Grant Agreement No.: 687645
Research and Innovation action
Call Topic: H2020 ICT-19-2015

ORPHEUS
Object-Based Audio Experience

Object-based broadcasting – for European leadership in next generation audio experiences

D5.4: Implementation and documentation of user interfaces associated with last personalisation and interaction mock-up scenarios

Version: v0.6

Deliverable type	DEM (Demonstrator, pilot, prototype, plan designs)
Dissemination level	PU (Public)
Due date	31/12/2017
Submission date	12/01/2018
Lead editor	Chris Baume (BBC)
Authors	Niels Bogaards (ECANDY), Chris Baume (BBC), Andrew Mason (BBC)
Reviewers	Andreas Silzle (FHG), Werner Bleisteiner (BR)
Work package, Task	WP5, T5.2
Keywords	Personalisation, interaction, user interface, usability, utility, intelligibility, dynamic range

Abstract

Two different user interfaces have been created to explore personalisation of, and interaction with, object-based broadcasts. An app on an iPhone demonstrates personalisation by enabling the listener to change the balance between different elements of a programme, or to have the sound adapted automatically to one of several pre-defined listening environments. The app also allows the user to choose a shorter or longer version of a programme, to jump to points of interest, and to see transcripts.

The web browser interface developed for The Mermaid's Tears drama allows the listener to choose between parallel narrative threads, supporting the story-telling with images. The choice of stereo, binaural, or surround sound reproduction is also offered.

Document revision history

Version	Date	Description of change	List of contributor(s)
v0.1	09/11/17	First draft	Chris Baume (BBC)
v0.2	07/12/17	Added iPhone interface description	Niels Bogaards (ECANDY)
v0.3	07/12/17	Added early draft of The Mermaid's Tears section	Chris Baume (BBC)
V0.4	03/12/18	Editorial improvements. Incorporation of abstract, summary, and conclusions sections	Andrew Mason (BBC)
V0.5	10/01/18	Incorporating comments from review process	Andrew Mason (BBC)
V0.6	12/01/18	Final editing	Anja Köhler, Uwe Herzog (Eures)

[1] http://creativecommons.org/licenses/by-nc-nd/3.0/deed.en_US

Executive Summary

Two different user interfaces have been developed to demonstrate and explore the possibilities that object-based audio broadcasting offers the listener in personalisation and interaction. As well as showing what can be achieved, the development of the interfaces increased the understanding of how best to present choices to listeners, and what features are likely to be most well-received.

An iPhone app, developed by ElephantCandy demonstrates how simple interfaces may be offered to the listener to control features with a wide range of technical complexity.

Object-based audio can send foreground and background sounds separately, mixing being done in the receiver. The app offers a simple interface to the user to improve clarity (intelligibility) which controls the balance between foreground and background. More specifically, it is possible to characterise some objects as speech, and give those more prominence (than, say, foreground sound effects).

Dynamic range can be reduced by an extent determined by the listening environment. Pre-defined types of environment can be detected and adaptation automated (e.g. by detection of home WiFi network).

The app can render the audio objects to different audio formats, such as stereo, binaural, or, particularly for forwarding to home HiFi, multichannel surround sound.

If interaction is enabled in the audio stream transmitted, the listener can use positional control to move sounds around in a 3D audio space. A choice of different spatial renditions is also offered for some programmes. The controls for this kind of interaction become more complex than an “on” or “off”, or “more” or “less” selection.

Variable-length playback control has been included in the app, using ‘level of importance’ tags defined during programme production. The full chain, (adding the tags, exporting them with the audio, then propagating them to the receiver) has not yet been implemented.

An interactive radio drama, “The Mermaid’s Tears”, was produced by the BBC to explore the possibilities of multi-threaded narrative, spatial audio, and additional data (images) in radio. The drama was streamed live, and later made available on BBC Taster (a platform where the BBC tries out new, experimental, ideas). More than 6,000 people tried the experience on Taster, and it was rated 3.9 out of 5, on average.

The drama offered three intersecting narrative threads, each corresponding to one of the three characters involved. The audio followed the character selected, as they moved from room to room in the drama, and so included their voice, those of others in the same room, and sound effects in that room. Images to support the narrative were triggered by sending URLs to the browser at appropriate times.

The audio was rendered in one of a number of available formats, according to the choice of the listener, and the capabilities of the receiving device (the number of audio outputs).

It can be concluded already that personalisation and interaction features are generally welcomed by listeners. Care needs to be taken in the design and presentation of controls so that they meet the requirements of utility and usability.

Table of Contents

Executive Summary **3**

Table of Contents **4**

List of Figures **5**

Abbreviations **6**

1. Introduction **7**

2. ORPHEUS iPhone demonstrator **8**

2.1. Clarity 9

2.2. Programme segmentation 10

2.3. Transcript navigation 11

2.4. Spatialisation 12

2.5. Variable-length playback 13

2.6. Location-based profiles 14

3. The Mermaid's Tears pilot programme **15**

3.1. Introduction 15

3.2. Interaction 15

3.2.1. Triggered images 15

3.2.2. Spatial audio 15

3.3. Implementation 16

4. Conclusions **17**

List of Figures

Figure 1: The “Clarity” screen in the ORPHEUS radio app 9

Figure 2: Points of Interest in the Playback screen 10

Figure 3: Live transcripts in two scenarios: an interview and a sports commentary 11

Figure 4: Various way of interaction with audio objects and immersive audio rendering 12

Figure 5: Design of the “Program Length” screen with a bar indicating the sections that are played and those that are skipped 13

Figure 6: The Profile selection screen and the conditions and settings a Profile contains 14

Figure 7: Interface of The Mermaid's Tears 16

Abbreviations

AAC	Advanced audio coding
ADM	Audio Definition Model
API	Application programming interface
OBA	Object Based Audio

1. Introduction

This document describes the implementation of two user interfaces – one in an iPhone app, the other in a web browser – for providing personalised and interactive audio experiences in the ORPHEUS project.

The implementation is preceded by the need to understand what should be offered to the listener to enable them to personalise the experience and interact with object-based audio features. The terminology to be used needs to be easily understandable, the controls need to be easy to use, and the end result needs to be desirable.

The implementations described below show the usability and utility of object-based audio. Their development has required an understanding of the underlying technology, and of user-interface design.

The remainder of this document shows the design of user-interface components in an iPhone app and a web browser to allow the user – the listener – to personalise and interact with object-based audio services, to improve intelligibility, to explore multi-threaded narratives, to tailor their experience, and to immerse themselves in audio to the degree possible with their device.

2. ORPHEUS iPhone demonstrator

The ORPHEUS Radio iPhone app, designed and implemented by ECANDY, serves as one of the reception endpoints of the Orpheus object-based broadcasting chain[2]. The objectives of the app are manifold, including:

- find understandable and useful terms and interfaces for new concepts in OBA Broadcasting
- validate the functioning of the object-based audio features proposed in ORPHEUS
- serve as a demonstration platform for the ORPHEUS project and its technologies
- investigate the use and usefulness of OBA broadcasting on mobile devices and in mobile scenarios

This chapter does not aim to provide a complete overview of all screens, settings and concepts used in the ORPHEUS Radio app. Instead, it focuses on a number of distinct features where a user interface was developed to highlight the end-user interaction with technological innovations.

[2] As was shown in Figure 8 in ORPHEUS Deliverable 2.2 "Interim Reference Architecture Specification and Integration Report"

2.1. Clarity

One immediately useful possibility of OBA is to increase the clarity or intelligibility of the audio mix according to the listener's preferences and environmental circumstances. Concrete examples of this are the compression of the dynamic range of the sound to reduce the masking of softer passages and sounds by environmental noise, and the relative amplification of speech by adjusting the balance between various objects in the audio stream.

In the ORPHEUS radio app, three tools are presented to adjust the clarity:

- Quiet parts louder: compress the dynamic range
- Speech enhance: apply compression, equalisation, and level adjustment of individual audio objects, based on their content
- Foreground/background balance: control the prominence of certain audio objects over others.

Currently only the Foreground/background balance is interfaced with the MPEG-H decoder. Experiments are on-going for the other functions. The user interface to the clarity controls is shown in Figure 1.

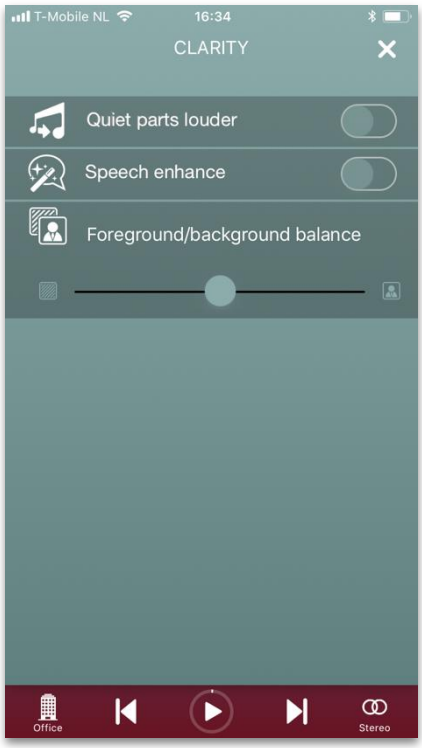

Figure 1: The "Clarity" screen in the ORPHEUS radio app

2.2. Programme segmentation

An important aspect of OBA broadcasting is the attachment of rich metadata to the objects and the programme. Metadata is needed to identify characteristics of the objects, but also to provide information about the objects in the programme that may be time-stamped and time-varying. In the ORPHEUS radio app, time markers can be used to identify points of interest in the programme, to which the listener can jump or from which extra information can be accessed. On the left in Figure 2 is a programme with 5 points of interest within it. On the right is the extra information displayed for one of them.

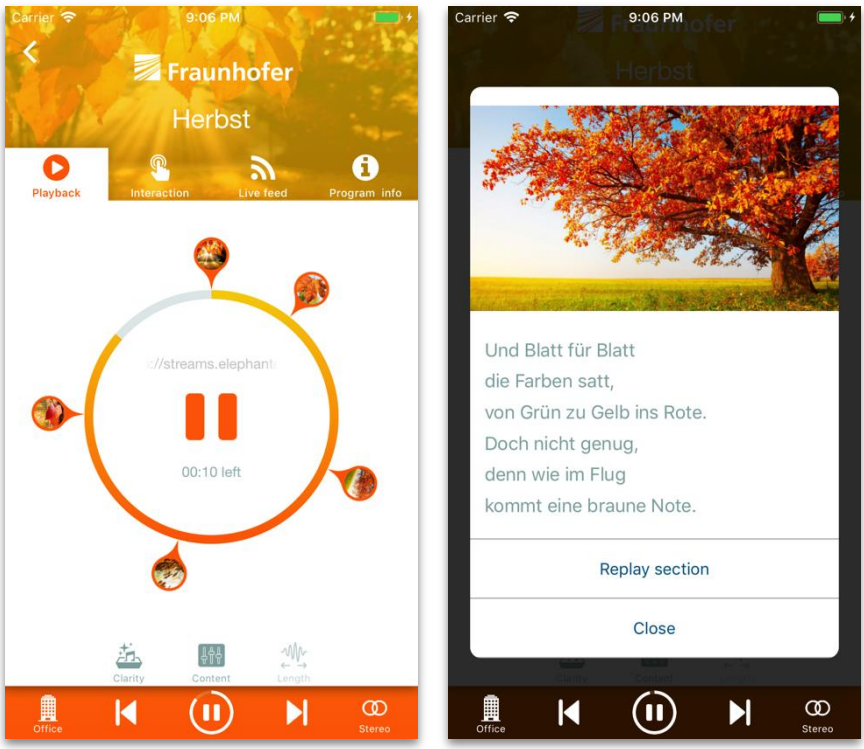

Figure 2: Points of Interest in the Playback screen

2.3. Transcript navigation

The ORPHEUS radio app can display a transcript of the current programme, updated live. This transcript can be used to catch up quickly with content that has already passed when the listeners joins a program, or to replay a section of interest. Additional metadata can convey information about the speaker, narrator, or music being played.

As in the audio stream itself, multiple languages are supported, and the preferred language can be set in order to show the text in a personalised way. Two examples of the transcript are shown in Figure 3: on the left, an interview; on the right, commentary on a sports event.

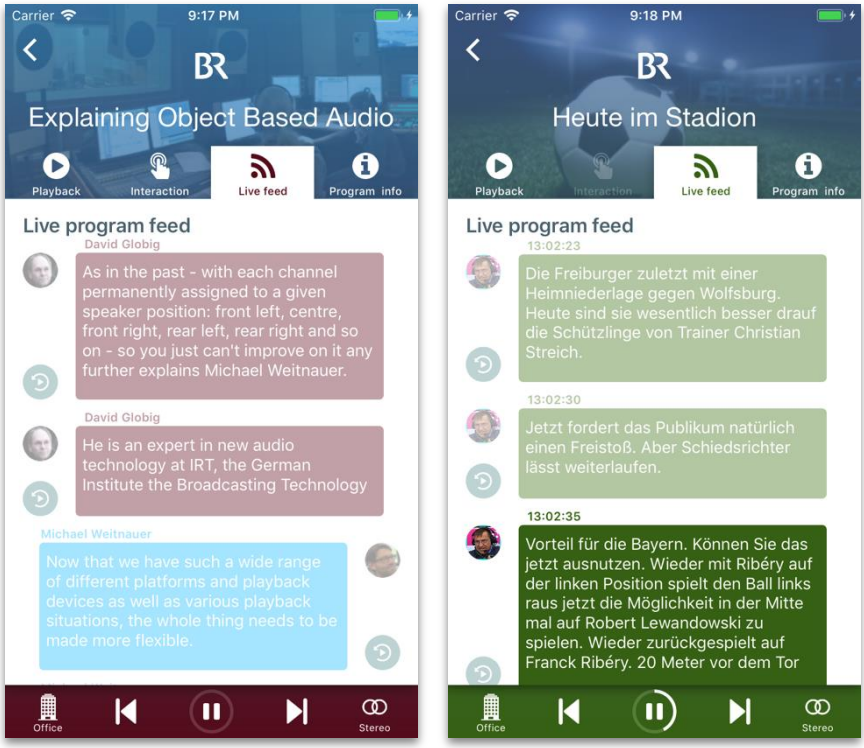

Figure 3: Live transcripts in two scenarios: an interview and a sports commentary

2.4. Spatialisation

ORPHEUS puts a large emphasis on spatialised immersive audio. In the app, three aspects of spatialisation are manifest:

- placement of objects in (3D) space in pre-defined scenes
- placement of objects in (3D) space interactively
- rendering of the stream to different audio formats or (channel-based) speaker layouts, such as stereo, binaural, or “5.1” surround

The presentation of the controls on the user interface for these three aspects are shown from left to right in Figure 4. On the left, the interface for this musical ensemble piece shows that the listener may choose to have the instruments (in this case, clarinet, violin, cello, and guitar) spatialized in one of four pre-defined arrangements. “Classic” puts them in a row in front of the listener; “Corners” puts one in each corner of an imagined square room; “Cross” puts one in front, one behind, and one at each side of the listener; “Rotating” causes the instruments to circle continuously around the listener.

The centre of the figure shows another programme, in which the listener may change the position of a single audio object in terms of prominence, elevation, and azimuth, using three sliders.

The right of the figure shows how the listener may select an audio format from those available. In this example, the app does not allow the selection of formats for which it cannot detect the capabilities needed (head-tracking, connection to surround sound loudspeakers).

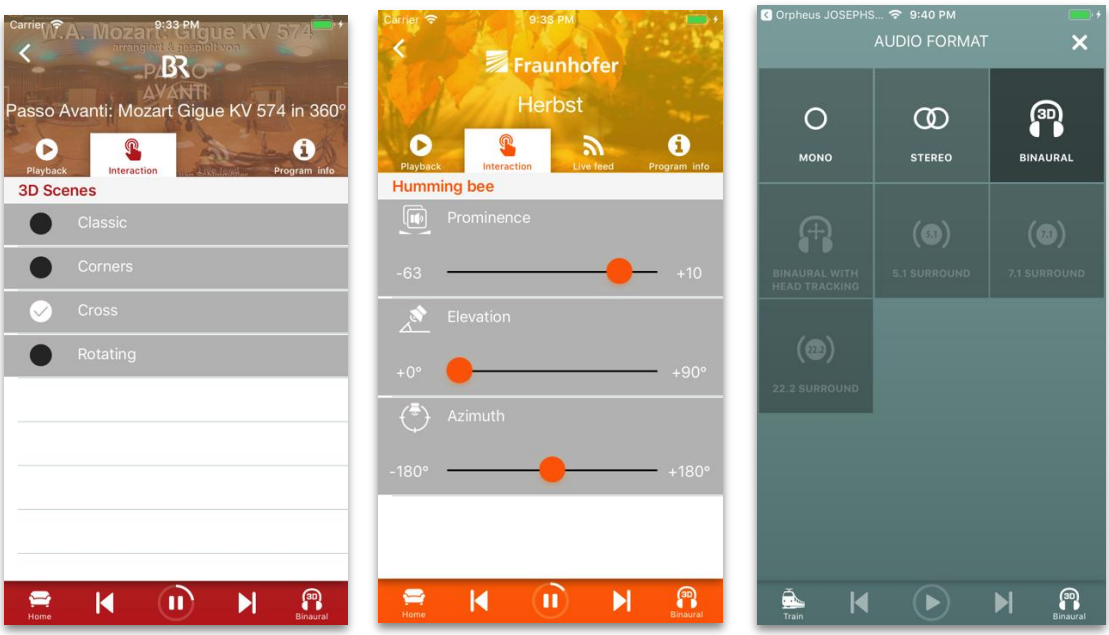

Figure 4: Various way of interaction with audio objects and immersive audio rendering

2.5. Variable-length playback

An important part of ORPHEUS is the ability to change intelligently the length of a radio programme, based on user preferences. The preference might be an interest in certain topics, the wish to skip music, or the need to fit the programme into the time available.

While this aspect of the ORPHEUS project is scheduled near the end of the project, and implementations have not yet been completed, the concepts have been explored in depth, and the user interface for the app has been designed. The scope of variable-length playback support in the ORPHEUS project has been limited to time or content adjustment within a single radio programme: there is no composition of a new programme by aggregating parts of interest from multiple programmes. Concretely, this implies that it is mainly about reducing the programme length in an intelligent way.

With this constraint in mind, the goal in the design has been to communicate which parts of the programme will be skipped as a result of the preferences of the listener. The most important concept here is the 'level of importance' of a section of the program. These levels can be defined at the production stage, where metadata is added to delimit and annotate sections, but they are also further interpreted at the listener's end, by applying constraints or preferences to the levels. In the user interface, the controls provided to alter the length are accompanied by a visual indication of the sections that will be played and those that will be skipped. This is shown in Figure 5.

Figure 5: Design of the "Program Length" screen with a bar indicating the sections that are played and those that are skipped

2.6. Location-based profiles

With the design of the user interface for the ORPHEUS radio app, it immediately became clear that object-based audio broadcasting can offer a very large number of options and possibilities to the listener, many of them involving complicated technical concepts. To reduce the amount of interaction needed and present information as simple as possible, the concept of *Profiles* was introduced. A *Profile* combines various settings (audio format, foreground/background balance, preferred language, etc.). Moreover, the settings in a *Profile* can be linked to a combination of technical factors, such as the Internet connection, the current device and reproduction hardware, and environmental factors, such as listener's location, current activity, and the time of day. The combination of all these factors is called a *Situation* and the app can automatically recall a Profile when the conditions of its *Situation* have been met. For instance, when the listener enters her home, an AirPlay[3] connection to the home hi-fi can be made and the audio format automatically switched to 5.1 Surround. On the left in Figure 6 is the interface for selecting a Profile. On the right are shown all the settings within that selected.

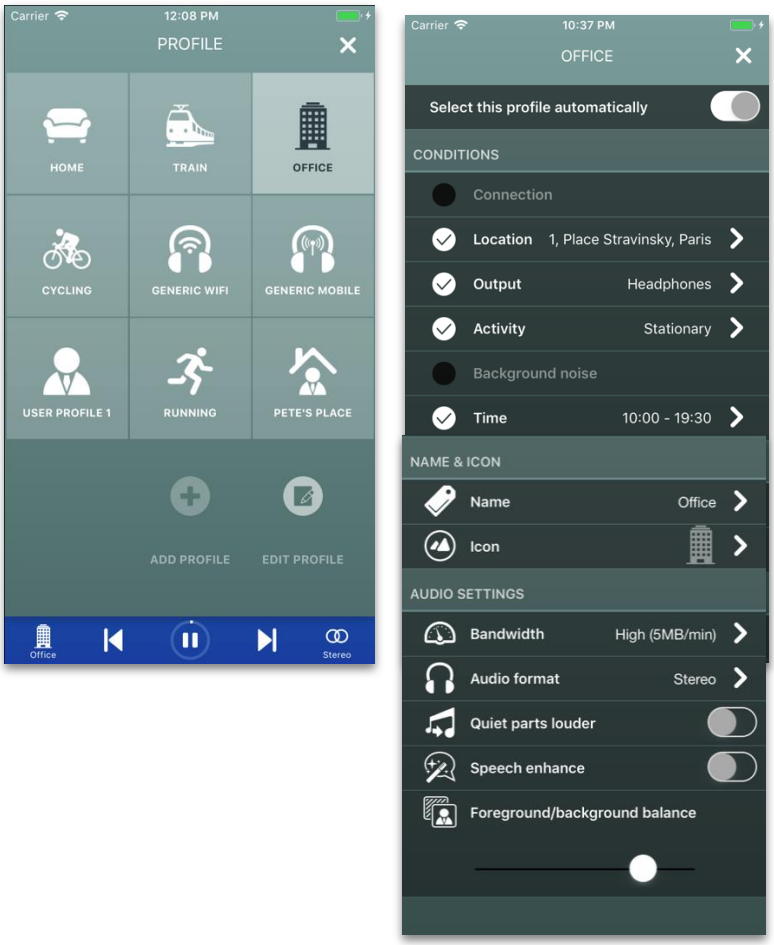

Figure 6: The Profile selection screen and the conditions and settings a Profile contains

[3] AirPlay is a protocol stack, developed by Apple, that allows wireless streaming between devices of audio and video

3. The Mermaid's Tears pilot programme

3.1. Introduction

"The Mermaid's Tears" was an interactive radio drama produced by the BBC as part of ORPHEUS pilot phase 1, and made available to the public via the BBC Taster website[4]. As of November 2017, over 6,000 people had listened to the drama and given it an average rating of 3.9 out of 5.

3.2. Interaction

The main interaction feature that was tested in ORPHEUS pilot phase 1 was switching between multiple parallel storylines. In "The Mermaid's Tears", this was realised through the use of three characters that moved between different rooms during the drama. Listeners could select one of the three characters at any time. When a character was selected, the listener would follow that character as they moved between rooms, and hear what was said and what happened in each room.

The use of parallel audio streams was designed from the very beginning of production. The writer of the radio play scripted the drama to make use of the fact that listeners could not hear more than one of the streams at any one time. In "The Mermaid's Tears", the characters Dee and Bill were police officers who were interviewing a mother called Lesley. At certain points in the drama, Dee and Bill would speak behind Lesley's back about their doubts about her innocence. As such, listening to different characters would give listeners a different perspective on the story.

In addition to having parallel audio streams, each story was accompanied by images of the rooms and the characters that were triggered at specific times, and the drama was available in multiple spatial audio formats. These are explained individually below.

3.2.1. Triggered images

With audio alone, it can be difficult to keep track of what characters are in which room at any one time. To assist the listener, the drama was accompanied by images that were triggered to appear at particular times in the programme. These were linked to the characters that the listener had chosen. For example, if the listener was following the character Dee, when Dee moved from the living room to the bedroom, the images were automatically updated to show the change of room. The images were made up of two layers – the background image displayed the room, and the foreground images displayed the characters in the room.

3.2.2. Spatial audio

The audio for the pilot programme was delivered using object-based audio and rendered in the browser. This allowed us to provide the user with multiple outputs for spatial audio playback. This was made available to user with a "change output" button positioned beside the play button. The options were:

- Stereo
- Binaural
- 5.0 surround sound (when using a device with 5 output channels)

[4] www.bbc.co.uk/taster is where members of the public can try, rate and share new ideas from the BBC and its partners

3.3. Implementation

The interface for the pilot programme was made available through a website delivered through BBC Taster. The system was designed so that users did not have to install any special software: it used HTML5 Web Audio, so was compatible with all 'evergreen' browsers such as Chrome and Firefox[5]. The system was also designed to use a 'responsive' interface that scaled appropriately for mobile screens. However, due to the iPhone web browsers being out-of-date, the system was not compatible with iPhones when it was first launched.

Figure 7 shows the finished interface. The character Dee is selected, and the image display shows all of the characters superimposed on the living room.

Figure 7: Interface of The Mermaid's Tears

The audio playback was implemented using a Javascript library developed by the BBC. This received the audio objects compressed using AAC, and the object metadata using serialized ADM, then rendered the final audio output using the Web Audio API. The Javascript library is currently being prepared for release as open-source software.

[5] 'Evergreen' browsers update themselves automatically to the latest stable release, without the user having to take action

4. Conclusions

The development of two new user interfaces to offer listeners personalised and interactive experiences has been crucial in increasing the understanding of how to present to the listener the multiplicity of possibilities that object-based audio enables. It has also increased our knowledge of what features are most appreciated by listeners.

The design of the user interfaces has taken into account the need to present to the user the necessary and sufficient controls to effect what is desired. Even where the underlying technology is complex, the ideal design can result in a trivially simple user control that is immediately understood.

The development of the user interfaces, and the authoring of the new programme material on which they operate, has proceeded in parallel, and continues to push forward the boundaries of knowledge in creating and presenting object-based audio.

Implementing the user interfaces, such that experimental programmes can be exposed to the public, in particular “The Mermaid’s Tears”, results in valuable feedback that can be used to assess the cost and benefit of providing new types of experience.

Some technical processes of personalisation and interaction are yet to be implemented, although the user-interfaces already exist – such as the end-to-end chain required for variable-length programming – and these will follow in the remainder of the project.

[end of document]