subtitles first from left to right and end by watching the interpreter on the right. One user pointed out that the position they favoured may have to do with their previous experience consuming sign language on TV.

Taking these data into account to first test the on-screen position parameter, we decided that the two variables that will be tested in lab conditions using the eye-tracker will be left and right positions to see if there is any particular location that is preferred or favours usability with no interaction with subtitles.

Both the Deaf consumers and the professional interpreters reported that the most important on-screen feature to grant accessibility was considered by all users to be the size of the interpreter. Most agreed that using a MS/MLS would be the ideal. Regarding the on-screen size most expressed that taking about a third of a vertically split screen would be good, but when showed with different screenshots no clear size was favoured by all the users.

Taking into account the data previously collected from the online platform Sign Language Television for the Deaf (http://signlangtv.org/), three sizes have roughly been identified: small (approximately 1/4 of the on-screen width), medium (approximately 1/3 of the on-screen width) and big (1/2 of the on-screen width). From these three different sizes only two can be included in the eye-tracking test, so that one should be left out in our first user tests. In order to decide which of the sizes will be tested a pretest was designed.

The experimental design for the on-screen size pretest was aimed to help us decide what two sizes would be included in the first sign language test in combination with the position parameter, right and left.

#### *6.1.2.1. Pretest report to evaluate the SLI sub-screen size on screen*

**AIM**

The aim of this pretest is to gather different opinions from a group of Deaf sign language users regarding the on-screen size of the sign language interpreter sub-screen. Their opinions will be used to curtail the number of sizes to be included as variables in the first sign language test. Three different sizes will be tested, namely, Small (approximately 1/4 of the on-screen width), Medium (approximately 1/3 of the on-screen width) and Big (approximately 1/2 of the on-screen width). In this pretest a group of 14 Deaf sign language users were surveyed using questionnaires aimed to assess their preferences and user experience while watching TV news clips in the three different sizes for the sign language interpreter sub-screen described earlier.

**METHODOLOGY**

**Participants**

14 Deaf sign language users participated in the pretest. A total of nine women and five men turned up. The age-range comprised between 23 and 73 year-olds. The distribution of the participants in the different age ranges was not uniform. Most of the participants belonged to either the youngest or the oldest age ranges (see Table 1). Among the youngest, all had access to higher education at university level whereas the oldest received either vocational training or compulsory basic education only.

| Age | > 40 (1942 - 1975) | < 40 (1976 - 1994) |
|---|---|---|
| absolute number | 8 | 6 |
| percentage | 57.14% | 42.85% |

**Table 2.** Distribution of participants' age.

All the participants were fluent in at least Catalan Sign Language and showed a certain degree of bilingualism with Spanish Sign Language. Furthermore, all but one of the participants acknowledged accessing some TV programs through sign language interpretation and manifested that they would enjoy more programs being accessible in this linguistic option.

**Sampling method**

In order to constitute a sample of the targeted population for the pretest, a videomessage in Catalan Sign Language was recorded and sent to the Catalan Federation of the Deaf (FESOCA), which disseminated the information to all its associate members, and Difusord, an important Deaf organisation that has an online newsletter for the Catalan Deaf Community. Both organisations disseminated the information through their social networks.

**Materials (apparatus, stimuli and questionnaires)**

Apparatus - A 19-inch computer screen was used for the user's clip visualisation. The computer screen was connected to a Mac laptop computer from which the interviewer controlled the input clips and the information being collected in the questionnaires. The online questionnaires were filled on an Ipad mini. We used the Google application "Forms" to collect and record the answers.

Stimuli - The input clips were created from different news programs broadcasted by the news channel RTVE Canal 24h on Thursday January 22nd 2015. The different clips were selected from three different times along the same day to allow the news to be similar in content but different in their linguistic form, presenter and duration. The times selected were in the morning from 07:32 to 08:18, afternoon from 13:04 to 14:39 and evening from 17:14 to 17:30. The raw broadcasted news was provided by the UPM team from which the contents of the three clips were selected. The contents for each of the clips were controlled to be similar in topic and duration as outlined in the following Tables 3 - 5.

| News topics | Duration |
|---|---|
| Employment survey results | 45" |
| Yihadist terrorism + King | 1'15" |
| Sports: Handball | 35" |
| Weather forecast | 50" |

**Table 3.** News Clip 1 (Total duration3'25").

| News topics | Duration |
|---|---|
| Yihadist terrorism + Parliament | 1'25" |
| Employment survey results + Ministry | 40" |
| Sports: Tennis | 20" |
| Weather forecast - Santander | 1'30" |

**Table 4.** News Clip 2 (Total duration 3'55").

| News topics | Duration |
|---|---|
| Yihadist terrorism + Parliament | 1'30" |
| Employment survey results + Ministry | 50" (+19") |
| Sports: Dakar | 20" |
| Weather forecast - Burgos | 1'20" |

**Table 5.** News Clip 3 (Total duration 3'59").

A professional sign language interpreter in collaboration with the bimodal bilingual researcher, from the UAB team, translated the linguistic content. A native Deaf sign language expert was consulted for specific neologisms, terminological items and name signs. A long mid-shot was used to record the interpreted signed version to allow a relatively bigger hand and face size on screen. Even though using a mid shot implies restricting the grammatical signing space, and thus sometimes losing naturalness, this is always preferred to having a longer shot that would make hand size look even smaller.

**Figure 12.** On-screen Small, Medium and Big sizes.

To conduct the pretest, a total of nine clips were edited by the UPM team, namely, each of the three clips in the three sub-screen sizes (see Figure 12). The clips in the test were randomised following a Latin square design to ensure that all participants were exposed to the same sub-screen sizes not repeating the content of the news clips.

Questionnaires. Two online questionnaires were designed using Google forms. The first questionnaire included 16 short questions aimed to collect the demographic features of the participants. The questions gathered basic demographic information: age, gender, education, deafness; linguistic information: sign language and written language skills; and TV accessibility preferences: subtitling and sign language.

The second questionnaire was designed to assess the different on-screen sizes, specifically, the user experience and the user preferences when watching news on TV (see Table 5). It included the same 9 questions for each of the clips and one more question at the end of the test.

| Question number | Question | Answers |
|---|---|---|
| Q1 | Did you have time to watch the sign language interpretation? | scale 1 - 10.<br>1 = I could watch all the interpretation<br>10 = I could not see any of the interpretation |
| Q2 | How easy do you think it is watching the sign language interpretation with this subscreen size? | scale 1 -10.<br>1 = very difficult<br>10 = very easy |
| Q3 | Do you think you missed important on-screen information because of attending to the sign language interpretation? | Yes<br>No |
| Q4 | How comfortable or enjoyable would you rate your experience watching news with this size of on-screen sign language interpretation? | scale 1 -10.<br>1 = comfortable / enjoyable<br>10= uncomfortable / unenjoyable |
| Q5 | How would you characterise the size of the subscreen for the sign language interpreter? | Very small, Small, Normal, Big, Very big |
| Q6 | What do you think about the size of the subscreen for the sign language interpreter? | Appropriate<br>Inappropriate |
| Q7 | What do you think about the relative size of the two subscreens? | Appropriate<br>I prefer a bigger SLI and smaller news<br>I prefer a smaller SLI and bigger news |
| Q8 | What do you think about the SLI screen shot? | Appropriate<br>Too short, I prefer a longer shot<br>Too long, I prefer a shorter shot |
| Q9 | What kind of TV programs would you like to watch in this size | News, Documentaries / educational, Series, Films, Magazine shows, Game shows, Reality TV, Sports, Others, None |
| Q10 | After watching the three clips, which on-screen size do you prefer? | Small<br>Medium<br>Big |

**Table 6.** Preference and user experience questionnaire.

**PROCEDURE**

To conduct the interviews three different sessions were organised. The surveys were held individually. For each of the participants a 30 minute slot was allowed. The participants chose both the date and time slot according to their availability. The three sessions were conducted by two bimodal bilingual researchers so that all the information and instructions were given using Catalan Sign Language, LSC, which was the language primarily used throughout the tests. The forms and questionnaires were originally written in Spanish and were sight-translated into LSC.

During the survey the participants were first welcomed by one of the researchers into a waiting room where she outlined the test components and objectives. The consent form and the first questionnaire were filled in this room.

After completing the first questionnaire, the interviews were held in a separate room. The interview room had a table and two chairs (one for the interviewer and one for the interviewee) near ample windows that

provided good light conditions for the test. The interviewer had a laptop computer with all the clips and presented them to the participants on a 19-inch computer screen in random order, following a Latin square design. The participants were placed in front of the screen and asked to watch the news clips. When participants reckoned that they had enough information they could stop the play of the clip and go to the questionnaire. After each visualisation they answered questions 1- 9 and only at the end of the test they would answer question 10 once.

**RESULTS**

This pretest was run with a small sample of participants, therefore we used non parametric tests to analyse the results. Our goal with this pretest was not to reach statistical generalisations, but rather getting insights from the participants' opinions that would lead us to choose the two sizes to be included in the lab-conditioned test.

The answers to the questionnaire were analysed with a non-parametric test for related samples, namely the Related-Samples Friedman's Two Way Analysis of Variance by Ranks test. In this section we will include the results from three questions: question 2 (Q2) and question 7 (Q7), which are the two surveyed questions that did show results statistically relevant, and question 8 (Q8) which turned out only one response.

The collected answers for question 2 (How easy do you think it is watching the sign language interpretation in this sub-screen size?) are shown in Table 7 below:

| Participants | small size 1/8 | medium size 1/3 | big size 1/2 |
|---|---|---|---|
| Total sum | 94 | 119 | 126 |
| Average | 6.71 | 8.5 | 9 |
| Possible answers | scale 1 = very difficult > 10 = very easy | | |

**Table 7.** Answers to Q2 - How easy do you think it is watching the SLI?

Friedman test analyses indicate that the differences in the answers for sizes Big and Small are statistically significant (p=.049) and the answers for Small and Medium show a trend to significance (p=.089).

The collected answers for question 7 (What do you think of the relative size of the two sub-screens?) are shown in Tables 8 and 9 below. The differences between small and medium size are not statistically significant (p=.131), but the differences between Medium and Big are significant (p=.008) and so are differences between Big and Small (p=.000).

| Participants | small size 1/8 | medium size 1/3 | big size 1/2 |
|---|---|---|---|
| Total sum | 22 | 28 | 38 |
| Average | 1.57 | 2 | 2.71 |
| Possible answers | 1 = I prefer a bigger SLI subscreen and smaller news subscreen<br>2 = Appropriate<br>3 = I prefer a smaller SLI subscreen and bigger news subscreen | | |

**Table 8.** Q7- What do you think about the relative size of the two subscreens?

Question 8: (What do you think about the SLI screen shot?) only produced one answer in all the three conditions by all the participants. In every case the shot size was considered 'appropriate'.

| Answer | small size 1/8 | medium size 1/3 | big size 1/2 |
|---|---|---|---|
| bigger SLI preferred | 6 | 2 | 0 |
| appropriate size | 8 | 10 | 4 |
| smaller SLI preferred | 0 | 2 | 10 |

**Table 9.** Q7- Absolute number of participants for each possible answer.

## CONCLUSIONS

When analysing the data from question 2 (How easy do you think it is watching the sign language interpretation in this sub-screen size?), the Friedman test analyses show that the differences between the answers are only statistically significant between the Small and Big sizes. But are similar when comparing the Small and Medium or the Medium and Big. Thus, according to the sampled users opinion, the bigger the sub-screen size is the easier it is to watch the interpreter on screen.

However, the results from the analyses on the participant's answer in question 7 (What do you think of the relative size of the two sub-screens?) show that both small-to-big and medium-to-big differences are statistically significant whereas the results comparing small and medium sizes are not. This seems to indicate that on the one hand, the bigger size is the least valued when taking into account the whole on-screen composition. On the other, it seems to indicate that the users do not make a great difference when judging the Small and Medium sizes.

Further analysing the results in this Q7, we noticed that none of the participants regarded the small-sized sub-screen as 'big' or 'very big'. This might indicate that the participants judge the 1/8-size the minimum comfortable on-screen size. Notwithstanding, smaller SLI sub-screens have been attested in the collected data for the project (included in Deliverable 6.1).

Sign language on-screen size not only depends on the size of the SLI sub-screen but also on the shot size. Our previous data from both the professional interpreter's interviews and the focus group (included in D6.1) suggested that the preferred shot size is a medium-shot/medium-long-shot (MS/MLS) with some space above head level to allow signs placed in the head area. Interpreters working on TV reported that when the sub-screen is too small they ask cameramen for a shorter shot size so that hand-size on screen is relatively bigger. Even though using a mid shot implies restricting the grammatical signing space this is always preferred to having a longer shot that would make hand size look even smaller. Furthermore, the answers from question 8 seem to indicate that a MS/MLS is found appropriate regardless of the sub-screen size.

In conclusion, the two sub-screen sizes that will be included in our future tests are the Small (1/8 of the screen width) and Medium (1/3 of the screen width). The screen shot size used will be a MS/MLS in all the cases. The aim of the test will be to evaluate using eye-tracking data which of the two sizes (Small or Medium) and positions (right or left) lead to better results along with the linguistic comprehension and memory tests.

### *6.1.3. Design and development of the stimuli and questionnaires for the sign language pilots*

In the past months the UAB team has been working towards designing robust tests to evaluate the formal parameters of sign language on TV. The materials are parallel in structure and content to those developed to evaluate subtitling in WP3 of the project. However, all the materials for the signing pilots are more complex in their design because they include not only a different language but one in a new language modality too and with no written form readily understood by the native users.

On the one hand, the content of the stimuli is the documentary "Joining the Dots" by Pablo Romero about audio-description. The central part of the documentary is an interview with Trevor, a man who became blind in his sixties. He explains his experiences and the importance of audio-description to access TV, cinema and theatre performances.

On the other hand, there are four different questionnaires. The first one is aimed to collect information on demographic features of the participants. The other three questionnaires have been designed to test the language comprehension and memory, the visual content memory and the user preferences for each of the conditions. In total 4 different variables will be tested: Small size, Medium size, Right position and Left position.

#### *6.1.3.1 Stimuli for the sign language pilot*

The translation of the linguistic text of "Joining the Dots" into Catalan Sign Language (or LSC) was first drafted from the written subtitles by a professional sign language interpreter and the UAB bilingual researcher. The final version of the signed text was interpreted from the re-voiced subtitles in Spanish by the SLI and controlled by both the UAB bilingual researcher and a Deaf sign language expert user.

The 4 stimuli video clips for the sign language pilot were created from the 11-minute documentary "Joining the Dots". In order to produce four different clips that would create similar conditions to test the four on-screen variables, several criteria were taken into account: duration, narrative content and linguistic content.

First, we wanted all the clips to have a similar duration. That meant that each clip would roughly be 2 minutes and 45 seconds. However, to meet the second criterion to create the stimuli clips the narrative structure of the documentary was used as a guideline. Thus preventing a scene being split into two or more different stimuli clips. This is important because during the pilot each participant will see the clips in a different order (not necessarily the narrative order) so that we wanted the clips to be as self-contained as possible. Following the third consideration, we used the version with Spanish subtitles developed for WP3 to take some basic measures on the linguistic content, namely, the total number of words and subtitles per clip, the total number of subtitles, the number of words per subtitle and the number of word per second for each of the clips. These figures can be seen in next Table 10.

| Clips | Start and end times | Duration | Number of words | Number of subtitles | words per subtitle | words per second |
|---|---|---|---|---|---|---|
| **clip 1** | 00:00 - 03:10 | 03’ 08’’ (188’’) | 297 | 47 | 631 | 1.58 |
| **clip 2** | 03:10 - 05:50 | 02’ 40’’ (160’’) | 328 | 46 | 7.13 | 2.05 |
| **clip 3** | 05:50 - 08:50 | 03’ 00’’ (180’’) | 289 | 45 | 6.42 | 1.61 |
| **clip 4** | 08:50 - 11:10 | 02’ 20’’ (140’’) | 227 | 36 | 6.31 | 1.62 |

**Table 10.** Stimuli clips design from the documentary "Joining the dots".

Once the signed text had been recorded and the clips designed. The UPM team has been working to edit the synchronised version of the clips with the interpreted stimuli in the four conditions that will be used in the pilot Small and Medium sizes and Right and Left positions (see Figure 13).

**Figure 13.** Stimuli clips in the four conditions to be evaluated.

### *6.1.3.2 Questionnaires for sign language users*

As mentioned in earlier sections of this document, different questionnaires have been designed to evaluate the effect of size and position of the sign language interpreter’s sub-screen on memory and linguistic performance by sign language users.

There are 4 different types of questionnaires: the demographic variables survey, the linguistic comprehension and memory test, the visual memory test and the user experience test. The demographic variables survey is designed to be completed once at the beginning of the test by each of the participants. It is the same that was

tested during the on-screen size sup-pilot (see section 6.1.2.1). The other three short questionnaires will need to be completed after the visualisation of each of the clips in the different conditions. On the one hand, there is the user experience test which includes different types of questions such as scales or multiple choice questions. The user test is the same for all the conditions to allow better analysis. The other two questionnaires are specific to the contents of the clips and include 5 linguistic memory questions and 5 visual memory questions. For every question three possible answers are presented together with a fourth option which is always 'I don't know / I don't remember'.

The questionnaires were first created in written Spanish and later translated into Catalan Sign Language by a Deaf sign language interpreter and the bilingual bimodal researcher from the UAB. Special attention was put to make the linguistic questions using the same lexical items that appear in the interpreted clips, thus controlling dialectal variation that could appear otherwise. It is of utmost importance that questionnaires and surveys are in the same language as the linguistic contents being evaluated. Failing to do so, it would be impossible to evaluate the participants performance.

However, this is technically challenging. The UPM team has been closely collaborating with the UAB team to define and develop bilingual bimodal online questionnaires for the sign language pilots. These questionnaires are connected to a database spreadsheet.

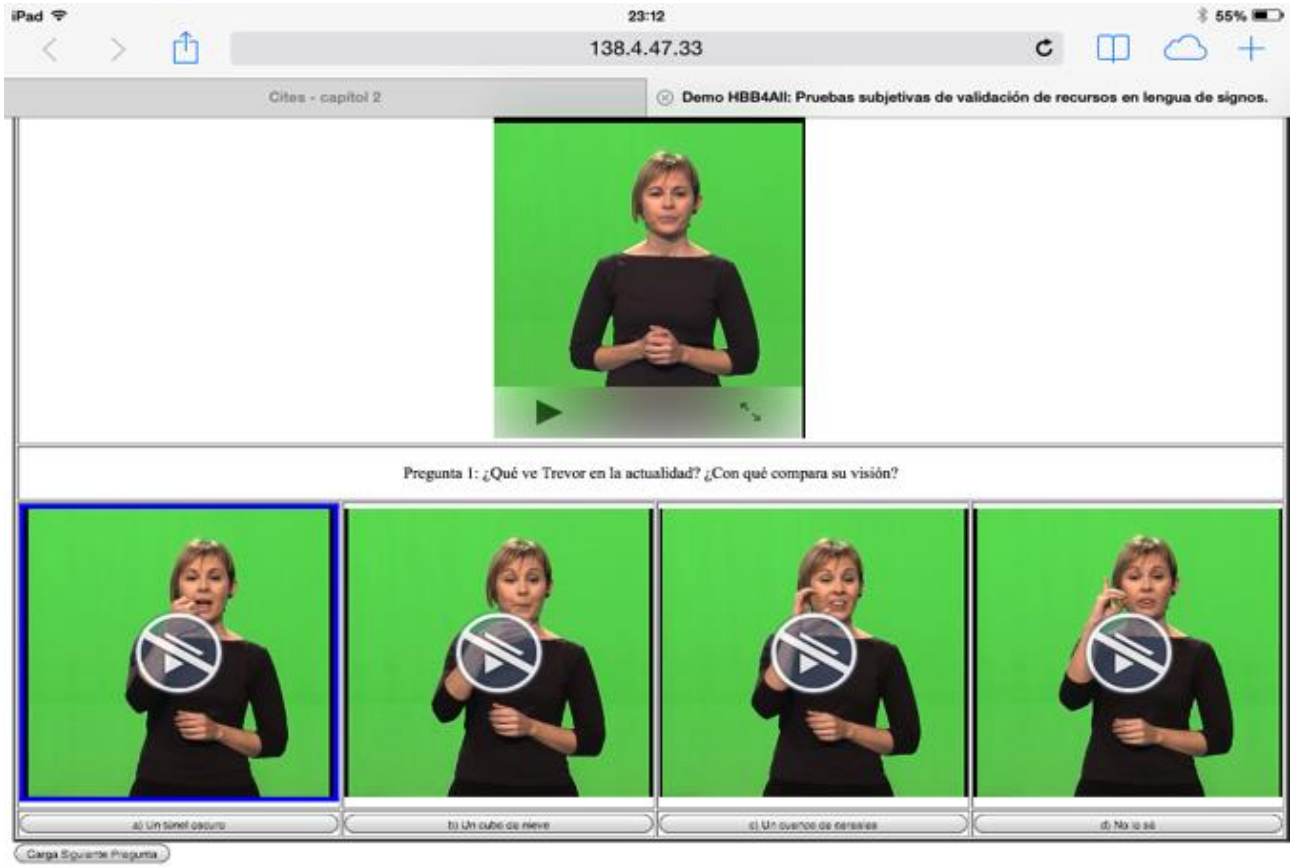


**Figure 14.** Example of question and four possible answers.

In this type of design (see Figure 14) there is no subordination of sign language with respect to oral / written languages. The same social and linguistic statuses are given to both modalities in the pilot materials. The participants can control the pace and replay any of the clips. Using bilingual materials there is no need for sight-translating the questionnaires. This gives a much more accuracy and consistency between the language variety use throughout the pilot. Thus making it possible to obtain reliable results. This methodology is innovative because it uses sign language as the main language for accessing, understanding and evaluating the information.

#### *6.1.3.3 Questionnaires for the sign language avatar pilot*

The questionnaires for the avatar pilot have been designed parallel to those for the sign language pilot. The demographics questionnaire includes 17 short questions. The questions are designed to collect basic demographic information: age, gender, education, deafness; linguistic information: sign language and written language skills; and TV accessibility preferences: subtitling and sign language. There is also one question aimed to the participant's previous experience using a signing avatar.

Three more questionnaires have been created to evaluate the avatar after visualising the weather forecast clip:

First, the user experience test that includes 13 questions either with scale answers or multiple-choice. Second, the linguistic comprehension and memory test that include 5 different multiple-choice questions with 3 possible answers each and a fourth option which is always 'I don't know'. Third, the visual memory test with five more multiple-choice questions with 3 possible answers. In this test the forth option is always 'I don't remember'.

The questionnaires were first written in Spanish and then translated into Spanish Sign Language (LSE) by a Deaf sign language interpreter in collaboration with the bilingual bimodal UAB researcher. In this case the sign language chosen was LSE to match the linguistic input that will be displayed by the avatar. In the linguistic questions, special attention has been paid to use the exact same signs that are used in the avatar dictionaries and other variants have been avoided.

The UPM team is developing an online bilingual bimodal questionnaire for the avatar pilot using the design in Figure 13.

## 6.2. Expert users' tests

### *6.2.1. Possible expert users' tests for the HbbTV signing workflow*

As described in section 4.1.1, RBB is considering adding a server-side video composition module to the workflow. This would allow for an automatic, command-line-based video mixing and would be triggered by the settings done in the HbbTV application. As this would potentially simplify the overall workflow and the video mixing, it must be thoroughly evaluated. For the case of the realisation evaluation objectives would be the cost efficiency, the picture quality and the affected service performance. The evaluation would be conducted with a small number of expert users from the production chain.

### *6.2.2 Complementary expert users' tests in lab conditions*

The UAB team has developed two different expert user test so far: the expert users' test for sign language on TV and the expert user pretest for the avatar developed by Vicomtech for weather forecasts.

***Interviews with SLI working on TV***

On the one hand, the experts that took part in our tests to assess the formal features of the sign language on TV were professional sign language interpreters that work (or have worked) on TV. The methodology used was a qualitative one, namely interviews. The interviews were held during the first year of the project. The methods and insights of these interviews were included in D6.1.

***Interviews with expert Spanish Sign Language (LSE) users***

On the other hand, UAB designed a pretest to help us to collect some guidelines to improve the signing avatar developed by Vicomtech. The report of this pretest is enclosed in section 5.3.2.

## 6.3. Summarised results and guidelines / recommendations

Only preliminar tests have been carried out in this Phase of the Project, we summarize their results and recommendations suggested in next Table 11.

| **Test** | **Results** | **Recommendation** |
|---|---|---|
| Screen size preferences | Users dislike biggest size tested | A SLI subscreen should be at least 1/8 of the total screen size and no bigger than 1/3 of the screen size |
| Screen shot preferences | Long medium shot preferred | Interpreter should be broadcasted in a medium/long shot |
| Avatar performance | Lack of facial expression that affects coherence and information in the discourse (morphology, lexical information, etc.) | Improve facial expression in the avatar in order to provide a correct SLI. |

**Table 11.** Summarised results and guidelines.

# 7. Conclusions

Work package 6 / Pilot D has advanced in a good manner and it is satisfying in general the work plan of the HBB4ALL project for this second period of the project. In the case of the provision of the IP/Web-based signing pilot a delay has appeared due to the lack of certain equipment in RTP. In any case, the subpilot will be deployed inside the Operational Phase of the project. Except this issue, activities in tasks T6.1 and T6.2 have evolved according to the plans.

As detailed in this deliverable, the main activities carried out during this period have been:

- Revision of the HBB4ALL signing workflow model. Particularly, the technical approach has been modified to gain in flexibility and to include the final implementations of the service pilots.
- Preparation of the HbbTV-based signing application and service pilot.
- Preparation of the IP/Web-based signing interface involving double screen feature and the respective service pilot.
- Creation of an avatar signing service for weather forecast.
- Complementary user tests to measure the optimal features of generic signing services.

User tests are a key part of HBB4LL. The UAB team has successfully conducted both the pretest for the signing avatar and the pretest to assess the size of the SLI sub-screen on TV.

The pretest for the signing avatar gathered useful insights by expert users from which the UAB team has provided Vicomtech with the necessary feedback with possible improvements for the second version of the avatar. (See section 5.3.3).

The pretest developed to assess the size of the SLI sub-screen on TV has effectively provided us with some useful insights. Finally, the myriad of formal features that could have potentially been tested in the sign language pilots of the project have been curtailed to two controlled variables. The two variables that will be tested are size, small (approximately1/4 of the screen width) and medium (approximately 1/3 of the screen width), and position, right and left of the screen. Additionally, the screen shot size has further been narrowed down to a MS/MLS, constituting one of the constant features of the future sign language test. (See section 6.1.2.1).

In the next few months the UAB team in collaboration with the UPM team will finish the bilingual stimuli and questionnaires for the sign language tests and the avatar test. The UAB team plan to start the participants' recruitment via the user organisations in September 2015.

# 8. References

[1] D6.1 - *Pilot-D Progress report*, HBB4ALL deliverable, December 2014

[2] D2.3.2 – *Common Technical Components (II)*, HBB4ALL deliverable, 2015

[3] ETSI TS 102 796 *"Hybrid Broadcast Broadband TV"*.

[4] ETSI TS 102 809 *"Digital Video Broadcasting (DVB); Signalling and carriage of interactive applications and services in Hybrid broadcast/broadband environments"*.

[5] Gutermuth, Silke (2011). Bickverhalten Gehörloser bei der Nachrichtenrezeption mit GebÄrdensprachdolmetscher - eine Pilotstudie am Beispiel PHOENIX TV. Unpublished M.A. Thesis. Johannes Guterberg-Universität Mainz