

COMPETITIVENESS AND INNOVATION FRAMEWORK PROGRAMME

ICT Policy Support Programme (ICT PSP)



ICT PSP call identifier: **ICT PSP/2007/1**

*ICT PSP Theme/
Objective identifier:* **2.1: Accessible digital Audiovisual (AV) system (Pilot B)**

Project acronym: **DTV4All**

Project full title: **Digital Television For All**

Grant Agreement no.: **224994**

Contract duration: **01.07.2008 – 31.12.2010**

Deliverable no.: **D2.3**

Deliverable title: **Interim Report on Pilot Services**

Nature of deliverable: **Report**

Dissemination level **Public**

Due date: **31.03.2009**

Actual delivery date: **30.06.2009**

Document title and history:

D2.3 – Planning evaluation of mature Pilot Services

<i>Version no:</i>	<i>Date:</i>	<i>Changes</i>	<i>Contributor</i>
001	06-05-2009	Structure of document in place (based on D1.1)	POL
002	07-05-2009	New iteration	POL
003	07-05-2009	Content	POL
004	15-05-2009	Added DR sections to report	POL
005	19-05-2009	Added RBB sections to the report	POL (BH)
006	22-06-2009	Added contribution from RedBee	POL/NR
007	28-06-2009	Executive summary added	POL
008	28-06-2009	Editorial revisions	TJO

<i>Document authors / organization:</i>	Pilar Orero – UAB, Peter Olaf Looms, DR
<i>Partners contributed:</i>	All
<i>Availability:</i>	Final version: Open
<i>Circulation:</i>	Partners

1 Executive Summary

This report covers the first six months' work on the planning and implementation of the Pilot of Mature Digital Television Access Services in four countries undertaken as part of the DTV4All project between 1 October 2008 and 31 March 2009.

Section 2 of the report introduces issues that highlight the need for the Pilot including:

- Inadequate statistics on the current access services provision in the 27 EU member states
- Low awareness of e-inclusiveness issues throughout the value chain of digital television Gaps in the current knowledge of the efficacy of existing access services
- The diversity of circumstances across Europe making it necessary to offer a range of access service solutions
- Changing priorities among key stakeholders including special interest groups representing those with physical and cognitive impairments
- The impact of the economic downturn on access service provision

Section 2 goes on to explain the detailed structure of the report in light of these issues.

Section 3 explains what has been planned and what activities are already underway in the four countries, namely, Denmark, Germany, Spain (Catalonia), and the UK. This includes the evaluation of Audio Description (AD) and Live Subtitling in Denmark; user preferences for DVB-Subtitling and Clean Audio in Germany; Audio Description and Signing on digital terrestrial television in Catalonia; and the evaluation of an access service awareness campaign focussed on audio description in the UK.

Another strand of the Pilot involves the undertaking of academic research on access service users' preferences on subtitles for the deaf and hard of hearing in seven countries to establish user preferences with respect to the legibility of the subtitles, their position, character identification, and context information. Work is also being done on the conventions underpinning the production of Audio Description. The aim is promote interoperability and the exchange of access services across Europe.

Section 4 describes the work completed by 31 March 2009 in the four countries.

Section 5 gives the preliminary findings of the Pilot in the 4 countries and some indication of how these will influence what is to be done in the next reporting period.

Section 6 includes preliminary conclusions from the first six months of the Pilot.

Section 7 discusses the implications of the work done so far for work in the next reporting period.

Sections 8 to 11 comprise appendices giving further details the work outlined in sections 3-6.

2 Introduction

2.1 What are the aims of this report

The vision underpinning DTV4All is one of e-inclusiveness. It is about a Europe in the not-too-distant future where as many Europeans as possible are able to access digital television

Access services for digital television are already available in many EU member states. DTV4All operates at the policy level and aims to identify measures to accelerate e-inclusiveness.

To improve the e-inclusiveness of digital television, action is required on three fronts:

1. In the short term, facilitating the take-up of mature access services on what the project team has termed first generation digital television. This will be operational over the timeframe 1997-2012 and is concerned broadcast systems based on MPEG2 technology.
2. Preparing for the second generation of digital television by assessing the viability of mature services on second generation digital television platforms.
3. Identifying and validating emerging solutions that will either replace mature access services, or extend the scope of access service provision, on second generation digital television platforms.

This report covers the Pilot of Mature Access Services (WP2) will contribute inputs to the first two action points.

A legitimate reaction to the notion of a Pilot of mature access services that has been noted several times by the project team since the project started is: “If the access services are mature, isn’t a pilot of them superfluous? As know what is needed, isn’t it just a question of getting started?”

What the work of the first six months of the DTV4All project has shown is that a Pilot *is* needed because:

- *The baseline for measuring success in increasing the roll out of access services on digital television is ill-defined* – adequate statistics are not available to establish what access services for digital television are currently offered in the 27 EU member states and the rate at which these services are being extended. Working with the European Broadcasting Union (EBU), the project has put in place a mechanism for an annual survey of EBU members so that comprehensive and up-to-date figures on what is being done will be available.
- *Low awareness* - knowledge of the access problems across Europe, their causes and potential solutions is patchy throughout the value chain of digital television. Even in states where there is a good offering of services like Audio Description, many of those who would benefit from them are unaware of their existence.

- *Gaps in knowledge of the efficacy of existing access services* – far too little is known about: user needs and preferences when it comes to the presentation of DVB-Subtitles, if and how users with varying hearing capabilities actually use pre-recorded and live subtitles.
- *One size does not fit all* – Europe is a culturally diverse continent where different access solutions have emerged in response to diversity. A good example is inter-lingual communication, i.e., making TV programmes in a foreign language accessible to viewers. Parts of Europe, such as the Nordic region and Benelux, have more than fifty years' experience of providing inter-lingual subtitles whereas most of the rest of Europe uses dubbing or lecturing. The role and perception of intra-lingual subtitles as an aid to those with a hearing impediment is clearly different in, say, Denmark where everyone is used to subtitles for inter-lingual communication and Germany, where subtitles are unfamiliar and are almost exclusively associated with hearing impediments. The DTV4All has to keep this diversity in mind and offer a strategic toolkit for access service implementation that is sensitive to cultural diversity and national circumstances.
- *Stakeholder perceptions of priorities are evolving* – when work started on proposal for the DTV4All project the consortium members were familiar with the priorities of the special interest groups representing those with physical and cognitive impairments but since mid-2008 changes have been noted in these priorities. An example of this is the interest in Spoken Subtitles as a complement to Audio Description. Feedback from the UK and Denmark indicates an increased interest in scaling up services for those with visual impairments in the direction of spoken subtitles for non-fiction and AD for TV fiction and drama.
- *The economic downturn and the need for more efficient workflows* – although producing access services accounts for a relatively small part of television production budgets, in the current economic climate there is pressure from within the organisations delivering digital television services to improve the efficiency of workflows and playout systems associated with access service provision and to contain their costs at a time where particularly those dependent on advertising revenue are struggling to make ends meet.

The planning and execution of the Pilot has had to be responsive to the above issues.

In this report the DTVAll project team aim to explain the following:

- what work was planned by:
 - The broadcaster partners in the project
 - The universities contributing to the work of the project
- what work has been done to date and the rationale for any changes or additions to the plan for the pilot since 1 July 2008
- what preliminary findings have been made
- which preliminary conclusions are emerging
 - how these can be incorporated into later outputs of the DTV4All project that can have a strategic impact on the efficiency and effectiveness of access service provision in the short term

2.2 Who this report is aimed at

- The DTV4All project partners
- The European Commission
- Other interested stakeholders

2.3 How to read this report

The *Executive Summary* contains a resume of work on the pilot from 1 July to 31 March 2009.

Sections 3-6 go into depth on the planning of the pilot and the work done on it in the period to 31 March 2009.

Section 3 gives a description of the project partners and their collaborators that are responsible for the evaluation of the mature access services to be piloted and the evaluation methodology they have adopted.

Section 4 explains any adjustments and refinements that have been made to the plan of the pilot and the rationale for them.

Section 5 covers preliminary findings and is primarily concerned with the validity of the pilot evaluation methodology but also reports some specific findings.

Concluding remarks are made in Section 6.

The *appendices* contain detailed background information on work referred to in sections 3-5.

Although this report only covers work on the Pilot to the end of March 2009 it will be followed by a second report covering the period to the end of August 2009 and a final report will be delivered in early 2010. This report will necessarily focus on process issues rather than results, but preliminary findings are already being fed into standardisation and dissemination work both within the access service community and the European Broadcasting Union.

3 What was originally planned

3.1 Broadcasters and the mature access services evaluated

3.1.1 DR

Evaluation of AD services

DR planned and carried out user consultations with *Dansk Blindesamfund (Landsforening af blinde og svagsynede i Danmark)* representing the blind and visually impaired on 3 December 2008 and again in March 2009 following the introduction on an experimental basis of AD for the TV series “Sommer” in the autumn of 2008 which was followed by regular AD support for the series “Krøniken” and “Album”.

The main conclusions of the user consultations were surprising. While the blind and visually impaired appreciated the existence of AD in connection with high profile drama series such as Somer [Summer] there were comments:

- On the marketing of the service itself and
- On the relative importance of drama vis a vis other genre

It was generally felt that spoken subtitles using speech synthesis in the TV receiver might be a higher priority than AD for genre like news, especially if this meant that the same budget could lead to the production of a greater number of hours of spoken subtitles. This conclusion is currently being discussed with the regulator, the Danish Ministry of Culture.

Evaluation of subtitling

DR has been running both in vision and opt-in Subtitling for the Deaf and Hard of Hearing (SDH) subtitles using Teletext for more than a decade, and using DVB-Subtitling since April 2006. A major user survey was conducted on subtitles in 1996 which forms the basis of policy and production priorities to this day. Live subtitling using re-speaking was introduced in 2006. DR’s public service agreement calls for an increase in subtitling by the end of 2010 so that nearly 100% of DR programming on DR1 and DR2 will have subtitles. This means a significant expansion in live subtitling for news, sports, current affairs and events. This area was identified as a major access service for evaluation.

Television programmes are increasingly broadcast with subtitles not just for inter-lingual communication (translating one language to another) but also for intra-lingual communication (to improve the accessibility of the soundtrack for those with hearing impairments - SDH).

In a number of European countries the trend is towards 100% SDH subtitling.

Providing live subtitling is a challenging matter. It normal requires either a stenography set-up where a team of up to six subtitlers take turns to manually key-in small chunks of the required subtitles or the subtitler listens to the programme and dictates the subtitles into a speech-to-text system (“re-speaking”). Note that in TV3 of TV de Catalonia they don’t use stenography but normal qwerty keyboards. Regardless of how the subtitles are produced, there is a delay in relation to the programme of six to ten seconds. In The

Netherlands public service broadcasters introduce a delay in live programming (on cable) of 10 seconds, allowing for the subtitles to be shown in synch with the content.

There are several problems with live subtitling: the quality of the subtitles themselves: the extent to which compression and the re-speaking system lead to semantic or factual errors, the way they are displayed on screen (*roll on* rather than *pop up*) and the delay in showing the subtitles in relation to the video and audio to which they refer.

Three international service providers that use dictation solutions, IMS, ITFC, and Red Bee, claim up to 96-97% content accuracy for widely spoken languages such as English. There are alternative solutions available for widely spoken languages such as English, but options for many less-widely-spoken European languages are either limited or non-existent.

Broadcasters offering live subtitles and using re-speaking systems report that there has been criticism of the quality of live subtitles, primarily semantic errors (misspellings, incongruous and omitted words).

Live subtitles also differ from their pre-produced counterparts in that text is presented as soon as it is available and the presentation conventions are different. The delivery rate may vary making demands on the viewer's reading speed.

The focus of the quality debate has been on semantic errors. Less emphasis has been given to presentation and the significance of the delay. There seem to be no known formal evaluations of live subtitling in the research literature, only informal feedback from call centres and broadcasters.

News has been chosen for study in the pilot by DR. The reasons for this are that:

- The main news at 18:30 already offers live subtitles
- News is one of the critical genres in terms of its complexity involving a mix of pre-produced and live subtitles

Other live genre, sport, factual, general election coverage, events, may well represent slightly different challenges for live subtitling, as they contain a higher proportion of spontaneous speech at a higher delivery rate. Even so, results for a news programme will be indicative of how users perceive live subtitles, and there is always the option of conducting additional studies on other genre if this is deemed necessary.

This is the rationale for an exploratory study with 30 Danish viewers with a range of hearing impairments to assess the relative importance of three issues:

- Semantic errors
- Presentation differences between live and pre-recorded subtitles and
- The delay between the programming and SDH subtitles.

The methods to be used are described in Section 8 Appendix 1.

3.1.2 RBB

RBB planned to test both DVB-Subtitles and Clean Audio technology from January till December 2009 with a user group of 50 hearing impaired and deaf users.

DVB-Subtitling

Even though DVB-Subtitling is basically a mature technology which has been in regular operation in the United Kingdom (BBC) and in Scandinavia for some time now, this technology is completely new to RBB. Up to date only ZDF has been broadcasting DVB-subtitles in Germany, only since October 2007 and only via DVB-S. In DTV4All RBB wanted to test this technology both from the broadcaster perspective, i.e., technological feasibility, maturity of existing solutions, costs, and from the end user perspective, i.e., optimum design under the given design options which are much more favourable than in ordinary teletext subtitles.

The plan was to broadcast DVB-subtitles (bitmap-based) via DVB-T from January till October 2009 simulcast with existing teletext subtitles. The users were to be equipped with a Set-Top-Box that they could keep once they had successfully completed the test running for 10 months. Testers would receive a simple questionnaire once a week either by mail, email or fax and would have to evaluate the subtitle design in each testing week. One or two design variables of the subtitles like font size, type of font, and background like outline, shadow (stripe/banner) and box, all in different grades of transparency, were to be changed each week. These tests were to be performed with a view to introducing DVB-subtitles via DVB-T (in the first instance) into regular operation. The idea was to have a validated running technology at the end broadcasting DVB-subtitles that have the optimum design according to the evaluation of the target group.

Clean Audio

The plan was to use November and December 2009 for testing mature clean audio technology with the same user group that was testing the DVB-subtitles. Certain pieces of content (films) were to be prepared to offer an extra audio channel with “clean audio”. The users were to evaluate this with a specific questionnaire.

3.1.3 Red Bee (UK)

Red Bee offered to contribute to the Pilot by providing a summary of work done in the United Kingdom to raise awareness of access services such as Audio Description. Section 4 includes a summary of the work done by Red Bee and the BBC as part of an OFCOM-funded study on access service awareness.

3.1.4 TVC (TV3)

Televisió de Catalunya will broadcast, consolidate procedures and workflows, and test functionalities during a 12 month period, for two pilot accessibility services:

- Audio Description
- Signing language

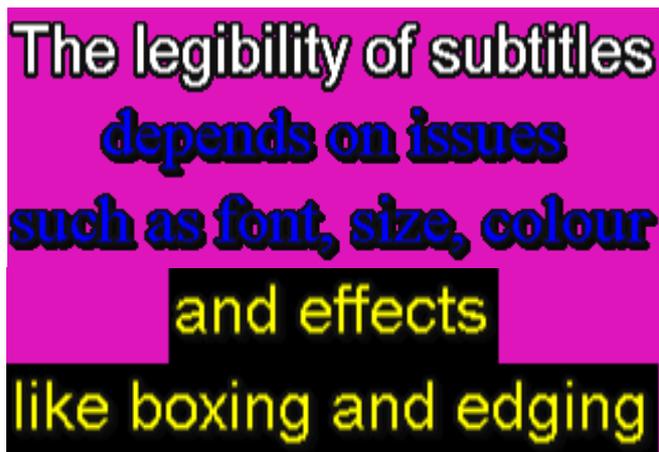
The content for AD will be broadcast over DVB-T and will be accessible by all DVB-T receivers except for a few integrated receivers that hide the services as in another language than the “preferred language” in the receiver firmware. The signing language is to be broadcast as a fixed window visible in analogue broadcasts as in DVB-T.

The services will be tracked along 2009 through specific tests taking advantage of the already established communication with users associations, and a pool of individuals that regularly maintain contact with and give feedback to the accessibility department. These two accessibility services have already started and are going to reach a stable status of exploitation and growing.

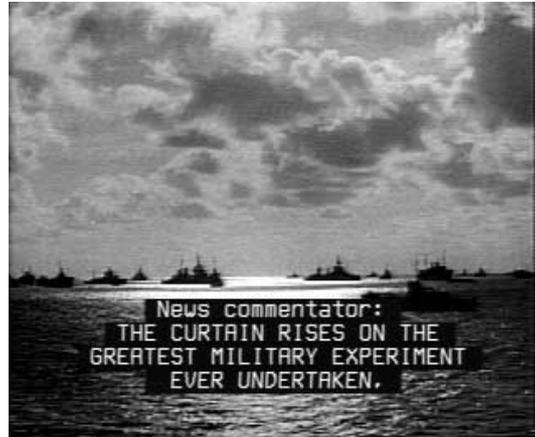
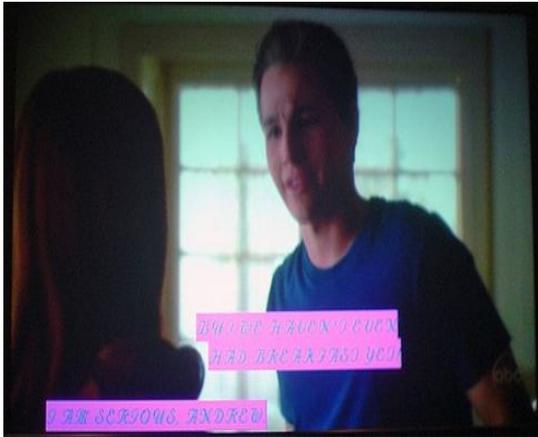
3.2 Universities and the mature access services evaluated

Universities had planned to check two mature services: Subtitling for the Deaf and Hard of Hearing (SDH) and Audio Description (AD).

One of the most common problems often attributed to **Subtitling for the Deaf and Hard-of-Hearing (SDH)** both as a research discipline and as a service for the viewers is the absence of a) a common standard within and across countries and b) active consultation with the Deaf and hard-of-hearing audience. The study proposed in this project is an attempt to resort to the latter as a means to explore the possibility of achieving the former. With this objective in mind, researchers from Universiteit Antwerpen (Belgium), Università degli Studi di Napoli Federico II (Italy), Hellenic Open University (Greece), Roehampton University (UK), Warsaw University (Poland), Copenhagen University (Denmark), and Universidad Autònoma de Barcelona (Spain) have carried out experiments in seven different languages to examine the viewers' preferences with regard to four technical parameters of SDH:



- Font and size



- Position



- Subtitles may appear at the top or bottom of the screen



- Character identification and identification by colour



· Identification by tag: (Eleanor) (Frederik) are the speakers



· Identification by superposition on characters



· Context information through tags [sighs] [groaning]



- Context information through icons

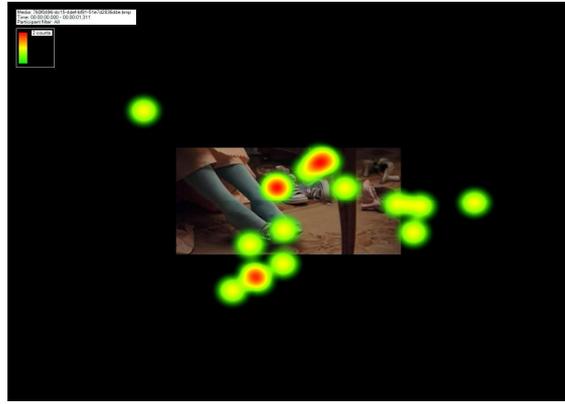
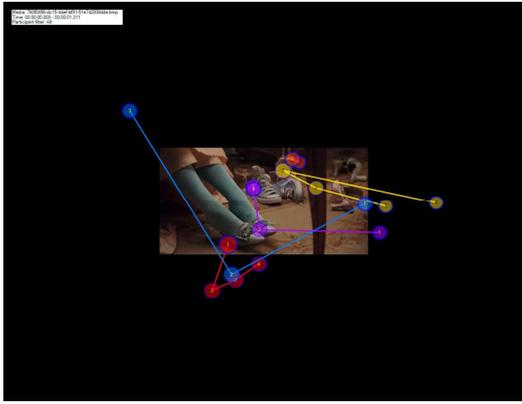
In every case, the audience was made up of three different groups: Deaf, hard-of-hearing and hearing viewers.

Film excerpts were subtitled, providing the participants with different variables for each of the four parameters. A number of questionnaires (see Section 9 Appendix 2 for a sample of the English and Spanish questionnaires) were drafted, adapted and localised to the social and cultural reality of each country and then handed out. They contain specific questions about readability, preferences and conventions, always taking into account the viewers' background (age, hearing capacity, education, etc).

For tests regarding audio description a project has been set up with strict considerations. Twenty participants, the viewers of the film, are female students of translation, consistent with the fact that in most translation faculties across Europe the percentage of female students is higher than that of the male. In addition to this homogeneous gender profile, a homogeneous age profile has also been built into the research design (18 to 23 years). For the research being carried out in all institutions (in Belgium, Greece, Ireland, Italy, Spain, Poland, UK and the USA¹) the data has to be gathered and translated into English by April 2009.

AD scripts are written in long hand then typed into a Word document which is then translated into English. Different templates have been created in order to gather relevant information. Other AD tests are being carried out with eye-tracking technology, whose primary aim was to identify priorities in perception by examining areas of interests in film viewing, to test the validity of the findings, as well as that of current AD practices, on different types of visually impaired viewers.

¹ The reason for including the USA is to be able to take into account all the results provided by the 1980 experiment set up by Chafe (Chafe 2002).



· Two eye-tracking experiments with four viewers watching the same image

4 What has been done to date

4.1 Broadcasters and the mature access services evaluated

4.1.1 DR

AD evaluation

This is mentioned in section 3.1.1. A round of consultations was completed, leading to the conclusion that there is interest in increasing the proportion of programming for which there are access services for the visually impaired, even if this means introducing audio subtitles at the expense of AD, if this would lead to an increase in service coverage.

Live-subtitling

Work until the end of March 2009 has concentrated on the following:

- 1) Putting together the sample user group with the aid of a partner with the necessary audiometric background data in order to assure that the sample is representative of Danish users with hearing impairments
- 2) Developing and validating a holistic tool to conduct user tests.
- 3) Developing and validating the test procedures.
- 4) Selecting personnel to carry out the audience research on live subtitles.
- 5) Improving the statistics for access services in Europe
- 6) Linking this work with dissemination activities to ensure a multiplier effect across Europe.

As of early April 2009, the following was completed:

Re. (1) the framework for selecting the sample user group of 30 viewers was completed with the aid of Oticon, a leading manufacturer of high-end hearing aids based in Denmark. The aim is to enlist their support to DR by asking a sub-set of their user panel whose hearing impairments are known to take part in the DR study. By the deadline, this commitment was still not in place.

Re. (2) after consultations with the DTV4All Project Manager, it was agreed in February to use some of DR's resources to produce an open-source evaluation tool for this and other access services. The tool was adapted from existing software for video editing and the first version was complete during this period. What remains to be done is to validate the tool and its reporting facilities.

Re. (3) the proposed evaluation procedures have been submitted to the head of DR Audience Research Lars Thunøe and have been approved in general terms. The details of the procedures are being vetted by Sofie Scheutz, head of qualitative research at DR

Re. (4) two graduate researchers with experience in usability testing (Anni Randers, Marguerite Johnsen) have been selected to deal with the implementation of the tests. Marguerite has the additional advantage in that she has a hearing impediment in both ears and is a regular user of audio link when watching television.

Re. (5) and (6) a meeting was held with Frans de Jong and Edward Wilson of the EBU on 12 March 2009 to discuss DTV4All and links to ongoing standardisation and dissemination work in the field of access services. Detailed minutes of this meeting are available.

The main conclusions include the following:

- The evaluation tool being developed by DR is of general interest to broadcasters and universities working on holistic measures of access and interactive services for digital television
- The EBU is willing to provide space to document the tool and make references to its web location in a recognised open source forum provided that there are no competition issues (i.e. commercial software doing the same job – this is not the case)
- The EBU included a slot in their Television Summit on 8 May 2009 in Lucerne for RRB and DR to present their current work and explain the implications of the open source software tool. The results of this gathering will be reported on in detail in the next progress report.
- It is imperative that the EBU come up with solid statistics on the trends in access service provision among its members in Europe. This would seem to be an area where DTV4All and the EBU have related interests in improving the data for Europe.
- Some existing draft recommendations for DVB receivers seem to have overlooked details to do with AD and subtitling. At the 12 March 2009 meeting with the EBU, the drafts were reviewed in detail and corrective action was taken.

4.1.2 RBB

DVB-Subtitling – preparations and start

As to DVB-Subtitles the original plan was kept to and realised. The only change was that the test operation and the start were delayed by one month, mainly due to delays in receiving the technology required for broadcasting the subtitles.

After a meeting with the disability organisations at RBB in October 2008, recruitment of the testers together with the disability organisations started. The user group recruited by the disability organisations was to be representative and to encompass members from Berlin and the area around the capital. The users repeatedly received detailed information from RBB about DTV4All and the testing procedures. All information was cross-checked with the disability organisations for clarity and understanding. Dates were fixed for installation of the devices at users' homes and sign language interpreters for the installation were hired. On 5 February 2009 broadcasting of DVB-T subtitles started. By 15 February 2009 each user had a set-top-box at home and was equipped with a specific manual written for them and all necessary instructions. Test Week One started on 23 February 2009. By then, a methodology had been developed and the questionnaire was ready. This questionnaire was slightly adapted after two test weeks as the first results showed that some of the questions were too ambiguous and not clear enough for what RBB wanted to know.

In parallel, from September 2008 onwards, “pre-testing” of technology started at RBB. A decision had to be made about which technology to use for DVB-subtitle production and about which set-top-box to give to the testers. For results please see chapter 5.1.3.

DVB-Subtitles Test-Method

The technology chosen for transcoding offers a number of parameters, such as:

- **Font**; choice: basically any font that can be stored on the machine. As the text is sent as bitmaps, the font does not have to be available on the DVB-T set-top box.
- **Font size**; the size is only limited by visual parameters. Too small would be illegible, too big could mean that the text doesn't fit on the screen so that, in the worst case, parts might be lost
- **Appearance/Layout**; there are a number of options for the appearance of the subtitles:
 - **Text only**
 - **Text with border (outline)/shadow**
 - **Text with a box**, i.e. text with a dark background. The background is generally black, but the transparency can be adjusted before broadcast. The "box" only "highlights" the text; if there is no text, there is no box; if the text is short, the box is short; if the text is shorter in one line, the box will be of different length in the two possible text lines ("stair effect")
 - **Text "with longest line box"**, i.e. a box where both text lines have the length of the longer line. If the text of one line is longer than that of the other the box will have the same length in both lines.

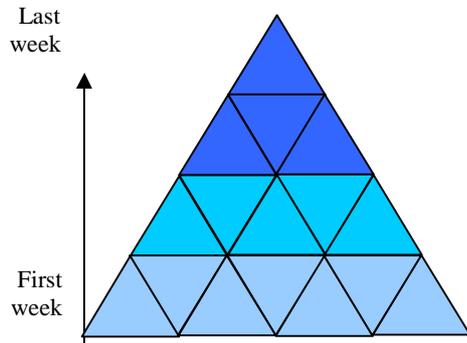
Approach

The selection of appropriate parameters is essential for the success of the test for various reasons. The key objective is to find out which parameters best suit the interests and needs of the users. A mathematical approach to join every parameter with every other to enable maximum comparability might seem useful for statistical analysis of the acceptance of every single option per parameter. However, offering a specific option again and again even though a majority of users repeatedly stated that they dislike this option, might cause test users to quit the testing community, a risk that might ruin the comparability of test data.

It was therefore decided to plan the tests only for a reduced number of weeks in advance, then adapt the selection of options according to the feedback received during the first test weeks, so that, eventually all "risky" options would be banned.

Thus we would come to a reduced number of optimum choices between which the editors and other decision-makers on the broadcaster side can decide.

Week by week, with every new questionnaire, the number of acceptable combinations (e.g. Arial, 40, Box, 0 Transparency) will be reduced by reducing the number of acceptable options (text only, Box within safe area, etc.).



· Optimisation Pyramid

The optimisation pyramid above shows how through the steady reduction of the number of options the tests will lead to a manageable number of combinations towards the end of the testing phase.

Test Plan

The general plan planned to test 5 different fonts and apply 5 parameter combination sets to each of the 5 fonts chosen.

1	Arial	32			Box longest line, Transparency 0
2	Arial		40		Box longest line, Transparency 160
3	Arial			44	Box, Transparency 160
4	Arial		40		Box, Transparency 120
5	Arial		40		Text with border/shadow
6	rbb Interstate	32			Box within Safe Area, Tr 0
7	rbb Interstate		40		Box within Safe Area, Tr 160
8	rbb Interstate			45	Box within Safe Area, Tr 224
9	rbb Interstate		40		Text with border/shadow
10	rbb Interstate		42		Band, Transparency 120
11	ARD Font	35			Box, Transparency 140
12	ARD Font		43		Band, Transparency 120
13	ARD Font			48	Band, Transparency 80
14	ARD Font		43 ¹		Box, Transparency 0
15	ARD Font		43 ¹		Text with border/shadow
16	Tiresias	36			Band, Transparency 0
17	Tiresias		45		Band, Transparency 120
18	Tiresias			50	Band, Transparency 140
19	Tiresias		45 ¹		
20	Tiresias		45 ¹		Text with border/shadow
21	Segoe Print	33			Band, Transparency 0
22	Segoe Print		41		Band, Transparency 120
23	Segoe Print			46	Band, Transparency 140
24	Segoe Print		41 ¹		
25	Segoe Print		41 ¹		Text with border/shadow
26	Preferred Font	Preferred Size			Position TOP

Table 1: Test Plan Layout (state: Test week 6)

Explanations:

Each of the five selected fonts was to be tested:

- a) In at least three different sizes,
 - The sizes are different per each font, as the “same size” e.g. 10 pt. may appear quite different on the screen. Therefore, the sizes were chosen so that the actual size on the screen is the same
 - Much greater than weeks 4-5 per font were left open to wait for feedback and adapt accordingly.
- b) With up to five different appearances,
 - “Box within Safe Area” is labelled as a variant of “Band”, so that each font would be tested with variations of “box” and “band”

Analysis:

This approach causes some difficulties when it comes to analysing the test results:

1. In the questionnaires the appearance was to be evaluated with only one check box, although it actually comprises two different parameters: the background type (box, band, and border) and the transparency of this background. Ideally, this should have been tested in all possible combinations in order to get separate results for the individual parameters. This, however, would have deterred numerous testers who tend to feel bored and under-challenged if they do not see the differences between test weeks, etc.
2. Some options have been omitted relatively early (e.g. “text only”, “band”) as the feedback in the comment fields soon showed clearly that these options would not be accepted. If these options cannot be replaced by others, but are omitted, the set of combinations will have to be adapted. Otherwise, either there would be repetitions of former combinations or sets would have to be omitted so that this font would be tested for less than five weeks.

To react to the users’ feedback as soon as possible, ideally almost instantly, complicates the analysis of the tests but is absolutely necessary as experience shows that the test users tend to feel not taken seriously if their comments are not responded to directly.

It is already difficult, if not dangerous, to ignore their more general recommendations, even though the introductory letters and talks about the test and its objectives stated clearly that the test will not involve editorial issues such as the number of programmes with subtitling or the ways the texts are written and presented. The general comments, however, are replied to instantly in personal letters / mails. They are also forwarded to the RBB subtitling editorial team which is very happy to hear directly what users may want or not want in terms of subtitle content.

According to regular analysis of the quantitative and qualitative feedback from the questionnaires, those options with the lowest marks and the most convincing contra-arguments will be taken off the list. Consequently, these options will be replaced by other comparable parameters.

Quantitative results, i.e. marks between 1 and 5 are being analysed regularly. This analysis compares the best marks (1-2) with the worst marks (4-5) and also considers the delta value, i.e. the difference between the two opposing value groups. The reason for taking the delta value into account is the fact that the most praised combination of parameters may as well be among the most discussed or even the most disapproved of. As avoiding dissatisfaction is more important for a Public Service Broadcaster than fancy design, combinations with even a few remarkably negative marks have to be handled very carefully.

Clean Audio – general remarks

Here, the plan has changed. During the first months of the project it turned out that more development work than had been envisaged is still needed to make Clean Audio a mature access service. Please see Chapter 6 of Deliverable D3.2 for an overview, especially concerning IRT's development work on Clean Audio in DTV4All. Therefore, Clean Audio is regarded as an emerging service in the project and dealt with in WP3. Still, RBB intends to take advantage of its user group recruited for the DVB subtitle field test for testing Clean Audio on a larger scale than just a laboratory test. The idea is to provide those users in the group that are hard of hearing (about 25) with a DVD offering different versions of Clean Audio for testing. Ideally, all the 25 hard of hearing people in this group will be involved. Preceding laboratory tests, however, will show whether the clean audio services are suitable for all grades of hearing impairments.

4.1.3 Red Bee (UK)

The awareness campaign and evaluation involved interviews with a representative sample of adult television viewers in the UK. The study was carried out by GfK and was funded by OFCOM. Red Bee and the BBC were involved in the preparation and transmission of spots for the campaign itself.

4.1.4 TVC (TV3)

TVC has established a feedback service with users, both at association level and with individuals, who write regularly offering their opinions regarding the service. Since 2006 all comments and suggestions made by users has been entered providing, in retrospect, a valuable source of information.

4.1.4.1 - AD

The aim of the test is to validate AD strategies and decisions concerning thorny issues and dilemmas which are part of any script process of AD.

The tests TVC shall take on board will have the following stages:

- 1) Classification of thorny issues at two levels:
- 2) From a technological point of view
- 3) From a script creation point of view
- 4) From the taxonomy of issues which will result from the previous point a questionnaire will be drafted. This questionnaire will request feedback on the issues identified.
- 5) Issues under analysis will be created in one or as many films as there are needed.

- 6) Questionnaires will be sent to users associations (ONCE and Associació Catalana de Cecs) and also to a control group (5 blind users).
- 7) Film/films will be broadcast and associations and users warned.
- 8) Questionnaires will be analysed.
- 9) Results will be compared with existing guidelines such as UK OfCom or Spanish UNE.
- 10) When some results show a significant deviation from the accepted norms, further tests will be created under laboratory control and methodology.

Number of users: 5 blind users as a qualified control group. Otherwise, the number of users is unknown because it will be the users associations that will be contacted and who will agree with TVC the format of tests.

Name of users associations: ONCE and Associació Catalana de Cecs

The model of AD services in TVC

The workflow for the production of Audio Description (AD) is similar to that for subtitling. TVC has developed with the collaboration of an external SME (Anglatecnic S.L., Barcelona) the necessary software for the production of AD.

The programs to be audio described are copied to MPEG 4, with time code, and accompanied with the text dialogue script and sent to the recording operators-editors of audio commentaries and speakers (they are the same person). Such workers are often experienced operators of subtitling. They have passed a selection for voice quality and a training period on operation. There are continuous discussions about the style of the descriptions, for example, when there are several things to describe about what the priorities are.

In the last auditions of the material, the operator gives a value for the volume of audio description in relation to the loudness of the soundtrack of the program. The possible levels of the soundtrack can be attenuated giving several fixed levels. The experience until now has lead to do a broadcaster mix with 3 levels of the soundtrack: no attenuation, 1.5 Db of attenuation, and 3 Db of attenuation. Audio Description remains at a fixed level. This is the current practice. In the past TVC assessed 3 Db and 6 dB attenuations. This is not discount implementing more levels of attenuation for the tests, or measuring continuously the loudness level of the soundtrack (Dolby LM 100) and using this data for the control of the soundtrack volume in relation to the AD signal to ensure the intelligibility of the AD in a more accurate way.

The result of the production of audio description is a *compressed file* containing:

- The audio descriptions (a long list of MP3 audio files)
- The time-code points where the descriptions are to be broadcast on air
- The volume value for each piece of description
- The metadata descriptive of program title, number of ident for archive purposes, etc.

This product goes to the Continuity Department and is automatically stored in a server ready for *automatic and manual check-ins*, and automatic broadcasting in parallel with the video program in the moment of play-out. It is at this moment when the system gives the volume value to the AD and creates the DVB –T transport stream with the AD channel codified in an optional secondary audio channel.

The things to be tested are the convenience of the workflow for efficiency, functional security and adequacy to the needs of the blind and visually impaired people.

4.1.4.2- Sign language

The content chosen for test by TVC is a group of 3 news leads summarising the lead stories at 3 moments of the day: 9:00, 10:00 and 18:00 hours. The duration of the leads is 5 to 10 minutes depending on their contents.

The translators to signing language are a pool of 3-4 persons with very close contact with the deaf associations, so they receive continuous close feedback from the users. The format of presentation is a fixed window of approximately 1/5 of the surface of screen. The signing signal is produced in a secondary studio with a neutral background. The colours of the translators' dresses are recommended to be dark to favour the visibility of hands and face of the translator. The signing video is superimposed on the main signal in a rectangular window in the main continuity video mixer.

There are meetings with some regularity in one of the associations for viewing in a group the signed news. This gives an idea of the possibilities for social contact with the users. Feedback is gathered from users groups primarily in an informal way but also in more systematic ways.

Tests are to be made with groups of 10 – 15 users, asking for preferences through comparison of pairs of videos and putting questions to the users to check their understanding, recall, and satisfaction with the format and the service. If the necessary equipment is available, additional tests under laboratory conditions can be made in collaboration with UAB. This is to be planned in the coming weeks with UAB.

4.2 Universities and the mature access services evaluated

The universities contributing to the work of the project have started drafting and translating questionnaires for SDH. They have also started establishing contacts with users associations, and deciding on the material which will be used for the eye-tracking tests. Subtitling this material in the different formats and languages is also well under way. Some tests have been done in order to ascertain the validity of test material and allow for adjustments to be done before embarking in a pan European experiment.

For AD the first deliverable is finished. The objective was to present the audio describing practices of three different companies which would shed some light on issues such as:

1. **AD services:** AD services provided so far and the companies' future prospects for AD
2. **The process of AD:** how AD is carried out.
3. **Software:** presents the specific software used for AD
4. **Fees:** if disclosed, the fees for ADs are reported.
5. **The translation of audio descriptions:** a new alternative for AD generation
6. **Technical aspects**

The repost was drafted by UAB based on input from:

- Televisió de Catalunya (TVC), from Catalonia (Spain);
- Red Bee Media, from the United Kingdom,
- Bayerisches Rundfunk, from Germany.

The following tables show the results at a glance:

(a) AD services: key dates and figures

TVC (ES)	1989: first broadcast of an AD film in Catalan 90s: occasional AD 2006: regular AD	2007: 81 hours of AD 2008: 111 hours of AD 2009. Aim: 160 hours of AD (4% of TV3 programming)
Red Bee (UK)	2000 (as part of the BBC): initial AD tests 2004: detailed records	2004-now: approx. 2,600 hours of AD/year AD business model driven by regulation (10% of programmes on certain channels)
BR (DE)	1997: first AD at BR	Total: 350 hours of AD BF: 4% of prime time programming has AD

(b) The process of AD

TVC (ES)	Creation of AD script by script writer → Quality check by another professional → Voicing (by the initial writer) → Final quality check
Red Bee (UK)	Creation of AD script by script writer → Voicing (by the initial writer or another professional) → Final quality check if two script writers or a trainee work on a project
BR (DE)	Creation of AD by teams of three people (2 sighted + 1 blind) → voicing by a professional voice talent → Final quality check

(c) Software and technical aspects

	Software	Technical aspects
TVC (ES)	Specific software	1989: stereo channels analogue: NICAM DUAL DTV (currently): different audio track on DTV
Red Bee (UK)	ISIS 5.5.6	Depending on the broadcaster. BBC: DTT/ receiver mix, DSAT/broadcaster mix.
BR (DE)	DVD player + Word	Broadcaster mix

(d) Fees

	90 minutes (average cost)
TVC (ES)	900 euros (paid to free-lancers)
Red Bee (UK)	Confidential information
Ofcom, 2004	1,070 euros (in 2004)
BR (DE)	4,000-5,000 euros

5 Preliminary Findings

5.1 Broadcasters and the mature access services evaluated

5.1.1 DR

Audio Description and Audio Subtitling

- Work done within the project DTV4All indicates that the DR approach to AD, while being of high quality from a user perspective will run into both production and transmission issues in the next 2-3 years due to the relatively high production costs and the need to transmit an additional set of stereo tracks at 256 kbit/sec. Awareness of other solutions from Red Bee and TVC opens up a discussion on different work flows and a migration from broadcaster to receiver mix.
- The user consultations from December 2008 onwards indicate that further clarification of the roles of Audio Description and spoken subtitles would be beneficial, as this would allow an increase in the services for those with visual impairments without any major increase in budget.
- In the medium to long term, the use of speech synthesis chips such as that used in the RNIB pilot project could accelerate the production and take-up of spoken subtitles if the subjective quality of such services meets user expectations.

Subtitling

- There is a general trend in the direction of almost 100% provision of subtitles for the deaf and hard of hearing in Europe.
- Work done in the first quarter of 2009 has highlighted gaps in European knowledge about live subtitling services and how these are perceived by users with varying degrees of hearing impediment.
- The project has also identified gaps in the availability and quality of re-speaking systems that can produce live subtitles, in particular in countries with languages that are less widely spoken.

5.1.2 RBB

Technical Feasibility

Results of technology pre-tests

The text in the following is based on detailed documents in the German language covering all major technology aspects which are available to interested parties if desired.

There are two ways of producing DVB subtitles. You can produce them with an original DVB subtitle authoring and production system and deliver them directly to the DVB subtitle encoder. The alternative is to use an existing authoring and production infrastructure for teletext subtitles and transcode these in real time. This latter option of transcoding existing teletext subtitles into DVB subtitles was RBB's choice for the DTV4All tests. The most important reasons for this decision were:

- Minimizing the costs

- Being able to use archived material subtitled with conventional teletext subtitles
- Continue to use our existing teletext authoring system
- Being able to author and playout both types of subtitles when simulcasting on analogue and digital channels

There was a long list of requirements for deciding on a transcoder system, the most important were these:

- Clean graphic presentation of the subtitles
- Keeping the original teletext formatting
- Maximum compatibility to existing end devices (set top boxes)
- Maximum variety of design features (e.g. graded transparency)

In the end RBB tested three transcoder systems with 13 set-top-boxes. These included popular older boxes and the most popular new boxes on the market in Germany. The objectives of this test was to find not only the best production system but also the “best end device”, a user friendly device which showed the DVB-subtitles broadcast by RBB in a problem free manner. This was to be the device to be given to the 50 RBB testers as a test device. For the tests, the different transcoder modules received a pre-produced test transport stream encompassing the teletext data. A DVB-T modulator then broadcast a DVB-T signal including DVB-subtitles to the different end devices. The result shown by each of the receivers with each of the transcoder modules was examined according a number of criteria like functionality, usability or readability. Both conventional screens and LCD screens were tested. All major different layout options were tested like stripe/banner, box, and outline/shadow with different fonts. Given the fact that we are talking about “mature technology” the test results were quite disheartening. In the end, only four set-top-boxes were found to bring satisfying results with the best transcoder module that was finally chosen by RBB. One of these four boxes turned out to bring slightly better results than the others and was chosen to be handed out to the testers as it was felt to be important for the testers to have a problem free depiction of the DVB-subtitles.

The most obvious problem was depicting a transparent background. Most boxes had problems with depicting that and showed corrupted letters without any background at all. Another problem was displaying a box or a banner/stripe, in some instances the TV screen went completely black and only the subtitles were shown. The “outline” mode was the best option for most boxes, meaning that the subtitles were readable but in some instances not shown as envisaged in the “outline” mode but in boxes or banners. Although, this mode was not displayed in the correct way in the case of many boxes at least the subtitles could be read when shown not in “outline” but in a box or other different ways. RBB will get back to the manufacturers of the set top boxes and recommend that all but one of the 13 boxes should be improved.

In the end, one transcoding system, the FAB Subtitler, and one testing device, the Philips DTR 220, were chosen.

Problems with subtitle trans-coding

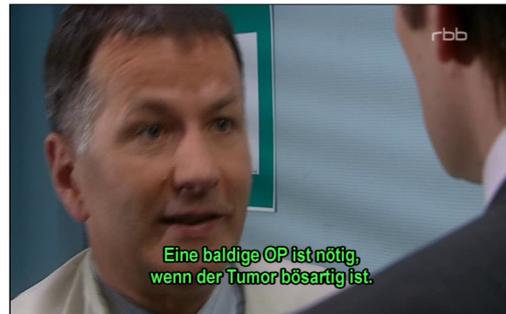
1. Three line subtitle problem

During the process of transcoding teletext subtitles to DVB subtitles RBB experienced a few difficulties. One immediate problem was the faulty presentation of triple-space subtitles. Triple space subtitles are only used in one regular programme originating from one of RBB's partner broadcasters. However, as this is the main nationwide news programme at 8pm people were considerably irritated. In teletext the triple-space text is realised by showing the first two lines as one teletext page and by instantly adding a third line as another page, but without removing the former one. The DVB subtitle transcoder did not accept the "do-not-remove" tag and deleted the first two lines once the third one was to be presented. Thus, only the third line was shown and most of the subtitle information was lost for the viewer. Additionally, there was a text positioning problem. The first two lines were presented in the middle for a few milliseconds, but the third line removed the first ones and appeared left-aligned, so that the viewer experienced not only fragmented but also jumping text.

The transcoder manufacturer was able to avoid this faulty presentation behaviour by changing the software into accepting the "do-not-remove" tag. Currently the triple-space DVB subtitle presentation is much better, both the alignment and the entirety.

2. "Jumping subtitles"

Another faulty pattern that should not have occurred at all was what we call "subtitle jumping". The picture below shows one example of this: the teletext subtitles are left aligned both as a text block and concerning the single lines, the DVB subtitles are centred and the lines are also centred. Something like this happened in different text positions. However, there was not a regular pattern at first sight, it all happened in a very arbitrary way sometimes the positioning was right, sometimes not, so the impression is of "jumping text" which is very irritating. Nearly all the different text positions of the teletext are presented in the wrong way in such an "irregular manner".



Of course the basic requirement was to have DVB-subtitles that looked exactly the same in terms of position as the original teletext subtitles. What was done to solve this problem was to analyse the teletext input going into the transcoder system. Seven such faulty positioning patterns were identified. The teletext authoring / editing system uses space characters to position the subtitles on the screen. Transformation of this alignment data is not really working inside the subtitle editing software. The DVB ST Transcoder cannot interpret several sequenced space characters. The solution is that the software reads and considers spaces and can calculate the subtitle position within the screen width. The manufacturer is now working on that problem. A possible forced centre alignment as a work around was not a satisfying option.

End user Feedback

As mentioned above: DVB subtitles offer a lot of different design options for broadcasters. In RBB's DVB-subtitle tests RBB wants to find out the optimal design from the end user point of view. The users are marking the individual design parameters and also the combination of these parameters. The subtitle design of each respective test week on the whole is marked using marks from 1 (very good) to 5 (dissatisfying). Users can also comment freely on the subtitles and their marks. The different parameters are described in detail in section 4.1.3.

At this early stage of the project, i.e., the reporting period until 31st of March 2009 which is only test week five of 36 test weeks of the RBB field tests, only a few tendencies can be described very cautiously. Up to the end of March, three font sizes, two font types, Arial and RBB interstate light, and five different background types have been tested. For the average font size, the given teletext size (40) seems to be judged better than larger or smaller sizes (87% of users rate it good or very good). Smaller and larger sizes got only about a 70 % good or very good rating. There is however a slight tendency to prefer larger rather than smaller letters. The font type Arial was judged good or very good by about 80 % of users, RBB interstate light was only tested for one week yet and judged by 70 % of users to be good or very good.

Concerning the background no tendencies can really be described as they were changed each week so far. All this needs to be tested repeatedly and cross checked. Best, so far is the black box (55% of users rate it good or very good), second best an average is a transparent box (52% of users rate it good or very good).

5.1.3 Red Bee (UK)

The BBC and Red Bee took part in a national campaign to promote awareness of AD on digital television in the UK which ran in the early part of 2008 and was evaluated for OFCOM by GfK².

Awareness of AD increased from 37% to 60% overall and from 43% to 72% among the visually impaired. Among those with a severe visual impairment, awareness increased from 61% to 82%, and went from 40% to 60% among those with a moderate impairment and from 26% to 66% among those with a mild impairment.

1 in 3 were aware of an advertisement featuring AD, 9 out of 10 were aware of a feature on TV and 38% took some sort of action as a result (e.g. calling the helpline or telling a friend about it).

The RNIB sold 300 of their set-top boxes within a month of the campaign and apparently some manufacturers have shown renewed interest in the market.

Audience satisfaction with the service was high, 92% of users feel it improves their enjoyment of TV.

² http://www.ofcom.org.uk/research/tv/reports/access_services_audio/

The negative comments mainly covered such issues as: there's not enough AD, the right sort of programmes are not AD'd and there's not enough variety.

OFCOM concluded that there was potential for an increase in usage of AD and more information about the service will increase access to it.

The results of this awareness campaign have implications not only in the UK but across the European Union. It would be useful to put together information on the extent to which such campaigns have been run elsewhere and what is known about their impact on popular awareness of AD.

5.2 Universities and the mature access services evaluated

In Subtitling for the **Deaf and Hard of Hearing** though still at a very early stage, some initial complications have been identified. The first problem has to do with the heterogeneity of the viewers and of the results obtained in the different countries.

Although participants in the experiment have been divided into three groups (hearing, deaf and hard-of-hearing) and all the groups have been narrowed down with regard to age, it is still difficult to have comparable groups across countries and, sometimes, to find common patterns in the results.

Another problem arises from the sheer number of contributors involved in the DTV4All project, from broadcasters to service providers and front-line researchers, which prevents the project from reaching consensus at a fast pace. However, these are minor drawbacks which may actually be regarded as advantages. First of all, although comparability and coherence are considered as important in the tests, the differences found across countries may reveal important national subtitling (or even cultural) practices or habits, all of which are to be taken into account when analysing the results and indeed when considering how to provide SDH for digital TV.

As for the number of contributors involved in the DTV4All project, although it may account for slow but steady progress in some aspects of the project's work, the involvement of broadcasters and service providers in the project guarantees that the work carried out by the academic researchers in the consortium of universities will have practical implications for the subtitles shown on digital TV across Europe.

As far as the tests on formal subtitling parameters are concerned, the analysis of the first results, however tentative, shows interesting patterns. With regard to the subtitling font, whether opting for Arial (Spain) or Arial/Verdana (UK) as first choice, all viewers seem to reject the use of Tiresias. Although further tests with larger samples of viewers should be carried out, this calls into question the suitability of this font, which was created precisely for subtitling purposes in the UK. With regard to font size, Spanish participants chose 32, whereas in the UK 28 was preferred. Much bigger or smaller sizes were ruled out. Also worth noting is that many of the British participants, especially the hard-of-hearing ones (aged 65+), found it extremely difficult to distinguish the different sizes (28, 32 and 36), which questions the validity of this particular test for these participants. With regard to position, convention proved to be a decisive factor. Thus, Spanish viewers

preferred the bottom and mixed positions shown on Catalan and Spanish channels respectively while British viewers chose only the bottom position.

Habits and convention may also explain why colours were favoured as first choice for speaker identification in Spain, whereas the results in the UK present a different and more complex landscape. Yet, some interesting implications are emerging here. For instance, the deaf, not only in the UK but also in the Spanish tests, seem to like displacement and colours, but not tags. In contrast, the hard-of-hearing clearly prefer tags and reject displacement. This may be due to the age factor. The deaf viewers taking part in the tests were aged between 22 and 45 years old and were used to reading colours and chasing information on a screen, whether on TV or on a computer screen. The hard-of-hearing are generally older and less used to doing this. For them, tags pose fewer problems for speaker identification. Needless to say, both groups are potentially within the subtitling audience, which illustrates the difficulty of finding a standardised form of SDH that can please all of its users.

Finally, a further comment is in order with regard to the importance of conventions in the tests. In the Spanish tests on position and identification, the participants' views regarding specific issues do not seem to correspond to their views on overall preference. In other words, whereas they identify a given method as the best position or the best way to identify speakers, they then choose another one, usually the one applied in their country, as their preferred position or identification method. It would appear that they objectively admit the validity of a given method but end up choosing what is conventional in their country. This mismatch has also been found in recent experiments with eye-tracking technology (Tuominen, 2008), where the viewers' preferences do not correspond to what has been identified as objectively better by the eye-tracker. In these cases, and given the above-mentioned importance given to empirical research, should broadcasters and service providers go with the eye-tracking evidence and against the viewers' stated preferences? Would this not be a form of enlightened despotism, where all is done for the viewers, but nothing by the viewers? Would it be better instead to disregard the empirical evidence and go with the viewers' stated preferences, given that they are the consumers? Although it is still early to answer any of these questions, it seems advisable to look for a happy medium between considering that the consumer is always right and regarding the eye-tracker as the be-all and end-all of the research. The approach adopted in the DTV4All project is to regard viewers' preferences and empirical research as complementary. The initial steps towards the optimisation of SDH in digital TV on this basis have been outlined.

In **audio description** further analysis, such as that carried out by corpus linguistics, and resulting data gathered through the experiment is expected to be studied through different methodological approaches such as eye-tracking. The outcome will show some possible universals, or laws, or regularities in the reception of visual input. This would have implications for the formulation of harmonised guidelines for AD scriptwriting across Europe, and the possibility of the translation of a source audio-description text. It may, on the other hand, show that while from a top down perspective it was thought that there existed a uniform society and culture, we Europeans are still different and perhaps need more time before we go in search of common cultural behaviour.

6 Preliminary Conclusions

At the time of writing this report tests are underway, but the project team can already provide some comments which have been observed, and some recommendations.

1 – For accessibility purposes a difference must be drawn between subtitling countries and dubbing countries (France, Germany, Italy and Spain). Special attention should also be paid to countries such as Poland where they neither subtitle nor dub but use lecturing/voice-over as their preferred mode “According to recent research (a poll by Inst. SMG KRC Poland, 2002) 50.2% of Poles prefer voice-over and 43.4% opt for dubbing; while subtitling is preferred by only 8.1%. A staggering 72.1% of Poles, when asked which type of AVT was the worst, chose subtitling. The latter is a standard in Polish cinemas. Intra-lingual subtitles seem to be gaining ground on Polish television in documentaries with authentic utterances played back from a low-quality recording, e.g. telephone conversations, and dubbed cartoons as well as certain commercials are gaining popularity, but documentaries and foreign films for television are voiced-over. This technique may be beneficial for foreign language learners, although subtitling is undoubtedly a better choice in this respect (Brett, unpubl.) and less costly than dubbing in that only one reader is hired, but its imperfections are many. Notwithstanding, it remains the main mode of transferring foreign programmes onto the Polish television market because of target audience expectations³.”

2- The BBC achieved by 1 April 2008 almost 100% subtitling. This was achieved thanks to the technique of “subtitling by respeaking”. Though the quality and accuracy of this type of subtitling is still to be improved subtitling by respeaking has proven to be a potent tool towards accessibility. In order to take on board this technique Speech Recognition (SR) Software must be available in the language used. Some commercial firms offer quality SR programmes, such as Dragon (Nuance) or ViaVoice (IBM), but for minority European languages there are no SR software available and its creation is not guaranteed. Hence one recommendation will be to create a pool of European languages SR software in order to promote 100% subtitling.

³ Bogucki, Lukasz (2004): “The Constraint of Relevance in Subtitling”. *The Journal of Specialised Translation*: 69-85.

<http://www.jostrans.org/issue01/articles/boguckien.htm#about> [Retrieved 6/05/2009]

7 Implications for Work in Next Period

The links between the pilot of mature services, dissemination and standardisation work will have to be re-examined and strengthened in the coming three-month period of the project if the project is to have a long-term impact on the take-up and use of access services for digital television in Europe.

8 Appendix 1 Design document for DR live subtitles

Summary

This proposal covers the experimental design of one of a number of user studies being conducted as part of the EC funded project DTV4All to improve the accessibility of television through assistive technologies on digital television.

The focus of this study is on quality issues to do with live subtitles for the deaf and hard of hearing.

The aim is to identify the frequency with which viewers with known hearing impairments react adversely to semantic errors, problems caused by presentation differences between live and pre-recorded subtitles and the problems caused by delay between the programming and SDH subtitles.

It also aims to identify the perceived importance of these three categories of problems, ranging from "cosmetic blemish" to "show stopper".

The results will be feed into reports to help European broadcasters with strategic planning for pre-recorded and live subtitling. Further studies conducted by DTV4All between now and 2010 will address related issues to do with the accessibility of Audio Description and signing.

Live subtitles for the deaf and hard of hearing (S) - the research problem

Television programmes are increasingly broadcast with subtitles not just for inter-lingual communication, translating one language to another, but also for intra-lingual communication to improve the accessibility of the soundtrack for those with hearing impairments (SDH).

In a number of European countries the trend is towards 100% SDH subtitling of television programmes.

Providing live subtitling is a challenging matter. It either requires a set-up like that at TV3 Catalonia where a team of up to six subtitlers take turns to manually key-in small chunks of the required subtitles, or the subtitler listens to the programme and dictates the subtitles into a speech-to-text system ("re-speaking"). Regardless of how the subtitles are produced, there is a delay in relation to the programme of six to ten seconds. In The Netherlands, public service broadcasters introduce a delay in live programming on cable of 10 seconds, allowing for the subtitles to be shown in synch with the content.

There are several problems with live subtitling: the quality of the subtitles themselves, the way they are displayed on screen and the delay in showing the subtitles in relation to the video and audio to which they refer.

Three international service providers that use dictation solutions, namely, IMS, ITFC, and Red Bee, claim up to 96-97% content accuracy for widely spoken languages such as English. There are alternative solutions available for widely spoken languages such as English, but options for many less-widely-spoken European languages are either limited or non-existent.

Broadcasters offering live subtitles and using respeaking systems report that there has been criticism of the quality of live subtitles, primarily semantic errors (misspellings, incongruous and omitted words).

Live subtitles also differ from their pre-produced counterparts in that text is presented as soon as it is available and the presentation conventions are different. The delivery rate may vary making demands on the viewer's reading speed.

The focus of the quality debate has been on semantic errors. Less emphasis has been given to presentation and the significance of the delay. There seem to be no known formal evaluations of live subtitling in the research literature, only informal feedback from call centres and broadcasters.

News has been chosen for the study. The reasons for this are that:

- The main news at 18:30 already offers live subtitles
- News is one of the critical genres in terms of its complexity (involving a mix of pre-produced and live subtitles)

Other live genre (sport, factual, general election coverage, events) may well represent slightly different challenges for live subtitling, as they contain a higher proportion of spontaneous speech at a higher delivery rate. Even so, results for a news programme will be indicative of how users perceive live subtitles, and there is always the option of conducting additional studies on other genre if this is deemed necessary.

This is the rationale for suggesting an exploratory study with Danish viewers with a range of hearing impairments to assess the relative importance of three issues:

- Semantic errors
- Presentation differences between live and pre-recorded subtitles and
- The delay between the programming and SDH subtitles.

Scoping the subtitling study

The case for the provision of access services is discussed in depth in an article from the forthcoming HCII 2009 conference in San Diego.

In countries such as the UK, the Benelux and the Nordic countries where inter-lingual subtitling and SDH are widespread, there is broad political and popular support for this kind of access service.

The take-up for SDH can be as high as 12-15% of the adult population. The elderly are the main target group, and their numbers are forecast to increase as mean life

expectancy increases and the elderly account for an increasing share of the total population. Other groups include immigrants, and work done in Catalonia suggests that subtitles can help promote social integration.

Broadcasters in “subtitling” countries are having to put in place live subtitling provisions and need to ascertain the effectiveness and efficiency of SDH subtitling.

- DR is currently reviewing three years’ experience with a ”re-speaking” system originally developed by a Philips subsidiary. In order to meet the requirements of its public service contract, it may be necessary to opt for a new solution. Work has started in the second quarter of 2009.
- In a Swedish government white paper published in 2008, a recommendation was made to earmark funding for the development of a new SVT live subtitling system.

Given the lack of research on existing live subtitling provisions, DTV4All will be conducting an exploratory study involving a minimum of 30 Danish adults with a range of hearing impairments. Similar user studies on access services are being done in Germany, Catalonia, and Italy.

In the Danish study, each viewer will be invited to watch DR1 TV Avisen [the main television news] on an individual basis from 6:30 to 6:55 pm. She/he will be asked to press a buzzer every time there is a difficulty following the programme, regardless of the reason.

At the end of the news, the observer goes through a recording of the programme with the viewer and discusses each point at which the viewer pressed the buzzer and together they note the cause of the problem. These annotations are stored with the recording for later analysis.

The study is primarily concerned with establishing the frequency and perceived importance of the three main types of problem: semantic errors, problems caused by presentation differences between live and pre-recorded subtitles and the problems caused by delay between the programming and SDH subtitles.

Detailed linguistic analysis to produce a typology of user problems is out of scope, but the research material can be analysed by academic researchers working within DTV4All. If the study shows that presentation and/or delays have the biggest negative impact on following SDH subtitles, a follow-up study will be planned and implemented in which different treatments of the same programme will be compared.

DTV4All will be doing a separate proof-of-concept test on buffering the video and audio of programmes with live subtitling as a possible feature of Personal Video Recorders. This approach is also being considered in NORDIG.

Research objectives

1. To identify the frequency with which viewers with known hearing impairments react adversely to semantic errors, problems caused by presentation differences between live and pre-recorded subtitles and the problems caused by delay between the programming and SDH subtitles.
2. To identify the perceived importance of these three categories of problems, ranging from "cosmetic blemish" to "show stopper".

Research design

The study is explorative only. We will need to ask participants about their prior experience with SDH subtitles, because familiarity with the differences in presentation conventions may improve SDH usefulness.¹

The design of the study is based on previous DR research from the mid nineties during which systematic work was done on subtitling using a holistic method.

Much academic research on subtitles "primes" viewers into providing specific feedback on, say, spelling mistakes, the use of colour, the use of two or three lines of subtitles, reading speed etc. The aim here is not to introduce bias of this kind, but to have some global measure of the accessibility of a given service.

The same set-up can be used to test Audio Description and other proposed services and, indeed, any kind of television programming requiring user interaction.

It should be noted that TV-Avisen contains a mix of inter- and intra-lingual subtitles, the former being pre-recorded, the later produced by "re-speaking". If we can find a solution for news, solutions for other programming genres should be less demanding.

Experimental set-up

1. Introduction to the study. A folder sent to participants in advance.
2. Briefing of individual viewers at DR (early evening)
3. Training in use of buzzer, an open source tool supported by DTV4All and the EBU for recording any output from a television set, marking events on the file and allowing for annotations.
4. Watch TV-Avisen
 - a. The observer makes notes on viewer reactions during the 25 minute period
 - b. The viewer presses buzzer
5. The observer and the viewer review the recording of the programme with time markers indicating points at which the viewer pressed the buzzer.
 - a. The observer takes the viewer through each buzzer point, discusses the viewers reactions and compares with own notes from observation
 - b. Observer and viewer agree on annotation that the observer records in the file
 - c. Observer saves the recording with all annotations for future reference
6. Observer administers a short questionnaire on Survey Monkey (same ID as file, no name) addressing:
 - a. Background parameters (demographic, use of access services)

- b. Self-assessment of capability using same scales as the largest European self-reporting study on capabilities and impairments.
- c. Rounding off – arranging for transport home for the viewer.

Research subjects (recruitment and selection of subjects; privacy)

The aim is to recruit not only persons unable to hear but also those with a range of hearing impairments. We need to consider a range of ages, too.

For these reasons, we hope to find subjects with a known range of hearing impairments in collaboration with Danish hearing aid manufacturers or resellers. Indeed, other parts of DTV4All will require collaboration with these organisations in order to assure, say, wireless access to Audio Description.

It goes without saying that special interest groups such as Dansk Døves Landsforbund [the Danish Association for the Deaf] and Ældresagen [the DaneAge Association⁴] will be kept informed of these studies. Recruitment to the study does not require their help.

We plan to ask each participant to assess their capabilities and to get some background information on them and their experience to date with access services. We will not be recording names, but will be pairing the video recordings with the ID tags for the questionnaires to respect confidentiality.

Data expected

30 data sets covering:

	Semantic problems	Presentation problems	Problems caused by subtitling delay
Frequency			
Importance			

Basic demographic parameters for each subject
 Prior experience with live subtitles and other access services

Study team

Peter Looms (design, supervision, analysis, report writing)
 Anni Rander and Marguerite Johnsen (conducting studies with 30 viewers under supervision of Peter Looms)
 Anni has completed her dissertation and graduates in interaction design from IT University this summer. She has experience in observation and ethnographic studies.
 Christoffer Godt-Hansen (peer review)
 Lars Thunøe (peer review)
 Sofie Scheutz (peer review)
 Peter Mølsted (technology, liaison with Nordija)

⁴ <http://www.aeldresagen.dk/Medlemmer/detgoervifordig/omos/english/Sider/Default.aspx>

Data analysis

Peer reviewed by DR's Audience Research and UAB (Pilar Orero).

Reporting

Internal reports to DTV4All. Public versions will be available in 2010.

9 Appendix 2 Questionnaires

- 1- Sample of questionnaires in English and Spanish for the SDH tests taken on board by universities.

Questionnaire

1) PERSONAL DETAILS

a. Gender: Male Female

b. Age: 17 – 24 / 24 – 35 / 40 – 59 / 60 → _____

c. Education (tick all the studies that you completed):

- Primary School
- Secondary School
- Higher (Tertiary education) School

Type of school

- Deaf school
- Mainstream school

Specify the number of years of education that you had:

d. Occupation:

- I study
- I work (please specify) _____
- I don't work
- I'm retired

e. Are you ...?

- Deaf
- Hard of Hearing
- Hearing

f. When did you become deaf?

- From birth
- 2 years old
- 2-4 years old
- 5-19 years old
- 20-29 years old
- +50 years old

When did you become Hard of Hearing?

- From birth
- 2 years old
- 2-4 years old
- 5-19 years old
- 20-29 years old
- +50 years old

Do you use a hearing aid/ implant?

- Yes
- No

g. Language used to communicate:

- Only BSL
- Only English
- Both

h. Sight:

- I use glasses/contact lenses
- No aid needed

i. Do you have difficulties reading English?

- Yes
- No

j. Do you have difficulties watching the TV screen or reading subtitles?

- Yes No Sometimes

2) GENERAL INFORMATION AND PREFERENCES

a. How many Deaf people live with you?

- 0 1 2 3 4+

b. How many Hard of Hearing people live with you?

- 0 1 2 3 4+

c. Which of the following do you have at home?

- TV PC DVD / VHS player Internet
 Mobile phone Other _____

d. How many hours a day do you spend reading newspapers, books...?

- 0 h 1-2 h 2-3 h 3-4 h 4-5 h 5-6 h+

e. How many hours a day do you watch TV?

- 0 less than 1h 1-2h 2-3h 3-4h 4+ h

f. How many hours a day do you spend watching subtitled programmes?

- 0 h 1-2 h 2-3 h 3-4 h 4-5 h 5-6 + h

g. If you watch TV, who do you usually watch it with? (you can tick more than one)

- By yourself Deaf and HoH friends/family Hearing friends/family

h. What programmes do you usually watch on TV?

- News Films and series Talk shows/ quizzes... Sports
Documentaries Soaps Others

Which of these types of programmes, if any, would you normally expect to be able to follow without subtitles?

i. Do you choose the programmes you watch based on whether they are subtitled or not?

- Yes No

j. What do you use subtitles for? (You may tick more than one option)

- They help me understand They are my only way to have access to the dialogue
 I use them for language learning

k. How do you know which programmes/films include subtitles?

- Teletext TV announcements TV guides Friends Other _____

l. What do you do when a programme doesn't offer subtitles?

- Switch the TV off or switch to another channel (for a subtitled programme)

Continue to watch....

- Guessing at speech through context
 Lip-reading / word processing
 Turning up the volume
 Relying on fellow-watchers to translate

m. What do you think is the best way to make audiovisual material accessible?

- Assistive device (wireless / infrared headphones) Subtitles Other

3) SUBTITLING

a. What do you think of subtitling in general (TV / DVD / cinema...)?

- Satisfactory Better than not having Unsatisfactory

b. What change would you like to see introduced in subtitles?

c. Are you aware of any conventions on how subtitles should be done?

- Yes No

d. If you know any of them, mention them:

e. What difficulties do you think are involved in producing subtitles?

4) SUBTITLES ON TV

a. Is it easy to find information on Teletext about which programmes are subtitled?

- Yes No Don't know

b. Which channels offer the best pre-recorded subtitles?

(Rate them from 1 to 6: 1- best subtitles; 6 - worst subtitles)

- BBC ITV Channel 4
 Channel 5 Sky Other

c. Why do you think this? (You may tick more than one)

- Amount of subtitles provided Language_____ _Synchrony
 Speed Other

d. Which channels offer the best live subtitles?

(Rate them from 1 to 6: 1- best subtitles; 6 - worst subtitles)

- BBC ITV Channel 4
 Channel 5 Sky Other

e. What channel offers the best subtitled news?

- BBC ITV Channel 4
 Channel 5 Sky Other

f. Why do you think this? (You may tick more than one)

- Amount of subtitles provided Language_____ _Synchrony Speed
 Other

5) SUBTITLES ON DVD

a. Where can you find information about which DVDs include SDH?

b. Do you choose the films you watch based on whether they offer subtitles for the Deaf and Hard of hearing people or not?

- Yes No Don't know

c. Finding subtitle options in DVD menus is:

- Easy Difficult

d. What type of SDH do you find easier to read / understand?

- TV DVD

e. Why? _____

6) SUBTITLING STYLES

a. Do you find the font used in teletext easy to read?

- Yes No Don't know

b. Do you find the font used in DVD/SDH easy to read?

- Yes No Don't know

c. When characters need to be identified, what system do you prefer?

- Only colours Only positioning subtitles next to / under the characters
 Only name tags Combining colours and positioning

d. The number of colours used is:

- Sufficient We could do with a wider range
 Too many and therefore difficult to read

e. Where do you prefer subtitles to be shown?

(assuming that whatever your general preference, the position would be changed temporarily to accommodate captions or strap lines).

- Bottom of the screen only Both top and bottom of the screen
 Top of the screen only Next to the character who speaks each time

f. For live events, how do you prefer subtitles to be shown?:

- Word by word Blocks

h. How do you prefer descriptions of sounds to be reflected on the subtitles?

- Explaining where the sound comes from Using words reproducing the sound
 Describing what the sound is like Pictograms

i. Where do you prefer sound-related information to be shown?

- Top-right side of the screen Bottom of the screen next to the subtitles
 Next to the source of the sound

j. Regarding information about the mood of the characters, how do you prefer that to be shown?

- With emoticons Explanation between brackets Nothing

k. When there is instrumental and background music in a film/TV series, what do you prefer?:

- To have the title of the song on screen

- To have information on what type of music it is
 - An icon indicating "music"
 - Nothing
-

i. When there are meaningful songs in a film/TV series, what do you prefer?:

- To have the title of the song on screen
 - To have the words of the song subtitled
 - To have information on what type of music it is
 - An icon indicating "music"
 - Nothing
-

m. Which of the options below do you prefer for pre-recorded subtitles?:

- Literal subtitles that contain absolutely all the information
- Not so literal but easier to read

n. Which of the options below do you prefer for live subtitles?:

- Literal subtitles that contain absolutely all the information
- Not so literal but easier to read

o. If you chose "literal subtitles" in the above questions, can you explain why you prefer them?

p. If it is not possible to represent everything in the subtitles, which is the most important thing to include?

(Please rate from (1) most important information to (4) least important.)

- Dialogue
- Names
- Sounds effects (thunder)
- Mood or way of speaking (eg: "shouting", "whispering"...) "whispering"...
- Expressions like "ok", "well..."...

q. What do you think about the usual speed of pre-recorded subtitles on TV?

- They are too fast
- They are OK
- They are too slow

r. What do you think about the usual speed of live subtitles on TV?

- They are too fast
- They are OK
- They are too slow
- I haven't seen subtitles in a live programme

s. What do you think about the usual speed of subtitles in DVDs?

- They are too fast
- They are OK
- They are too slow

- TV TV digital Ordenador Portátil Vídeo
 DVD Satélite Parabólica Internet Teléfono
 PDA Móvil Ipod Otros _____

f. ¿Ves la televisión sólo?

- Sí No Casi siempre Casi nunca

g. ¿Con quién sueles ver la TV?

- Sólo Amigos / Familia sordos Amigos / Familia oyente

h. ¿Cuántas horas de televisión ves al día?

- 0 1 2 3 4 5 +6

i. De las horas que ves la tele, ¿cuántas están subtitradas?

- 0 1 2 3 4 5 +6

j. ¿A qué horas ves la tele?

- 07:00 – 12:00 12:00 – 17:00 17:00 – 21:00 21:00 – 01:00

k. ¿Eliges el programa dependiendo de si tiene subtítulos o no?

- Sí No

l. ¿Qué tipo de programas sueles ver?

- Noticias Deportes Películas y series Documentales
 Concursos

m. ¿Por qué usas los subtítulos?

- Ayuda a comprender Única forma de comprender
 Aprender lenguas Otros _____

n. ¿Cómo sabes qué programas / películas / DVDs tienen subtítulos?

- Teletexto TV Guías Amigos Otros _____

o. ¿Qué haces cuando el programa / película no tiene subtítulos para sordos?

- Apagar el TV Cambiar (Buscar programa subtitrado) Leer los labios
 Interpretación de un acompañante Subir el volumen (TV / Audífono) Imagino por el contexto Usar subtitrado normal (si tiene)

p. ¿El subtitrado es lo mejor para ver la tele?

- Sí No NS / NC

q. ¿Si piensas que NO, ¿qué consideras más útil?

- Intérprete Subtitulado adaptado Intérpretes virtuales

3) Subtitulado

a. ¿Qué piensas sobre el subtitrado en general (TV / DVDs / cine...)?

- Está bien Mejor que nada No está bien
 Otras _____

b. ¿Qué cambiarías / quitarías en el subtitrado que se hace?

c. ¿Sabes si existe una normativa sobre subtitrado?

- Sí (¿Cuál? _____) No

d. ¿Sería más fácil leer subtítulos si fuesen igual en todas partes (cine / TV / DVD...)?:

- Sí No NS / NC

e. ¿Crees que es difícil hacer subtítulos? ¿Por qué?

TV

f. ¿Es fácil saber qué programas están subtitrados con el teletexto?:

- Sí No NS / NC
- g. ¿Qué cadenas tienen mejores subtítulos? (Numerar del 1 al 6: 1-mejor servicio; 6-peor servicio)**
- TVE (La Primera) TVE (La 2) Antena 3 Cuatro
- Telecinco La Sexta Otras (Autonómicas) _____
- h. ¿Por qué?**
- Más programas subtitolados Sincronía Velocidad
- Lenguaje más fácil Otros _____
- i. ¿Los informativos de qué cadena te gustan más?**
- TVE (La Primera) TVE (La 2) Antena 3 Cuatro
- Telecinco La Sexta Otras (Autonómicas)
- j. ¿Hay diferencias entre los subtítulos de las cadenas? Si piensas que SÍ, ¿cuáles son las diferencias?**
- colores colocación de los subtítulos explicación de sonidos
- tipo de letra tamaño de la letra velocidad de subtítulos
- Otros _____

DVD

- a. ¿Es fácil encontrar información sobre los DVDs que tienen subtitolado para sordos?:**
- Sí No NS / NC
- b. Si piensas que SÍ, ¿cómo encuentras esa información? _____**
- c. ¿Dónde sueles comprar / alquilar / conseguir los DVDs?**
- Asociación Centro comercial Videoclub Amigos
- Venta por Internet Videotecas públicas No veo DVDs
- d. ¿Eliges las películas sólo si están subtitoladas para sordos?**
- Sí No NS / NC
- e. Encontrar las opciones de subtitolado en los menús es:**
- Fácil Difícil
- f. ¿Cuál es el principal fallo del subtitolado para DVDs?**
- Colores Descripción de sonidos Rapidez
- Que toda la información aparezca abajo Otros _____

4) Estilos

- a. ¿La letra del subtitolado es fácil de leer?**
- Sí No NS / NC
- b. ¿Qué letra es más fácil de leer?**
- Teletexto normal Teletexto digital DVD
- c. ¿Qué prefieres para distinguir a los personajes?**
- Colores El subtítulo al lado del personaje
- El nombre Colores y subtítulo al lado del personaje
- d. Los colores que se utilizan son:**
- Suficientes Se podrían utilizar más Difíciles de ver
- e. Prefieres que los subtítulos salgan:**
- Sólo abajo Abajo y arriba (como en TV) Sólo arriba
- Al lado del personaje que habla
- f. Cuando se subtítulo algún acto en directo, ¿cómo prefieres que aparezca el texto?:**
- Palabra a palabra Todo a la vez (como el teletexto normal)
- g. El subtitolado en directo sale un poco tarde, ¿hasta cuántos segundos podrías esperar ese retraso? _____**

- h. Cuando hay sonidos, ¿cómo prefieres que se subtitulen?**
- Explicando qué pasa Imitando el sonido (onomatopeyas)
- Describiendo cómo es el sonido Dibujos
- i. ¿Dónde prefieres la información sobre los sonidos?**
- Arriba a la derecha Abajo al lado del texto Al lado de lo que produce el sonido
- j. Para explicar cómo habla una persona (si está gritando, triste...), ¿cómo prefieres que se haga?**
- Emoticonos (:-D) Explicación entre paréntesis Nada
- k. Cuando hay música en la película / serie..., prefieres que:**
- Escriban el título de la canción Subtitulen la letra de la canción
- Expliquen qué tipo de música es Incluyan un icono de música
- Otros _____
- l. Prefieres que se subtitule:**
- Absolutamente todo (literal) Menos texto pero más fácil de leer
- m. En caso de seleccionar la opción “Absolutamente todo”, ¿por qué?**
-
- n. Si por motivos de espacio no se pudiese, ¿qué prefieres que se subtitule? (Indicar de más importante (1) a menos importante (4).**
- Diálogos Nombres Sonidos Estado de ánimo
- (“grita”, “susurra”...) Expresiones tipo “vale”, “bien”...
- o. ¿Qué piensas de la velocidad de los subtítulos para TV?**
- Son muy rápidos Están bien Son muy lentos
- p. ¿Qué piensas de la velocidad de los subtítulos para DVD?**
- Son muy rápidos Están bien Son muy lentos
- q. ¿Qué piensas de la velocidad de los subtítulos en programas en directo?**
- Son muy rápidos Están bien Son muy lentos

10 Appendix 3 Production of access services

Access services at **DR** grew out of inter-lingual subtitling that dates back to the beginning of television broadcasting in Denmark in the fifties. Subtitling for the Deaf and Hard-of-Hearing were offered shortly after the introduction of teletext services in the early eighties. Signing has been offered on analogue TV for nearly three decades. With the introduction of digital terrestrial television the opportunity arose to offer signing on three public service television channels (DR1, DR2 and TV2) by creating a virtual signing channel in DTT mux 1. Live subtitling for the main TV news was introduced in 2005. Audio description (broadcaster mix) was introduced on an experimental basis in the autumn of 2008 and is now in normal operation, currently with mainstream Danish language drama series.

Bayerisches Rundfunk created their Access Services Department (Begleitdienste Fernsehen) and started producing AD in April 1997. The first audio described film was aired in July 1997: it was a TV-movie called *Alles auf Anfang* (All back to the beginning). Some movies and a TV series called *Russige Zeiten* (Sooty Times) followed. Previously, a few AD presentations in cinemas had been offered, in 1989, 1990 and 1993, and another German broadcaster (ZDF) had also aired one or two audio described films per year from 1993.

BR has audio described around 190 programmes, most of them movies or TV movies, totalling around 350 hours of television. The programmes are aired on the two channels offered by BR (Bayerisches Fernsehen and BR-alpha) and on other channels of the public ARD network, including Das Erste —Germany’s premier television channel—, Arte, 3Sat, Kika and other regional public broadcasters in Germany. Bayerisches Fernsehen (Bavarian TV) offers AD of 4% of its prime time television programmes.

Concerning the exchange of audio description, it must be stressed that other broadcasters in Germany get BR’s descriptions free of charge, as BR gets their descriptions for free in return. Austrian and Swiss TV (German-speaking region) also use and contribute to the pool of all German language audio description, hence producing AD for films that can also be broadcast in Germany. Sometimes BR sells AD to be used on DVD or is asked by DVD companies to deliver AD for a film. In these instances, if the AD can be used in BR’s programmes, the cost is shared.

Regarding future prospects, BR aims to increase the number of audio described hours, and also aims to co-produce with DVD companies and increase the presence of AD in film festivals.

TVC

In November 1989 **Televisió de Catalunya (TVC)** broadcast the first audio described film in Catalan (*The Ten Commandments*), becoming the first Western country to offer audio description. In the 90s more audio description was aired, such as some episodes of the Catalan series *Estació d’enllaç*, although not regularly. In the 21st century, thanks to cooperation with the Spanish blind association ONCE, AD was included in the

Catalan sitcoms *Plats Bruts* —later released in DVD format— and *L'un per l'altre*, and the Catalan series *Majoria absoluta*.

In 2006 a new project was launched and the audio description of *La Gran Pel·lícula* (*The Big Film*, a selection of box-office successes broadcast on Friday night) started to be aired weekly. The first publicised film was *Something to Talk About* (February 16th 2006), followed by *The Majestic* and *The Pledge*, among many others. An unannounced AD of the film *Mystic River* had been broadcast for testing purposes, so that the Catalan association for the blind and visually impaired, ACCDV, and the Spanish association ONCE could give feed-back to the television network about the service.

The AD of a weekly film has been expanded in 2008 with the AD of three children's programmes which are shown over the week-end: *Hotel Zombie*, *Being Eve* and *King Arthur's Disasters*, and the Catalan mini-series *Serrallonga*, which will soon be available on DVD. Whilst in 2007 81 hours of television were audio described, in 2008 this had increased to 111 hours, and TVC's aim is to offer 160 hours of audio described television in 2009, which represents approximately 4% of the new emissions broadcast on its main channel, TV3.

11 Appendix 4 Workplan for evaluation

UAB

Activity S-5: Compilation of long questionnaires

Data obtained from long questionnaires on the viewers' habits and general opinions regarding subtitling will be compiled and processed (Spain, France, UK, Belgium, Italy, Greece and Denmark).

Internal deliverable S5: report on S-5 (month 16)

Months: 5-6-7-8-9-10-11-12-13-14-15-16

Participants: UK; Italy; Spain; Belgium; Denmark; France; Greece

Test: Compilation of long questionnaires

Team Leader: Stavroula Sokoli

Dissemination: Results to be presented at the CESyA conference in June 2009

At this stage, a view of the (subjective) preference of viewers from different European countries regarding SDH (resulting both from the tests and the long questionnaires) will be emerging. The next stage is to test this empirically using eye-tracking technology.

Internal milestone 1: Report on the subjective preference of viewers from different European countries regarding SDH (resulting from both the tests and the long questionnaires) - Participants: Spain, France, UK, Belgium, Italy, Greece and Denmark (month 16)

Activity AD-2: Pear Stories 2

Activity AD2.1: A preliminary comparison of transcripts according to Tannen's features.

Delivery of the whole corpus gathered with the Pear Tree projects with back translations and analysis of the data.

Dissemination: Project setup to be presented at the International conference La traduction audiovisuelle: approches pluridisciplinaires (Montpellier, June 2008)

Activity AD2.2: An in-depth comparison of how transcripts select and present information about the Pear Story film.

Internal deliverable AD2: report (month 10)

Months: 1-2-3-4-5-6-7-8-9-10

Institutions: All

Test: analysing the reception of the same film by different cultures

Team leader: Andrew Salway

Dissemination: Results to be presented at the CESyA conference in June 2009

WP 2.4 Implementation of Pilot Evaluation

Data gathered in this task will be used to identify obstacles and constraints to the introduction of sustainable access operations of a technical, organisational, legal, cultural and context-specific nature. As the evaluation seeks to generate inputs to help optimise mature services, data on the effectiveness and efficiency of existing services based amongst other things on the outcomes of evaluation by user groups will also be collected and collated. The deliverable here will be a draft report for each of the four territories of the pilot and will include generic results of relevance throughout Europe.

subtitling

The results of tests carried out with deaf, hard-of-hearing and hearing participants across three countries, Belgium, Spain, and Italy will be compiled. The tests will obtain viewers' opinions on five formal subtitling parameters: font, size, position, justification, character identification, and the use of boxes, borders, shadows, icons and emoticons in subtitles. To validate the data obtained, the results of comprehension tests and validation tests using eye-tracking technology will be collated to produce empirical evidence of the formal subtitling parameters preferred by the viewers.

The results of tests about the viewers' comprehension of subtitles across seven countries: Spain, France, Belgium, Italy, Greece, and Denmark, will be compiled. The tests will include the viewers' opinions on three different subtitle formats (verbatim, standard, and adapted to sign language) and will provide data about their comprehension of the different formats.

Activity S-1: Testing of formal subtitle parameters (1): Layout

Tests will be carried out with Deaf, hard-of-hearing and hearing participants across three countries: Spain, UK and Italy. The tests will gather the viewers' opinions on five formal subtitling parameters: font, size, position, justification, and character identification.

Internal deliverable S1: report on activity S-1 and article (published by Peter Lang) on viewers' opinions about subtitling layout (font, size, position and character identification) (Month 5)

Months: 1-2-3-4-5

Participants: UK; Italy; Spain

Test: Testing of formal subtitle parameters (1): Layout

Team Leader: Carlo Eugeni

Dissemination: Presentations of set-up and scope in La traduction audiovisuelle: Approches pluridisciplinaires (Montpellier, June 19-21).

Activity S-2: Testing of formal subtitle parameters (2): Legibility

S2.1 Tests will be carried out with Deaf, hard-of-hearing and hearing participants across one more country: Belgium. Once again, the tests will cover font, size, position, justification and character identification.

S2.2. Tests will be carried out in Belgium and Spain on the use of boxes, borders and shadows in subtitles. (Spain will join this phase as a control group.)

Participants: Belgium; Spain

Test: Testing of formal subtitle parameters (2): Legibility

Team Leader: Gert Vercauteren

Dissemination: Presentation of results at Languages & The media (Berlin, 29-31 October).

Internal deliverable S2:

a) Report on viewers' opinions about subtitling layout (font, size, position and character identification) -Participants: Belgium

b) Report on viewers' opinions about legibility (use of boxes, borders and shadows in subtitles - Participants: Belgium and Spain

(month 5)

Months: 3-4-5

Activity S-3: Testing of formal subtitling parameters (3): Innovation

Tests will be carried out with Deaf, hard-of-hearing and hearing participants across four countries (Spain, UK, Belgium, Italy) on the use of icons and emoticons in subtitles.

Internal deliverable S3: Report on viewers' opinions about innovation in subtitles (use of icons and emoticons) -Participants: Spain, UK, Belgium and Italy (month 8)

Test: Testing of formal subtitle parameters (3): Innovation

Team Leader: Verónica Arnáiz

Dissemination: Results to be presented at the CESyA conference in June 2009

Activity S-6: Validation tests using eye-tracking technology

Activity S-6.1: Validation tests using eye-tracking technology will be carried out. These tests will produce empirical evidence of the formal subtitling parameters tested in activities S-1, S-2 and S-3 across groups A (UK, Italy) / C (France) / D (Spain). Comprehension tests will need to be passed and supported by eyetracking in order to validate data.

Activity S-6.2: Validation tests using eye-tracking technology will be carried out. These tests will produce empirical evidence of the formal subtitling parameters tested in activities S-1, S-2 and S-3 across B (Belgium) / C (Denmark / Greece). Comprehension tests will need to be passed and supported by eyetracking in order to validate data.

Internal deliverable S6:

a: Report with results on validation tests using eye-tracking technology -Participants: UK, Italy, Spain, France (Month 12)

Months: 8-9-10-11-12

Test: Validating tests using eye-tracking technology

Team Leader: Verónica Arnáiz

b: Report with results on validation tests using eye-tracking technology - Participants: Belgium, Denmark, Greece (Month 16)

Months: 8-9-10-11-12-13-14-15-16

Test: Validating tests using eye-tracking technology

Team Leader: Henrik Gottlieb

Dissemination: Results to be presented at the International Conference Media for All (Antwerp, October 2009).

Internal milestone 2: Report on the objective preference of viewers from different European countries regarding SDH. Validation by means of eye-tracking technology of the data presented in internal milestone 1 -Participants: Spain, France, UK, Belgium, Italy, Greece and Denmark (month 16)

Activity S-7: Testing language comprehension

Tests will be carried out of the viewers' comprehension of subtitles across seven countries: Spain, France, UK, Belgium, Italy, Greece and Denmark. The tests will obtain the viewers' opinions on three different subtitle formats (verbatim, standard and adapted to sign language) and will provide data about their comprehension of the different formats.

Internal deliverable S7: Report on viewers' comprehension of subtitles (whether verbatim, standard or adapted to Sign Language) -Participants: Spain, France, UK, Belgium, Italy, Greece and Denmark (Month 22)

Months: 16-17-18-19-20-21-22

Test: Language comprehension

Team Leader: Jean-Marc Lavour

Audio description / audio subtitling

A professional describer creates an AD from scratch whereas a professional AV translator translates and adapts an AD. A test will be carried out to compare the two approaches in terms of time they take. The films to be used for this test have still to be determined.

Tests using an eye tracker will be carried out to compare how different viewers focus their attention on the screen while verbalising what they have seen. These tests will allow relevant visual inputs to be matched to language outputs.

Activity AD-3: The translation of audio descriptions

Internal deliverable AD3: report (month 11)

Months: 5-6-7-8-9-10-11

Institutions: All

Test: a professional describer creates an AD from scratch whereas a professional AV translator translates and adapts an AD. Comparisons will be made in terms of time. The films to be used will be defined as part of this activity.

Team leader: Aline Remael

Dissemination: Presentation of results at the International Conference Media for All (Antwerp, October 2009)

Activity AD-4: Perception of visual inputs: scientific testing with eye tracking

Internal deliverable AD4: report (month 24)

Months: 1 to 24

Institutions: all partners

Test: Using an eye tracker tests will compare how different informants focus their attention on the screen while verbalising what they have seen. These tests allow the relevant visual inputs to be matched with language output.

Team leader: Pilar Orero

Dissemination: The results will be presented in two international conferences fully dedicated to Media Accessibility, such as Media for All (Antwerp, October 2009) and the Spanish CESyA in 2009.

WP 2.5 Analysis of and recommendations on mature Access Services

A draft report provided by WP 2.4 will be circulated to representatives of a wide range of stakeholders to gauge their support for the recommendations made in the draft report. These recommendations will fall into three main categories:

- Improvements to existing services
- Ways of addressing the obstacles to the general take-up of mature access services in Europe.
- Ways of optimising services and their reception

The draft report will be revised in light of stakeholder feedback to provide the deliverable of this task, namely, a report making recommendations known to have broad stakeholder support.

subtitling

Results from all the other subtitling related tasks of the project will be considered to determine the validity of the following hypotheses:

- there is a need and the potential for Europe-wide SDH guidelines
- there is a need to determine the relevance of users' feedback to the setting of the guidelines
- there is a need to empirically determine the technical parameters of SDH

In the light of the findings, recommendations will be made relating to both the format and the content of SDH as well as the training of subtitlers.

Activity S-8: Final report

Results from all tasks will be considered to address the following hypotheses:

- There is a need and the potential for Europe-wide SDH guidelines
- There is a need to determine the relevance of users' feedback in the set-up of guidelines
- There is a need to empirically determine the technical parameters of SDH

In the light of the findings, the report will make recommendations relating to both the format and the content of SDH, as well as the training of subtitlers.

Internal milestone 3: Final report with recommendations regarding both the format and the content of SDH as well as the training of subtitles – Participants: Spain, France, UK, Belgium, Italy, Greece, and Denmark (month 24)

Audio description / audio subtitling

Results from all tasks will be considered to address the following hypotheses:

- there is a need and the potential for Europe-wide audio description guidelines
- audio description can be translated effectively and efficiently
- it is cost effective for the same person to audio describe and subtitle

In the light of these findings the report will make recommendations relating to the audio description workflow, the content of audio description and the training of audio describers.

While audio subtitling is not yet widely available for broadcasting across European countries, it is used when foreign films are audio described into English. A report will be written describing the many possible combinations of AD found in the following UK available DVDs: *The Passion of Christ*, *Hero*, *Syriana*, *Borat*, *Letters of Iwo Jima*, *Everything is Illuminated* and *Volver*. The report will provide guidelines for good practice when audio subtitling foreign language productions either for broadcasting or DVDs.

Activity AD-5: Final Report

Results from all the AD tasks will be considered to address the following hypotheses:

- There is a need and the potential for Europe-wide audio description guidelines
- Audio description can be translated effectively and efficiently
- It is cost effective for the same person to audio describe and subtitle

In the light of the findings the report will make recommendations relating to the audio description workflow, the content of audio description and the training of audio describers.

Audio Subtitling (AS)

Activity AS-2: While audio subtitling has not been widely available, as yet, for broadcasting across European countries, it is used when foreign films are audio described into English.

Internal deliverable AD5: A report will be written describing the many possible combinations found in the following UK available DVDs: *The Passion of Christ*, *Hero*, *Syriana*, *Borat*, *Letters of Iwo Jima*, *Everything is Illuminated* and *Volver*. In the light of these findings the report will provide a guide to good practice when audio subtitling foreign language productions either for broadcasting or DVDs.

The contributing researchers are:

Pilar Orero, UAB

Aline Remael, University College Antwerp, Belgium

Gert Vercauteren, University College Antwerp, Belgium,

Sabine Braun, University of Surrey, UK

ⁱ Our working assumption is that there is synergy between reading subtitles at the same time as watching lip movement and hearing the audio channels. For this reason, for those with average to good reading skills, we think it likely that delays are likely to have a bigger impact on accessibility than either semantic errors or presentation differences.