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.:	**D2.5**
Deliverable title:	**Final Report on Pilot Services (2 of 2)**
Nature of deliverable:	**Report**
Dissemination level	**Public**

Due date:	**31.05.2010**
Actual delivery date:	**29.06.2010**

Document title and history: **D2.5 –Final Report on mature Pilot Services**			
Version no:	*Date:*	*Changes*	*Contributor*
001	15-05-2010	Structure of document in place	FD
002		Inputs from Pablo Romero	
003		Input from Agnieszka Szarkowska	
004		Input from Germany	
005		Input from Pilar Orero	
006		Input from TVC	
007	29-06-2010	Editorial revisions	TJO

Document authors / organization:	**UAB**
Partners contributed:	**All**
Availability:	**Draft: Open**
Circulation:	**Partners**

Table of Contents

1 Executive Summary

This report represents part 2 of a series of two reports representing the full and final results and conclusions of the work done to date on mature services by the DTV4All project. This report, which can be read as a supplement to the 'D2.5 part 1 Final Report' includes the final results of the eye-tracking work that was conducted by the Universities Sub-titling team. This represents work that was delayed due to issues encountered with the eye-tracking technology that was a key element in the tests that were carried out. The Final Conclusions and Findings are available within the D2.5 part 1 Report.

The studies included here represent a rich and varied collection of research both in terms of their methods, findings and presentation. They also address important lacunae in our understanding of the workings of mature access services. The subtitling studies included in this report refer to internal deliverable 3 of the Subtitling team i.e.

Report on eye-tracking tests, which includes a) opinion (referred to as post-test opinion, as these questions were asked after the test), b) comprehension (what the participants understood?), c) pure eye-tracking data (fixations and so on).

Tests carried out in the Universities, approximate 40,000 subtitles read by hearing, hard of hearing and deaf participants, which constitutes the largest corpus of its kind (eye-tracking plus comprehension plus opinion) and a treasure trove of information for research, further projects, etc. We can now attempt to map out, for the first time with this size of eye-tracking and questionnaire-base data, how hearing, deaf and hard of hearing viewers read and comprehend subtitles.

Section 2 – Roehampton Universty's interim results and fundings on subtitling were presented in the D2.5 Part 1 report. This Part 2 of the report presents the full and final analysis.

Section 3 – The University of Warsaw's reports presents the results of eyetracking tests carried out in Poland as part of the DTV4All project work on subtitling.

Section 4 – The Germany study conducted at the Humboldt University Berlin represents the first study of subtitles using eye-tracking technology in Germany.

Section 5 – TVC Audio Description Evaluation Report is based on analysis of questionnaires.

Section 6 – Audio Description; The Belgian team completed a brief report on spoken subtitles provided in this sectoin.

2 UK Evaluation Report (Roehampton)

2.1 Introduction

This section presents the results of eye-tracking tests carried out in the UK as part of the DTV4All project. The report draws on the following data:

1. Information derived from questionnaires collected before eye-tracking tests (pre-test questionnaire)
2. Eyetracking data
3. Results of comprehension questions obtained during eye-tracking tests
4. Information on preferences concerning particular subtitling variables in each of the parameters tested in the study (post-test preferences)
5. Further research: the notion of watching speed

The section is divided in three parts: The first one (subsection 2.2) deals with live subtitling; the second one (subsection 2.3) with pre-recorded subtitling; and the last one (subsection 2.4) with further research derived from the eye-tracking studies.

2.2 Live subtitling

From quantity to quality

The provision of live subtitles is a service many companies and broadcasters could do without. It is expensive, it requires skilled professionals and it is bound to be flawed.

Yet, the introduction of EU and national legislation on Subtitling for the Deaf and Hard of Hearing (SDH) means that live subtitling is no longer a privilege but a right for deaf and hard-of-hearing viewers, and therefore an obligation for subtitling companies and broadcasters. In many cases, this legislation sets targets of specific number of hours that must be subtitled (live and offline) depending on the country, type of channel, means available, etc. In other words, from the beginning the emphasis was placed on quantity. As a matter of fact, until recently, subtitling surveys often identified the lack of subtitles as the viewers' main concern regarding live programmes.

Now that respeaking seems to have consolidated as the most cost-effective method to provide live subtitles and companies and broadcasters are beginning to meet their targets, it may be time to change the focus from quantity to quality. In the UK, where the BBC already subtitles 100% of their programmes, it seems the obvious step forward. For other countries where live subtitling is still growing, it makes sense to apply quality standards now before "bad habits" are acquired.

But how do we measure quality in live subtitling and, in this case, in respeaking?

Most subtitling companies limit their quality assessment to error calculation, often carried out by trainers or respeakers as a respeaking skill to be applied after the process. Sometimes, this is completed with views gathered from the audience, be it through consultation with target groups or setting up an email address where viewers can express their opinion about respoken subtitles. However, an in-depth analysis of quality

in respeaking requires a different effort and the will to invest time and money on research. With the exception of some companies such as Swiss TXT, it is mostly scholars and research groups at university who embark upon this kind of research.

The aim of this section is to cast some more light on the quality of respoken subtitles by focusing on the viewers. For this purpose, this section includes a series of experiments about respoken subtitles in the UK that were recently carried out at Roehampton University as part of the research group Transmedia Catalonia and in the framework of the EU-funded project DTV4All (http://www.psp-dtv4all.org/). Having assessed the characteristics and the accuracy of respoken subtitles, attention is directed now to the audience and, in this case, to three different aspects: their comprehension of these subtitles, the manner in which they read/view them and their preferences/views.

2.2.1 Viewers' comprehension of respoken subtitles

Some surveys or opinion polls on subtitling include questions where participants are asked about the extent to which different types of subtitles enable comprehension. The problem in this case is that, as has been noted by many researchers (Tuominen 2008)[1], it is not uncommon to find discrepancies between opinion/preference and performance in this and other areas. In other words, the viewers' opinion about say a certain subtitling convention may be conditioned by different factors, such as for example habit, and does not always correspond to the convention that enables better comprehension. It is for this reason that the study of how much information viewers obtain through respoken subtitles needs a different experiment with a different approach. The following subsections include a description of such an experiment, the results obtained and their implications with regard to how hearing, hard of hearing and deaf viewers comprehend respoken subtitles.

2.2.1.1 Description of the experiment

The aim of this study was to find out how much visual and verbal information hearing, hard of hearing and deaf viewers obtain from news programmes in the UK. For this purpose, four clips from the *Six O'Clock News* broadcast on 4 July 2007 by BBC1 were shown to 30 hearing viewers, 15 hard of hearing viewers and 15 deaf viewers. The hearing participants were between 20 and 45 years old, native or near native in English, proficient readers and habitual subtitle users. Half of them were postgraduate students currently doing an MA on Audiovisual Translation at Roehampton University and the other half was formed by lecturers and professional subtitlers. The hard-of-hearing participants were over 60 years old, the most common age range for viewers with this type of hearing loss, and all of them but two became hard of hearing after the age of 50. Most of them were frequent readers and subtitle users. Finally, the deaf participants were between 20 and 45 years old. Most of them were oralist (i.e. use English as their first language) and only two were signing (use British Sign Language as their first language). All 15 were university students, frequent readers and habitual subtitle users.

As far as the methodology is concerned, participants were shown two clips with two news items each and were asked to answer questions about one of them. The clips were

[1] Tuominen, Tiina (2008), 'Reception or resistance? Some observations on the reception of subtitled films', paper presented at *Multidisciplinary Approaches*, University of Montpellier 3, on 19 June 2008.

subtitled by respeaking at two different speeds, 180 wpm, the usual speed in the UK, and 220 wpm, so as to ascertain the effect of speed on comprehension.

In order to carry out a quantitative analysis of the amount of information retrieved by the viewers, the two news clips were notionally divided, drawing on Chafe's (1980)[2] concept of *idea units*, into 14 semi-units: 8 verbal units and 6 visual units. In (very few) cases in which participants retrieved in their answers a semi-unit that was not included in these 14, the new unit was also factored in the analysis. For the purpose of the analysis of the findings, a simple division was made whereby any result between 0% and 25% is regarded as zero to poor information retrieval; 25%-50% goes from poor to sufficient; 50%-75% from sufficient to good; and 75%-100% from very good to perfect information retrieval.

Finally, a further problem was posed by the absence of a yardstick with which to compare the results obtained by participants watching subtitled news. Can we indeed expect viewers under normal conditions (no subtitles) to obtain 100% of the visual and acoustic information of a news clip? In order to answer this question, a preliminary test was run with 15 other students (from the above-mentioned class at Roehampton University) who watched the same clips with sound but no subtitles and were asked the same questions.

2.2.1.2 Findings

The following graphs and tables show the results obtained in the study, firstly with hearing participants and no subtitles and then with hearing, hard-of-hearing and deaf participants and subtitles at 180wpm and 220 wpm:

- No subtitles (hearing viewers)

No subtitles Performance	
Perfect	0%
Very good	93.3%
Good	6.7%
Almost good	0%
Sufficient	0%
Less than sufficient	0%
Poor	0%
Very poor	0%

[2] Chafe, Wallace L. (1980) "The Deployment of Consciousness in the Production of a Narrative", in Wallace L. Chafe (ed.) *The Pear Stories: Cognitive, Cultural and Linguistic Aspects of Narrative Production*. Norwood, N. J.: Ablex Publishing Corp, 9-50.

- Subtitles at 220 wpm (hearing, hard-of-hearing and deaf viewers)

	Hearing		Hard-of-Hearing		Deaf	
Good	0%	20%	0%	20%	0%	13.3%
Almost good	6.7 %		6.7%		6.6%	
Sufficient	13.3 %		13.3%		6.6%	
Less than sufficient	20%	80%	30%	80%	26.7%	86.7%
Poor	30%		30%		26.7%	
Very poor	30%		20%		33.3%	

- Subtitles at 180 wpm (hearing, hard-of-hearing and deaf viewers)

	Hearing		Hard-of-Hearing		Deaf	
Good	3.3%	46.7%	3.3%	46.7%	0%	46.7%
Almost good	6.7%		6.7%		6.7%	
Sufficient	36.7%		36.7%		40%	
Less than sufficient	20%	53.3%	20%	53.3%	13.3%	53.3%
Poor	20%		13.3%		20%	
Very poor	13.3%		20%		20%	

2.2.1.3 Discussion

As may be expected, hearing viewers watching the news with no subtitles did not manage to retrieve 100% of the visual and verbal information conveyed in the clips. Short term memory plays an important factor here. Yet, their results show very good comprehension (an average of 80%), particularly of the images (90.5%, as compared to 73.2% of the verbal information), which is normal considering that no subtitles were displayed.

As far as the study with subtitles is concerned, two elements are particularly striking: the overall poor average comprehension obtained and the similarity of the results across viewers regardless of the type hearing loss. The latter may be due to the fact that all participants taking part in the experiment were very used to watching subtitles on TV, be it because they study them or produce them (hearing) or because they use them as a means to access the news on a daily basis (deaf and hard-of-hearing). In any case, this makes the low overall score regarding comprehension even more puzzling.

As for the test with subtitles at 220 wpm, only 20% of the participants obtained sufficient information and none obtained good information. Besides, 60% could only give a poor or very poor account of the news. Although not surprising, given the high subtitle speed, these results warn against the possibility of producing verbatim subtitles for certain programmes such as debates, interviews and weather reports, which are sometimes spoken at this rate. Indeed, most viewers (76%) considered these subtitles to be too fast. Many of them also added that it caused them ‘stress’ and ‘headache’ and pointed out that the images were too fast, which, although not true (they were as fast as

in the other clips), goes to show how the speed of subtitles can affect the overall perception of an audiovisual programme.

The test with subtitles displayed at 180 wpm is more significant, as respoken subtitles are often displayed at this speed in some sport programmes and many news programmes, interviews and debates. In this case, most participants (66%) were happy with the speed of the subtitles and yet more than half of them (51%) did not obtain sufficient information. This suggests that viewers may be unaware of how much information they are losing due to the speed of respoken subtitles. Thus, although most of them regarded the speed as acceptable or even too slow, only 3% obtained good information and 31% got poor or very poor information. More worryingly, 1 out of 3 participants acquired incorrect information, believing, for example, to have seen the President of Nicaragua or Tony Blair, neither of whom appeared on the news.

Considering that these participants were highly literate and frequent subtitle users, viewers who are not used to subtitles or signing deaf viewers, for whom English is a second language and whose reading skills are often regarded to be poorer, can hardly be expected to obtain better results. Why do programmes with these respoken subtitles trigger such mediocre comprehension results? A possible answer to this question may lie in how viewers read and process these subtitles, which can be investigated by means of eye-tracking technology.

2.2.2 Viewers' processing of respoken subtitles

2.2.2.1 Eye-tracking and subtitling

Despite its obvious potential for the study of Audiovisual Translation and more specifically for that of subtitling, eye-tracking research in this area is still in its infancy. Following the initial studies by D'Ydewalle *et al*. (1987)[3] and Jensema et al. (2000)[4], it seems that an increasing number of scholars are turning their attention to this technology in order to find out how viewers read and comprehend subtitles and to assess their quality. Looking precisely at how subtitles are read, Jensema et al. (2000:284) found that

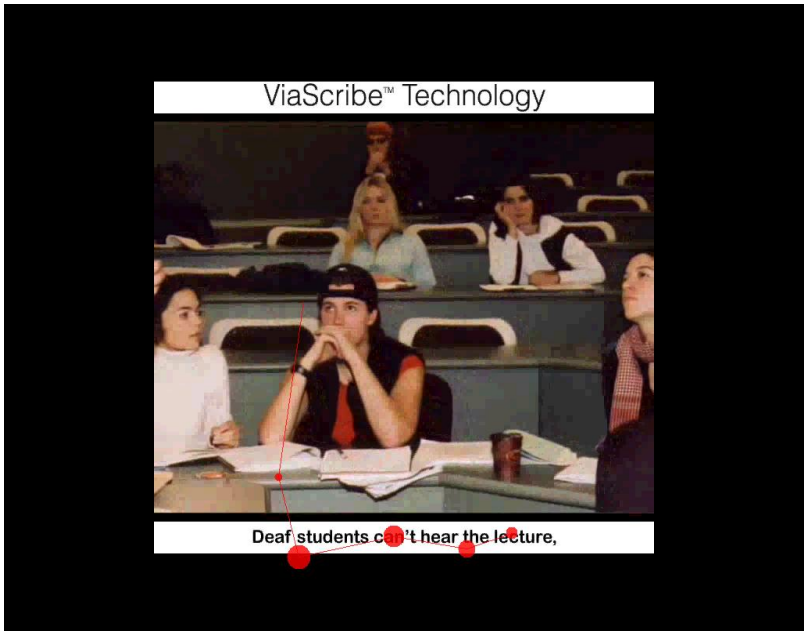
> "When captions are present, there appears to be a general tendency to start by looking at the middle of the screen and then moving the gaze to the beginning of a caption within a fraction of a second. Viewers read the caption and then glance at the video action after they finish reading."

Yet, reading is far from being a smooth process. Rather than moving continuously across the page/screen, our eyes pause and focus on specific parts and then jump across words and images. The visual information necessary for reading is obtained during those pauses, known as fixations, which typically last about 200–250 ms. The jumps between fixations are known as saccades, which take as little as 100 ms and are the

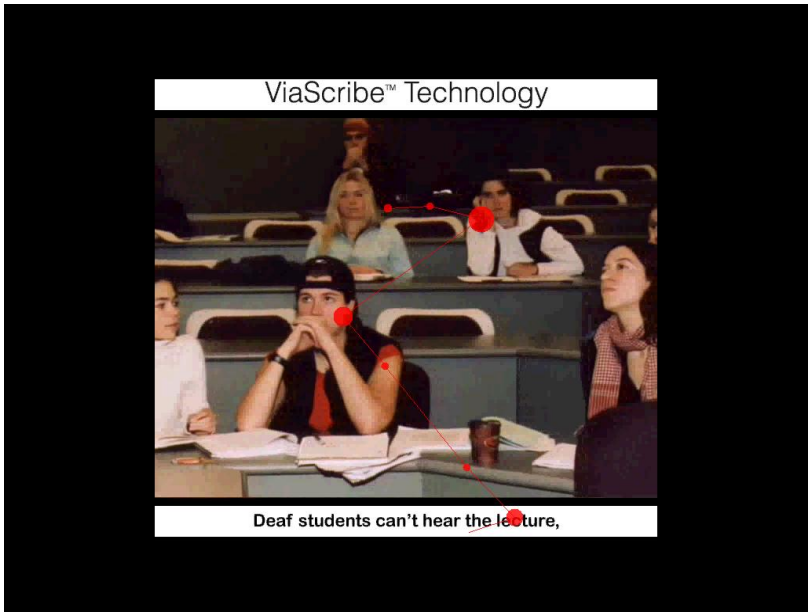
[3] D'Ydewalle, Gery, Johan van Rensbergen and Joris Pollet (1987) "Reading a message when the same message is available auditorily in another language: The case of subtitling", in John Kevin O'Reagan and Ariane Lévy-Schoen (eds.) *Eye Movements: From Physiology to Cognition*. Amsterdam / New York: Elsevier Science Publishers, 313-321.

[4] Jensema, Carl, Sameh El Sharkawy, Ramalinga S.Danturthi, Robert Burch and David Hsu (2000) "Eye movement patterns of captioned television viewers". *American Annals of the Deaf* 145(3), 275-85.

fastest movement the human being is capable of making. During saccades, vision is suppressed and no useful information is obtained, which is known as the saccadic suppression. But even though we cannot read during saccades, the eyes need not fixate on every word when reading a subtitle. In the following pictures, for example, reading the subtitled line in four fixations (picture 2.1) enables the viewer to turn quickly to the image (picture 2.2):



Picture 2.1



Picture 2.2

There has been no need to fixate on the words "students" or "hear" because a) they may be guessed by the context, particularly by the preceding words ("deaf" and "can't"), and b) they can be seen with peripheral vision, given that our global perceptual span, the area from which useful information is obtained during a fixation, comprises up to 14 or 15 characters to the right of a given fixation. In this regard, Rayner (1998)[5] explains that with the fovea (the part of the eye responsible for sharp central vision) we determine the location of a fixation, the foveal area, which spans 6 to 8 characters around the fixation point. But then, the so-called parafoveal area extends up to 15 characters to the right of fixation (Häikiö et al. 2009). This peripheral vision, which allows faster reading by not having to fixate on every word, applies to print and block subtitles. But what happens when we are reading subtitles that are displayed scrolling word-for-word on the screen, as is the case in respoken TV subtitles in the UK and in the above experiment on comprehension? How are these subtitles processed by the viewers?

Although not exactly applied to subtitles, the news coming from the field of psychology in this regard is discouraging. Experiments conducted by Rayner et al. (2006:321)[6] demonstrate "the importance of the continued presence of the word to the right of fixation […] in order for fluent reading to occur". It would seem that when our eyes are fixated on the foveal word (n), we have enough preview benefit of the next word, the parafoveal word (n+1), to pre-process it, which is crucial to maintaining normal patterns of reading behaviour. Needless to say, in scrolling subtitles, this word to the right of fixation, the n+1 word, is often unavailable for viewers, as words are displayed one at a time. In Rayner et al.'s (2006) study, the absence of this word causes regressions (the eye moves back to previous words already read) and considerable disruption to reading, slowing down reading speed significantly.

The aim of the following experiment is precisely to look at how viewers process respoken subtitles displayed in scrolling mode (as opposed to respoken subtitles displayed in blocks) to determine whether this may have been a contributory factor in the poor results obtained in the comprehension tests.

2.2.2.2 Description of the experiment

Conceived as an initial application of eye-tracking to research in respeaking, the present experiment was conducted with 30 of the 60 participants who took part in the comprehension tests described above: 10 hearing, 10 hard of hearing and 10 deaf viewers. Participants were shown two news clips from the *Six O'Clock News* (4 July 2004) subtitled by respeaking. The first clip was subtitled in scrolling mode (word-for-word); the second, in blocks. Eye movements were monitored via a non-intrusive tracker, which was used to determine a) the number of fixations per subtitled line and b) the amount of time spent on images as opposed to the time spent on subtitles. The equipment used was a Tobii X120 series eyetracker, at a frame rate of 50Hz and 35 ms latency. Viewing was binocular and the images were presented on a 17" monitor at a viewing distance of 60 cm. The computer kept a complete record of the duration,

[5] Rayner, Keith (1998) "Eye movements in reading and information processing: 20 years of research". *Psychological Bulletin* 124, 372-422.

[6] Rayner, Keith, Simon P. Liversedge and Sarah J. White (2006) "Eye movements when reading disappearing text: The importance of the word to the right of fixation". *Vision Research* 46, 310-23.

sequence, and location of each eye fixation, as well as a video recording of the participants. Tobii Studio was used to analyse all data recorded.

2.2.2.3 Findings

The following two tables show the results obtained in the study, namely the number of fixations per subtitled line and the time spent on both block subtitles and scrolling subtitles:

Number of fixations

	Blocks	Scrolling
Hearing	3.75	6
Hard-of-Hearing	3.75	6.5
Deaf	3.9	6.5

Time spent on images

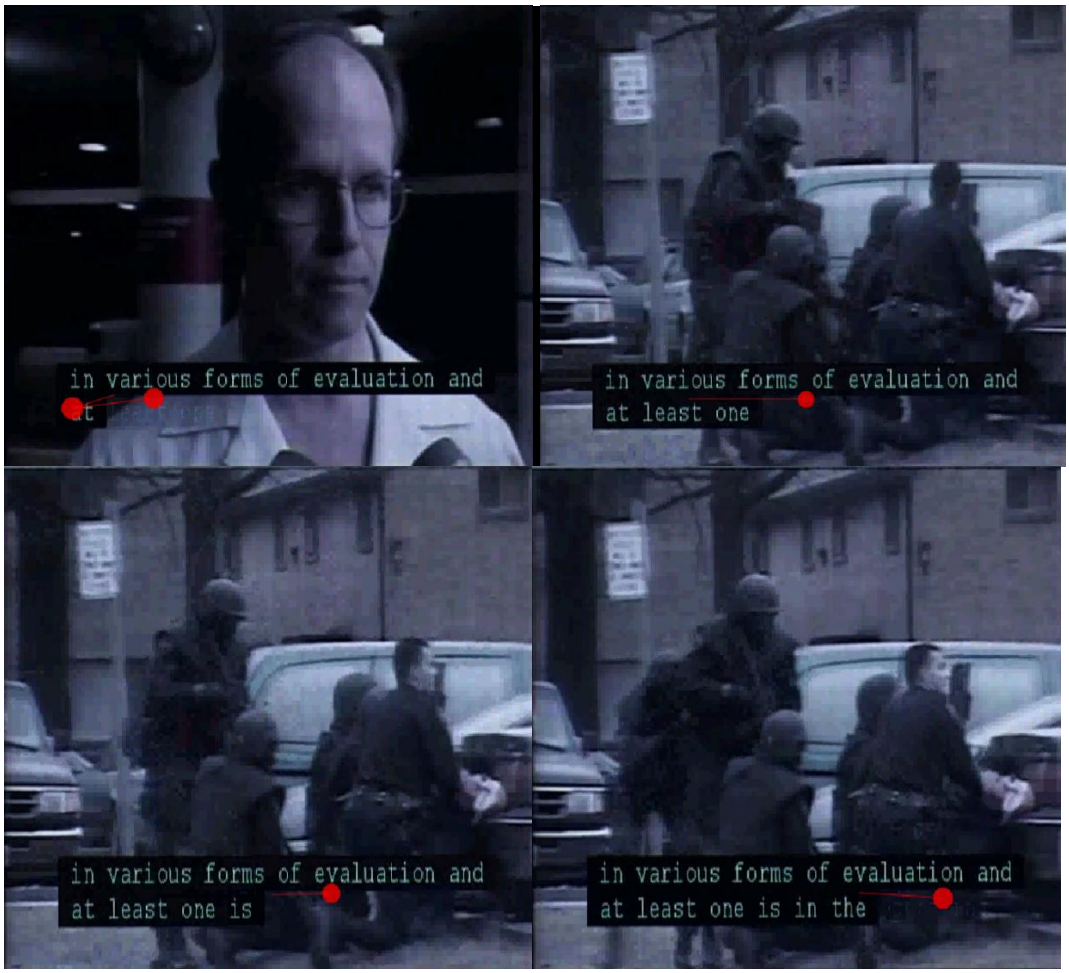
	Blocks	Scrolling
Hearing	33.3%	11.7%
Hard-of-Hearing	33.2%	11.4%
Deaf	31.7%	14.3%

2.2.2.4 Discussion

In line with what was described regarding the comprehension test, the results are fairly consistent across hearing, hard of hearing and deaf viewers. Scrolling subtitles cause almost twice as many fixations as block subtitles. The number of fixations per subtitled line in scrolling mode ranges from 3 to 10, with an average of 6 for hearing viewers and 6.5 for hard of hearing and deaf viewers. Given that the average number of words per line in the clips analysed is 6, it would seem that hearing viewers fixate on every word of every scrolling subtitle and deaf and hard of hearing viewers feature even more fixations than words. In contrast, the numbers of fixations in block subtitles ranges from 2 to 6, with an average of 3.75 fixations for hearing and hard of hearing viewers, and 3.9 for deaf viewers. In other words, viewers skip almost every other word of the subtitle when reading it. Needless to say, this has a direct impact on the time viewers spend looking at the subtitles and the time they devote to the images. Viewers of the scrolling mode spend most of their time reading the subtitles (an average of 87.5% versus 12.5% spent on the images), whereas viewers of block subtitles have more time to focus on the images (an average of 67.3% on the subtitles and 32.7% on the images).

The analysis of the reading patterns of every participant reveals another interesting element. Rather than differentiating the participants in hearing, deaf and hard of hearing, the results seem to establish a distinction between fast and slow readers.

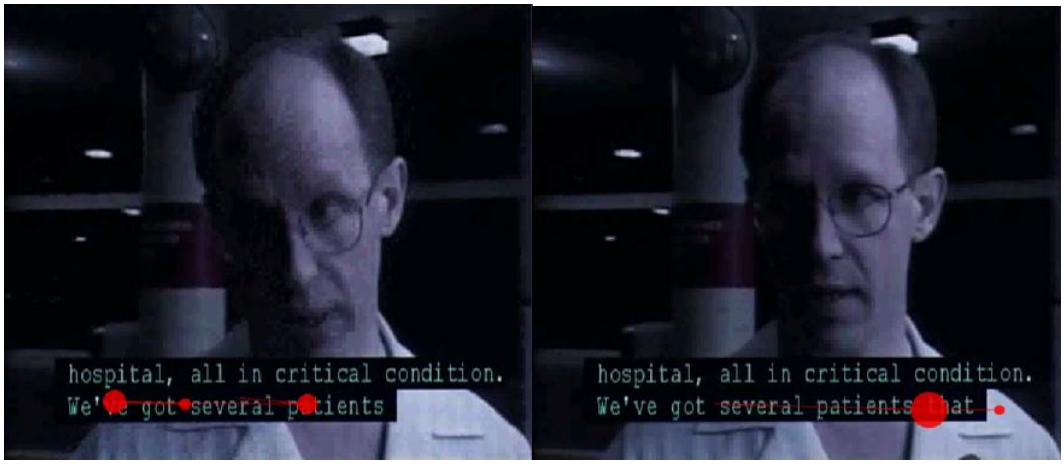
Besides, there seem to be two phenomena, astray fixations and regressions, that may explain the viewers' difficulty reading scrolling subtitles and perhaps the poor comprehension results obtained in the previous experiment. As for fast readers, they often get ahead of the subtitles and cast their eyes on gaps where no word has been displayed yet which results in astray fixations, instead of finding solid ground (a word or a whole line), the viewers' gaze falls on a sort of quicksand, which causes them to lose precious time in their reading process. In the following example, this "quicksand effect" occurs in four out of five attempts of the viewer to read the line 'at least one is in the operating room', the viewer ends up wasting a whole second (0.250 ms per each of the four astray fixations) when reading this line:



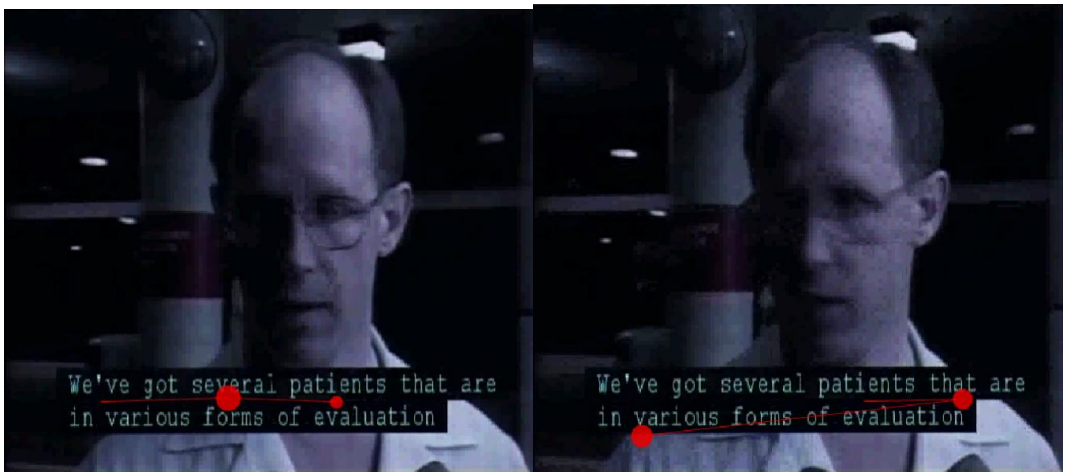
On average, these fast readers incur in 2 astray fixations per subtitled line. Half of the times this happens, they go back and re-read at least one word, which means they incur 1 regression per subtitled line. The other half of the time, they decide to go on reading the subtitle.

In contrast, slow readers do not get ahead of the subtitles (they usually lag behind them) and therefore their patterns do not feature astray fixations and the quicksand effect. However, their eyes often "land" on words in the middle of a subtitle which are not meaningful enough to make sense of what is being said. In order to go on reading, they have to go back and re-read previous words, which has happened 1.5 times per line in the subjects analysed.

In the following example, the viewer, who has been looking at the images, casts his/her eyes on the word "patients". Not being able to retrieve the information of the subtitled line by reading this word, s/he goes back to the previous one ("several") and yet one more time to ("we've got"), which finally provides enough information to go on reading after "patient". By then, though, the viewer has spent over a second reading a subtitle backwards:



In contrast to the chaotic patterns shown in scrolling subtitles, the reading pattern of block subtitles seems faster and more organised. Corroborating Jensema et al.'s (2000)[7] observations, viewers' gaze turns quickly to the subtitles, where this time they find firm ground on which to cast their eyes before looking up to the images. Thus, the same line as before ('we've got several patients that are') displayed in a block is read by a typical viewer in only four fixations (on 'we've', 'several', 'patients' and 'that'). There is no need to read all words and considerably less time is spent on the subtitle, which allows more time to focus on the image:



In other words, as anticipated in the literature on psycholinguistics and corroborated by the experiments reported here, it would seem that scrolling word-for-word subtitles cause very chaotic reading patterns. Fast readers get ahead of the subtitles and cast their

[7] Jensema, Carl, Sameh El Sharkawy, Ramalinga S.Danturthi, Robert Burch and David Hsu (2000) "Eye movement patterns of captioned television viewers". *American Annals of the Deaf* 145(3), 275-85.

eyes on gaps without words (astray fixations), whereas slow readers lag behind and constantly go back to re-read words (regressions). Either way they all waste precious time by chasing subtitles which seem to be playing hide-and-seek with them and which prevent them from looking at the images.

Needless to say, this chaotic reading pattern and the almost non-existent time left to look at the images may go some way towards explaining the poor comprehension results obtained by these participants in the comprehension test described.

What remains to be seen now is what viewers think about this and other types of respoken subtitles. Are they happy with them? Do they realise that this display mode may be hindering their comprehension of live programmes?

2.3 Pre-recorded subtitling

2.3.1 Introduction

This subsection presents the results of eye-tracking tests carried out in the UK for pre-recorded subtitles as part of the DTV4All project. The subsection draws on the following data:

1. Information derived from questionnaires collected before the eye-tracking tests (pre-test questionnaire)
2. Eye-tracking data
3. Results of comprehension questions obtained during eye-tracking tests
4. Information on preferences concerning particular subtitling variables in each of the parameters tested in the study (post-test preferences)
5. Background information on subtitling practices in the UK

2.3.2 Participants

20 volunteers participated in the study: 13 men and 7 women.

2 participants were excluded from the statistical analysis because of technical problems (it was not possible to obtain any eye-tracking data due to the glasses they were wearing).

Participants were initially divided into three groups: 1) deaf, 2) hard of hearing and 3) hearing, based on their self-description of hearing loss. Five people were considered deaf (3 men and 2 women), 10 described themselves as hard of hearing (8 men and 2 women) and 5 participants were hearing (2 men and 3 women).

The participants were divided into four age groups:

Age	**Deaf**	**Hard of hearing**	**Hearing**
15-24	2	0	2
25-39	1	2	2
40-59	1	2	1
60+	1	6	0
TOTAL	**5**	**10**	**5**

All participants attended both primary and secondary school. 13 of them have university degrees. 10 of them are currently studying and the other 10 are unemployed or retired.

2.3.3 Methodology

2.3.3.1 Tests

Upon arrival at the laboratory, each participant was given a pre-test questionnaire to complete. The questionnaire included a number of questions pertaining to personal information concerning the participant, such as hearing loss, age, education as well as preferences and general views on subtitling. The questionnaire can be found here:

https://rnid.wufoo.com/forms/what-are-your-views-on-tv-subtitling/

After completing the questionnaire, participants were introduced to the procedure of the eye-tracking test, including the test structure and the calibration process. The eye-tracking test consisted of nine parameters, with three or two variables per parameter (presented in the following order):

1. Character identification
    a. Colours
    b. Tags
    c. Speaker-dependent placement
2. Subtitling style
    a. Verbatim
    b. Standard
    c. Edited
3. Position
    a. Bottom
    b. Mixed
    c. Top
4. Emotions
    a. Description
    b. Emoticons
    c. Nothing
5. Sounds
    a. Description
    b. Icon
    c. Nothing
6. Justification
    a. Left-aligned
    b. Centred
7. Borders
    a. Borders
    b. No borders
8. Box
    a. Box
    b. No box
9. Shadows
    a. Shadows
    b. No shadows

Altogether, there were 23 clips. After watching each clip, the participant had to answer three comprehension questions: 1) about general understanding of the clip, 2) about textual elements included in subtitles, and 3) about visual elements in the clip. After each parameter, the participant was asked to choose the most preferred version (i.e. variable) and to comment on their preferences. On average, it took from 60 to 90 minutes for each participant to complete the entire procedure. Unfortunately, due to several problems with the calibration and the overall performance of the eye-tracker, a great deal of the information obtained could not be used. As a result, further tests were scheduled the resulting data is still being analysed. Included here are the two parameters, subtitling position and character identification, which yielded reliable data.

2.3.3.2 Eye-tracking data

Eye movements were recorded with a Tobii X120 series eyetracker, at a frame rate of 50Hz and 35 ms latency. Viewing was binocular and the images were presented on a 17" monitor at a viewing distance of 60 cm. The computer kept a complete record of the duration, sequence, and location of each eye fixation, as well as a video recording of the participants. Tobii Studio was used to analyse all data recorded. Participants were presented with 23 clips (each lasting about 1 minute) during which their eye movements were recorded. After three or two clips (i.e. one parameter), participants answered additional three questions concerning their subtitles.

Eyetracking data were analyzed on the basis of three different criteria: time to first fixation, observation length and fixation count.

Time to first fixation was calculated in milliseconds (ms) and defined as the time which elapsed between subtitle onset and the first fixation that entered the area of interest (AOI) with subtitles. Since in each clip subtitles appeared at different times, specific Interest Periods were created; thus, time to first fixation reflected how much time after the subtitle onset it took the participants to look at the subtitle AOI.

Observation length: also measured in ms, the observation length indicates the period of time that participants spent on reading subtitles compared to the time spent on watching the whole clip. It provides data on reading speed, which may lead to further research.

Fixation count: in this case, the data obtained provides information on the average number of fixations on subtitles vs. the number of fixations on images as well as the average number of characters and words per fixation.

2.3.4 Results

2.3.4.1 Subtitle position

Subtitles in the UK are usually positioned at the bottom of the screen. This applies both to interlingual subtitles for hearing viewers in cinemas and on DVDs as well as SDH on television.

Bottom – the standard position of subtitles in the UK. Subtitles are sometimes moved to the top in order not to cover an important piece of information placed at the bottom, be it a caption, notice or other crucial element.

Mixed – a combination of subtitles containing dialogue placed at the bottom of the screen with a description of sounds placed at the top. Non-existent on the UK audiovisual market, this style of subtitling was a complete novelty to the participants.

Dialogue	Sound information
	

Top – all subtitles, including dialogue and sound description, placed at the top of the screen, as is the case in some live programmes such as football matches.

2.3.4.2 Comprehension

As was expected, bottom subtitles scored better than the other two positions. Subtitles with mixed position had very low comprehension and top subtitles had very good results, particularly among young and deaf viewers.

	Bottom	Mixed	Top
All participants	75%	62%	73%

	Bottom	Mixed	Top
Hearing	73%	65%	74%
Hard of hearing	74%	58%	71%
Deaf	78%	65%	74%

	Bottom	Mixed	Top
Young	77%	69%	76%
Old	73%	55%	70%

2.3.4.3 Eye-tracking data

Time to first fixation (ms)

	Bottom	Mixed	Top
All participants	296	352	305

	Bottom	Mixed	Top
Hearing	290	360	301
Hard of hearing	299	328	314
Deaf	285	290	280

	Bottom	Mixed	Top
Young	250	282	272
Old	305	335	320

Bottom subtitles have generally faster reaction times, with the exception of top subtitles for deaf participants. The reason for this pattern seems quite obvious: bottom position is the one where people expect subtitles to appear and this is where they look in search for subtitles. Mixed subtitles consistently require more time for the eye to react. As for top subtitles, considering that they are not very common, they yield very good results which support its use in certain types of programmes. Age, rather than hearing ability, seems to be the most determining factor to group the different types of viewers.

Mean reading time

	Bottom	Mixed	Top
All participants	51.34%	54.1%	53.18%

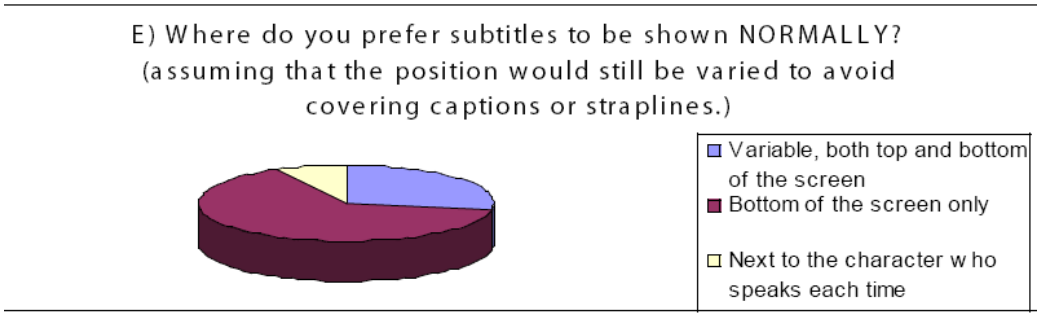
	Bottom	Mixed	Top
Hearing	50.2%	53.1%	52.1%
Hard of hearing	50.1%	53.1%	53.3%
Deaf	53.3%	55.1%	54.6%

	Bottom	Mixed	Top
Young	48.4%	52.6%	52.1%
Old	54.4%	56.2%	54.3%

Bottom subtitles lend themselves to faster reading than mixed and top subtitles, the latter being once again positioned as second best option. As was the case regarding time to first fixation, age proves more important than hearing ability.

2.3.4.4 Preferences

Pre-test preferences



Post-test preferences

	Bottom	Mixed	Top
All participants	17	0	3

	Bottom	Mixed	Top
Hearing	4	0	1
Hard of hearing	10	0	0
Deaf	3	0	2

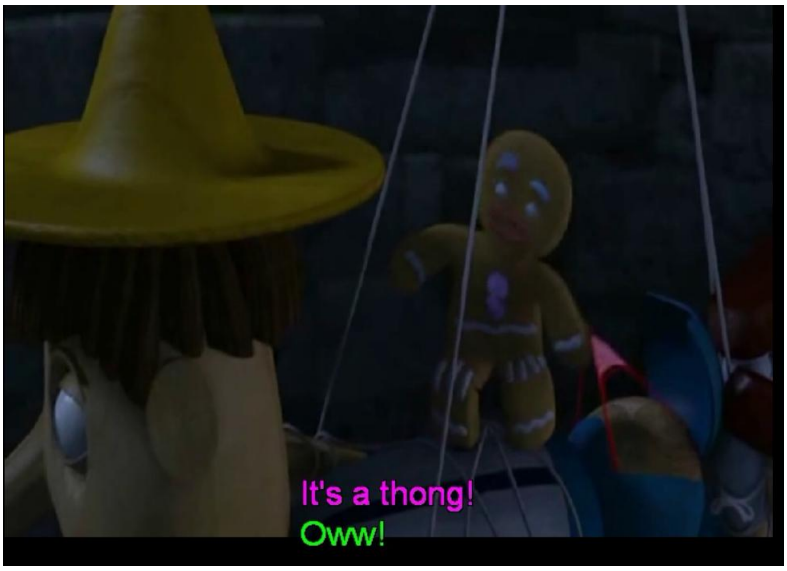
	Bottom	Mixed	Top
Young	3	0	3
Old	14	0	0

Post-test preferences regarding subtitle position are fairly similar to the views expressed by these and other participants before the eye-tracking tests with bottom being the preferred choice followed by top and mixed position.

2.3.4.5 Character ID

For this parameter, the following three variables were tested:

Colours



Colours are frequently employed in SDH on UK TV. The seven colours in use on public TV are white, yellow, cyan, green, magenta, red and blue.

Speaker-dependent placement

Speaker-dependent placement is used on UK TV on certain occasions, almost always combined with colours.

Tags

Tags are normally used in DVDs but they are not as common on TV in the UK.

Comprehension

The variable which rendered the highest comprehension scores was colours, closely followed by tags and displacement. This is not surprising considering that viewers are used to all three variables when watching SDH on TV (colours and displacement) and DVD (tags).

	Colours	Displacement	Tags
All participants	72%	70%	71%

	Bottom	Mixed	Top
Hearing	73%	72%	70%
Hard of hearing	70%	68%	73%
Deaf	73%	72%	68%

	Bottom	Mixed	Top
Young	76%	74%	71%
Old	68%	66%	71%

2.3.4.6 Eye-tracking data

Time to first fixation (ms)

	Colours	Displacement	Tags
All participants	440	355	390

	Colours	Displacement	Tags
Hearing	440	356	425
Hard of hearing	460	359	315
Deaf	420	352	430

	Colours	Displacement	Tags
Young	350	314	416
Old	530	396	346

The longest time to first fixation was obtained attained by participants watching subtitles with colours. Tags and especially displacement were noticed faster. Younger viewers seem to perform better with displacement whereas older viewers have shorter reaction time with tags.

Mean reading time

	Colours	Displacement	Tags
All participants	53.84%	55.1%	57.8%

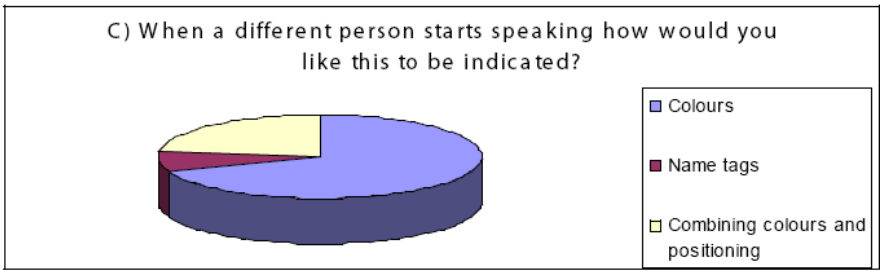
	Colours	Displacement	Tags
Hearing	53.81%	54.81%	57.7%
Hard of hearing	54.19%	56.3%	57.8%
Deaf	53.72%	55.1%	57.9%

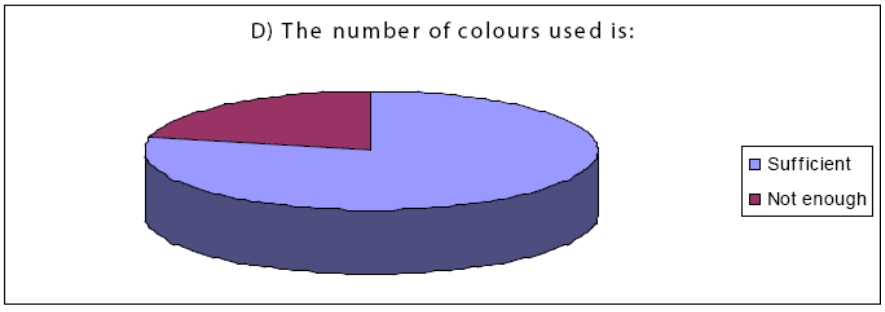
	Colours	Displacement	Tags
Young	51.7%	51.9%	54.9%
Old	55.9%	59.1%	60.1%

Participants spent more time reading subtitles with tags than with displacement and colours, the latter yielding the shortest reading times. This may be due to the fact that the use of colours is the standard convention on TV in the UK, sometimes combined with displacement. Viewers take more time to read subtitles with tags, probably because the subtitles themselves were longer (tag + subtitle) and different from the standard TV subtitles.

2.3.4.7 Preferences

Pre-test preferences





Post-test preferences:

	Colours	Displacement	Tags
All participants	11	2	7

	Colours	Displacement	Tags
Hearing	3	2	0
Hard of hearing	3	0	7
Deaf	5	0	0

	Colours	Displacement	Tags
Young	6	0	0
Old	5	2	7

Although colours are still the first choice to identify characters, after the eye-tracking tests some viewers, mostly hard of hearing, chose tags as their first choice. This is in line with some of the data obtained in the pre-test long questionnaire, according to which hard of hearing people are more likely to choose traditional SDH (tags) rather than innovative conventions (in this case displacement).

2.4 A new take on reading speed: watching speed

Among the most commonly debated topics of discussion in the subtitling literature, speed has always occupied a privileged position. This may be explained by the fact that it is the speed of subtitles that determines whether they can be verbatim or edited. Fast subtitles can convey every single word of the dialogue whereas slower subtitles typically summarise or condense what is being said. Often considered very important in "standard" subtitling (interlingual subtitling for hearing viewers), this issue becomes critical when applied to subtitling for the deaf and hard of hearing (SDH), hence Ofcom's (2005:11)[8] description of speed as "arguably the key underlying issue behind nearly every important issue" in SDH. Speed in SDH is as much a technical matter as it is economic (broadcasters, service providers), political and ideological (deaf associations).

Firstly, broadcasters, under pressure to provide more SDH, support verbatim subtitles, as they require less effort on the part of the subtitlers and are thus more economical than edited subtitles. Secondly, and surprisingly, most deaf viewers (or rather deaf associations) also demand verbatim, and therefore faster, subtitles. In this case, the reason is not financial, but political. There is among these viewers a great deal of sensitivity and antagonism towards the idea of editing, regarded as "a form of censorship and 'denying' deaf people full access to information available to the hearing population" (Ofcom, 2005:17). Finally, a third group is formed by scholars and researchers, the only ones who usually support edited subtitles. They often agree with Sancho-Aldridge (1996:24)[9], who calls for the need to "disentangle the politically sensitive issue of 'access' from the practical issue of which style, in real terms, provided deaf viewers with most information". Among scholars, there seems to be consensus

[8] OFCOM (2005), Subtitling – An Issue of Speed? London: Office of Communications.

[9] Sancho-Aldridge, Jane and IFF Research Ltd (1996) *Good News for Deaf People: Subtitling of National News Programmes*. London: Independent Television Commission.

as to the fact that verbatim subtitles are often too fast to provide full access for many deaf viewers.

As a result, the paradox remains that whereas scholars support editing (and thus slower subtitles) to provide full access for the deaf, the latter line up with broadcasters to push for verbatim, which may not give them full access.

At the centre of this controversy is the notion of **reading speed**, which supporters of verbatim subtitles push to 180wpm. In contrast, those who prefer edited subtitles recommend lower speeds of approximately 150wpm. The problem in this case is that this concept of reading speed does not come from subtitling, but from psycholinguistics, where it has always been applied to reading print. Useful as it may be, its application to subtitling should account for the audiovisual nature of this new medium, which has so far never been the case. Let's dwell on the difference:

Reading speed in a book: a given person reads text in a book while comprehending it (answering comprehension questions correctly). The reading speed this person has chosen to read the text at is the figure we're looking for.

Reading speed in subtitling: a given person watches a subtitled programme on a screen (text plus images) while comprehending it (answering comprehension questions correctly). The reading speed has already been set by the subtitler.

In this case, and as has been the case traditionally, experiments on reading speeds in subtitling have been conducted as follows:

Viewers watch programmes subtitled at different speeds and then answer comprehension questions. If comprehension is good, then reading speeds are fine.

But how about the images? Were viewers able to watch the images as well as reading the subtitles? If most questions are usually about the subtitles, how do we know that they didn't spend most of their time bogged down in the subtitles? Could it be the case that a given speed yields decent comprehension results and at the same time uncomfortable reading for the viewers?

Now that eye-tracking technology is available for research on subtitling, perhaps it's time to introduce the audiovisual factor in the study of reading speed. In other words, as well as different speeds and comprehension, we can take into account the amount of time devoted to images and subtitles.

Instead of considering "x speed yields y comprehension", we could consider "x speed yields y comprehension and z time on images and subtitles". This could be called **watching speed**, the speed at which we watch a scene with subtitles and images. Reading speed would be left for the time it takes to read the actual subtitle.

Having looked at hundreds of subtitled scenes watched by the participants in the present experiment, a pattern has emerged, whereby a given watching speed of a subtitled scene yields similar percentages in terms of the time devoted to subtitles and images:

Watching speed	Time on subtitles	Time on image
120wpm	40%	60%
150wpm	50%	50%
180wpm	60-65%	40-35%
200wpm	80%	20%

Needless to say, there are plenty of exceptions, but many more are the cases which conform to this pattern. The reading or in this case watching speed recommended in traditional guidelines, 150wpm, seems to provide viewers with the same amount of time to enjoy subtitles and images. Higher speeds near 180wpm, common both on TV and on DVD these days, break this balance forcing viewers to spend 60-65% of their time on the subtitles and only 40-35% on the images. Faster subtitles at 200wpm may leave viewers with as little as 20% of the time to devote to the images. It is important to highlight here that although subtitles at 120wpm allow viewers to spend more time on the images than on the subtitles, lower speeds (usually slower than 100wpm) increase the time the subtitles appear, which remain on screen for too long and are thus re-read by the viewers.

Reading speeds, that is, the time devoted to reading subtitles, excluding images, range between 300wpm and 400wpm or even 500 wpm (for instance for overviewing a text), which is in line with what has so far been found in the literature on psycholinguistics about print reading.

Needless to say, the table presented here is only preliminary, based on a reduced number of participants and does not include the many exceptions found, such as different types of subtitles, viewers, viewing conditions, etc. Yet, all these exceptions and reservations can also be applied to the notion of reading speed as it has traditionally been applied to subtitling, which, as opposed to the proposed watching speed, does not take into account the visual elements.

The aim of this work was to

a) confirm (or refute) the above patterns of watching speed – time on subtitles/images with further analysis of the data obtained in the UK and the other countries involved in DTV4All.

b) Apply these patterns to the professional practice of subtitling.

The second step has already been tested in the subtitling module of the MA in Audiovisual Translation at Roehampton University. Students were given the above table. When subtitling continuous and rather fast dialogue, they chose different reading (or rather watching) speeds not only on the basis of how much they wanted to edit their subtitles, but also on the importance of the images. For instance, dialogue in shot changes showing new images were subtitled at 150wpm. Some editing was required, but many students deemed it preferable as it would allow viewers to spend 50% of their time on the new images. Subsequent subtitles shown on the same image were subtitled at 180wpm, no editing required, as providing viewers with only 35-40% of the time to watch images was not a problem, given that the images were already known.

Should the data included in the above table be confirmed, the idea would be to present it as one more element to take into account when determining the (watching) speed of subtitles: different speeds not only involve more or less editing but also more or less time on the subtitles/images.

2.5 Conclusions

As far as live subtitling is concerned, now that respeaking has consolidated as the preferred method and that many broadcasters are meeting the targets set by European and national legislation, it may be time for research in this field to focus on the quality rather than on the quantity of real-time subtitles. Viewers seem to share this view as their complaints about lack of subtitles in live programmes take a back seat to other issues such as the delay of respoken subtitles, the number of mistakes, etc.

Equally important is to adopt a broad approach to the assessment of the quality of live subtitling. Significant as they may be, the viewers' preferences are but one element to take into account. This may be complemented by other elements such as comprehension tests or eye-tracking studies that can cast some light on the extent to which live subtitles are understood or on how they are viewed / perceived. In the case of the respoken subtitles currently provided in the UK, the tests included in this report suggest that there is much room for improvement. The results obtained by the hearing, deaf and hard of hearing participants in the comprehension tests are worryingly low. As shown by the eye-tracking study, this may be due to the scrolling display mode of respoken subtitles, which causes unnatural and chaotic reading patterns, the viewers chasing the subtitles and having no time to focus on the images.

With regard to pre-recorded subtitles, the comprehension and eye-tracking tests conducted for this study favour the use of bottom subtitles instead of mixed subtitles, standard in countries such as Spain. As for top subtitles, considering that they are not very common, they have yielded very good results, which support its use in certain types of programmes such as sports. Age, rather than hearing ability, seems to be the most determining factor to group the different types of viewers.

In terms of character identification, the variable which rendered the highest comprehension scores was colours, closely followed by tags and displacement. This reflects the particular reality of SDH in the UK, where viewers are used to all three variables when watching SDH on TV (colours and displacement) and DVD (tags).

Finally, the eye-tracking study has revealed a series of common patterns regarding viewers' reading speed. This suggests that this controversial issue, which has traditionally been tackled on the basis of whether different speed yield better or poorer comprehension, could take into account one more essential element: the extent to which different speeds allow viewers to spend more / less time on the subitle / images.

3 The Polish Evaluation (University of Warsaw)

Agnieszka Szarkowska
Izabela Krejtz
Anna Wieczorek
Zuzanna Kłyszejko
Joanna Pietrulewicz

Introduction

This section presents the results of eyetracking tests carried out in Poland as part of the DTV4All project. The section draws on the following data:

- Information derived from questionnaires collected before the eyetracking tests (pre-test questionnaire)
- Eyetracking data
- Results of comprehension questions obtained during eyetracking tests
- Information on preferences concerning particular subtitling variables in each of the nine parameters tested in the study (post-test preferences)
- Background information on subtitling practices in Poland

Method

Upon arrival at the laboratory, each participant was given a pre-test questionnaire to complete. The questionnaire included a number of questions pertaining to personal information concerning the participant, such as hearing loss, age, education as well as preferences and general views on subtitling.

After completing the questionnaire, the participant was asked to sign an informed consent to take part in the study and was then introduced to the procedure of the eyetracking test, including the test structure and the calibration process. The eyetracking test consisted of nine parameters, with three or two variables per parameter (presented in the following order):

1. Character identification
    a. Colours
    b. Tags
    c. Speaker-dependent placement
2. Subtitling style
    a. Verbatim
    b. Standard
    c. Edited
3. Position
    a. Bottom
    b. Mixed
    c. Top
4. Emotions
    a. Description
    b. Emoticons
    c. Nothing
5. Sounds

    a. Description
    b. Icon
    c. Nothing
6. Justification
    a. Left-aligned
    b. Centred
7. Borders
    a. Borders
    b. No borders
8. Box
    a. Box
    b. No box
9. Shadows
    a. Shadows
    b. No shadows

Altogether, there were 23 clips. After watching each clip, the participant had to answer three comprehension questions: 1) about general understanding of the clip, 2) about textual elements included in subtitles, and 3) about visual elements in the clip.

After each parameter, the participant was asked to choose the most preferred version (i.e. variable) and to comment on their preferences.

On average, it took from 60 to 90 minutes for each participant to complete the entire procedure.

3.1 Participants

Forty-two volunteers participated in the study: 24 women and 18 men.
Three participants were excluded from the statistical analysis because of technical problems (e.g. one person had the involuntary eye movement condition known as nystagmus).

3.1.1 Hearing loss

The participants were divided into three groups: 1) deaf, 2) hard of hearing and 3) hearing, based on their self-description of hearing loss. As a result, 12 people were considered deaf (7 women and 5 men), 19 described themselves as hard of hearing (8 women and 11 men) and 11 participants were hearing (3 men and 8 women).

3.1.2 Age

The participants were divided into the four age groups:

Age	**Deaf**	**Hard of hearing**	**Hearing**
15-24	1	6	3
25-39	9	8	4
40-59	2	1	3
60+	0	4	1
TOTAL	**12**	**19**	**11**

Participants by age (n=42)

3.1.3 Education

All of our participants attended both primary and secondary school. 15 people declared themselves to be graduates of higher education institutions.

In the group of deaf people, six had secondary school education (two of them attended deaf schools) and four had higher education while one is still studying. All the hard of hearing participants were educated in secondary schools (two of them went to deaf schools) and seven out of 12 completed studies, while seven are still studying. Seven hearers are secondary school graduates and five have an academic degree. Five are still studying.

Three participants are unemployed and seven retired.

3.2 Eyetracking tests

3.2.1 Apparatus

Eye movements were recorded with an EyeLink CL eyetracking system (SR Research Ltd.). The EyeLink system uses infrared, video-based technology to compute the distance between the pupil centre and corneal reflection. Signals were generated by the eyetracker at a frequency rate of 500 Hz when a stable corneal reflection was obtainable from a participant, allowing fixation latencies to be calculated with 2 ms temporal resolution.

The experiment was run under Microsoft Windows Vista in Experiment Builder (SR Research Ltd). The program also recorded button press accuracy for comprehension questions.

3.2.2 Procedure

Stimuli were presented on a 17-inch LCD colour monitor with 60 Hz refresh rate. Participants were seated in a chair in front of the monitor positioned at an eye level at a viewing distance of approximately 60 cm, maintained by a forehead and chin-rest. Participants were presented with 23 clips (each lasting about 1 minute) during which their eye movements were recorded. After two or three clips (i.e. one parameter), participants answered additional three paper and pencil questions concerning their subtitle preferences. Then recalibration was performed. Drift corrections were performed before each trial.

3.2.3 Analysis of eye tracking and behavioural data

Eyetracking data were analyzed for fixations and saccades using the Eyelink DataViewer (SR Research Ltd). For each trial, areas of interest (AOI) were drawn based on prior hypotheses. We compared mean reading time and time to first fixation to the AOIs. For the behavioural data, the percentage of correct answers to comprehension questions was calculated.

3.2.3.1 Time to first fixation

Time to first fixation was calculated in milliseconds (ms) and defined as the time which elapsed between subtitle onset and the first fixation that entered the AOI with subtitles. Since in each clip subtitles appeared at different times, specific Interest Periods were created; thus, time to first fixation reflected how much time after the subtitle onset it took the participants to look at the subtitle AOI.

3.2.3.2 Mean reading time

Mean reading time was defined as the period of time that participants spent on reading subtitles compared to the time spent on watching the whole clip. Precisely, we calculated mean dwell time, i.e. the sum of durations of all the fixations, in the subtitle AOI and compared it to the dwell time of all fixations across the whole clip.

Results

3.3 Overall comprehension

Out of the three groups of respondents, hearing people had the highest rates of correct answers (72%). Hard of hearing respondents scored slightly higher than deaf respondents, 66% and 64% respectively. These results may be attributed to the fact that hearing participants and to a lesser degree the hard of hearing obtain information from two major channels: visual and auditory, while deaf participants only have the former channel at their disposal. In other words, deaf people cannot take advantage of what is known as intersemiotic redundancy, i.e. excess of information between different communication channels, and thus they cannot compensate for the information they miss out on in the visual channel.

What is more, the *Shrek* trilogy that was used as a basis for the SDH in our tests is well-known and a large number of hearing participants had seen the dubbed versions of these films. In contrast, few hearing-impaired respondents had watched the films, which might have had a bearing on the final results of comprehension questions.

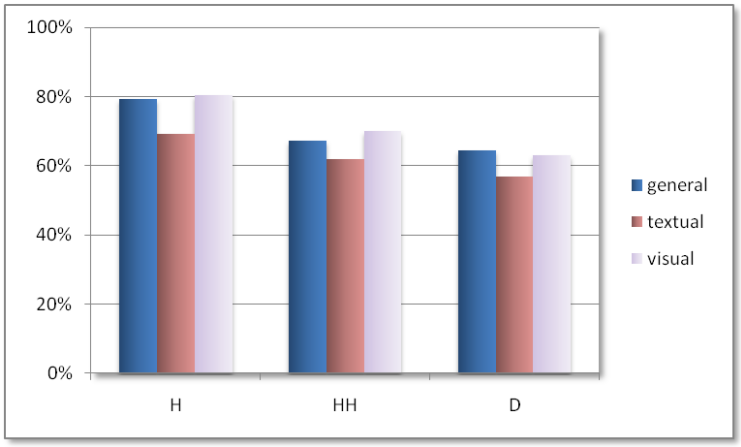
Hearing	Hard of hearing	Deaf
72 %	66 %	64 %

Comprehension: overall results from the three respondents groups (H/HH/D)

As stated above, there were three types of comprehension questions, relating to 1) general understanding, 2) textual elements included in subtitles and 3) visual elements in the image. The most difficult type of question turned out to be those relating to textual information, with less than two-thirds of the answers to these questions being corrrect.

General	Textual	Visual
70 %	62 %	71 %

Comprehension: Percentage of correct answers by question type



Comprehension questions: percent of correct answers

	Hearing	Hard of hearing	Deaf
General	79 %	67 %	64 %
Textual	69 %	61 %	56%
Visual	70%	71 %	71 %

Comprehension: Percentage of correct answers by type of question

Our initial hypothesis was that comprehension, especially in the case of deaf and hard of hearing viewers, is dependent on their looking at (i.e. fixating on) particular areas of interest. Put differently, we assumed that if a person looked at a particular visual element or a word in a subtitle, the comprehension score of that person is likely to be higher than that of a person who did not look at that element. To illustrate this, let us have a look at the following example with a shot of Puss in Boots at the beginning of the clip with standard style subtitles. In the comprehension question on visual elements we asked who was standing outside the window. Analysis using Pearson's correlation[10] allowed us to determine that in the case of deaf participants, there was a strong and significant correlation between looking at the cat and answering the comprehension question correctly. However, in the case of hearing as well as hard of hearing participants, there was no such correlation.

[10] r=0.68, p<0.05

The comprehension scores for this example, in terms of percentage of correct answers, for this visual element were: 54% for hearing viewers, 45% for hard of hearing viewers and 44% for deaf viewers. These scores are significantly lower than those for the questions on general understanding and textual elements. Another factor coming to play here is memory: despite having seen the cat, some participants may have forgotten about this by the time they reached the comprehension questions.

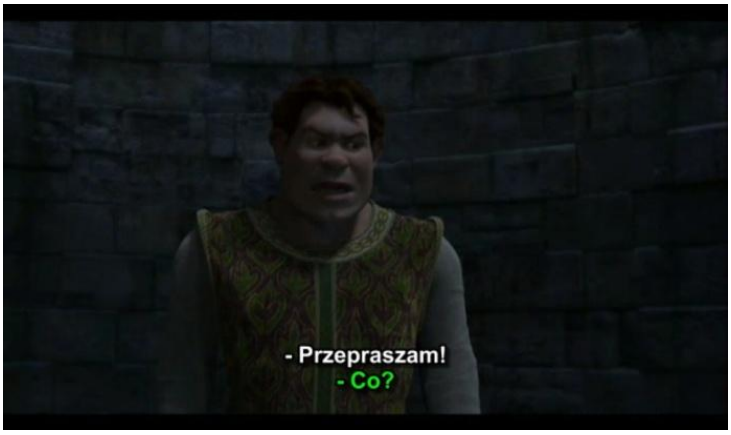
One needs to be wary, however, about jumping to hasty conclusions on the direct correlation between fixating on an item and its comprehension. This problem is best illustrated by the following example. In one of the clips, Shrek is touching his belly and this is what we asked in the comprehension question: (*Which part of his body is Shrek touching?*).

As a result, in the hard of hearing group, we obtained a significant *negative* correlation between fixating on our area of interest and answering the comprehension question correctly, this means that a hard of hearing person was likely to answer the question correctly if they were NOT looking at the picture. This may be explained by the fact that the subtitle includes sound information coming from the belly ('GURGLING IN STOMACH') so it may have been enough for respondents to read the subtitle to be able to answer the question correctly.

1. Character identification

For the colour identification parameter, the following three variables were tested:

1) Colours



Colours are frequently employed in SDH on Polish TV. The three colours in use on public TV are: yellow, green and blue; therefore, these colours were used in our tests.

2) Tags

Tags are not usually used in SDH on Polish TV. This variable was the one that viewers were mostly unfamiliar with.

3) Speaker-dependent placement

Speaker-dependent placement is used in SDH on Polish TV in the case of feature films and TV series. Most of the time it is combined with colours, i.e. colour subtitles are placed either to the left or to the right of the screen, depending on the speaker's position. Subtitles are always placed at the bottom of the screen, here our test differed slightly from the one carried out in Spain, where displaced subtitles were also moved vertically more to the centre of the screen (this observation may be important in measuring time to first fixation).

3.3.1 Comprehension

The variable which rendered the highest comprehension scores was tags, followed closely by speaker-dependent placement. This is quite surprising considering the unusual structure of such subtitles: a name tag plus dialogue. In the opinion of many viewers, however, it was subtitles with names tags that were the most unequivocal.

	Colours	Tags	Placement
Hearing	46 %	79 %	61 %
Hard of hearing	53 %	62 %	60 %
Deaf	52 %	63 %	59 %

Comprehension: Character ID (%)

The low comprehension rates of subtitles with colours may stem from the fact that it was the very first clip to be shown to the participants and although they were told in advance they would have to answer comprehension questions, many appeared slightly surprised at the level of detail we were asking about.

More detailed analysis of particular questions has revealed that the question on textual comprehension turned out to be significantly more difficult than others, with a comprehension score of 12% versus 67% for general comprehension and 72% for visual elements. This, in turn, has lowered the overall comprehension rates for the colour variable.

3.3.2 Eyetracking data

3.3.2.1 Time to first fixation (ms)

	Colours	Tags	Placement
Hearing	330	419	319
Hard of hearing	206	489	313
Deaf	243	581	161

The longest times to first fixation were attained by participants when watching subtitles with tags. Colours and placement were noticed faster, especially in the case of hearing-impaired viewers. Hearing people had a longer time to first fixation in the case of colours and placement than hearing-impaired people as they may have looked at the subtitles only when they heard the sound.

3.3.2.2 Mean reading time

	Colours	Tags	Placement
Hearing	30%	36%	52%
Hard of hearing	39%	50%	57%
Deaf	46%	57%	63%

Participants spent more time reading subtitles with speaker-dependent placement than subtitles with tags and colours. This may be due to the fact that such subtitles appeared in different places on the screen and viewers had to move their eyes from the left- to the right-hand side of the screen, travelling further distances, as opposed to subtitles with colours and tags, which were all centred.

Subtitles with colours turned out to be the ones with shortest mean reading time scores. Viewers read subtitles with tags for longer, most probably because the subtitles themselves were longer (tag + subtitle) and different from the standard subtitles that Polish viewers are used to.

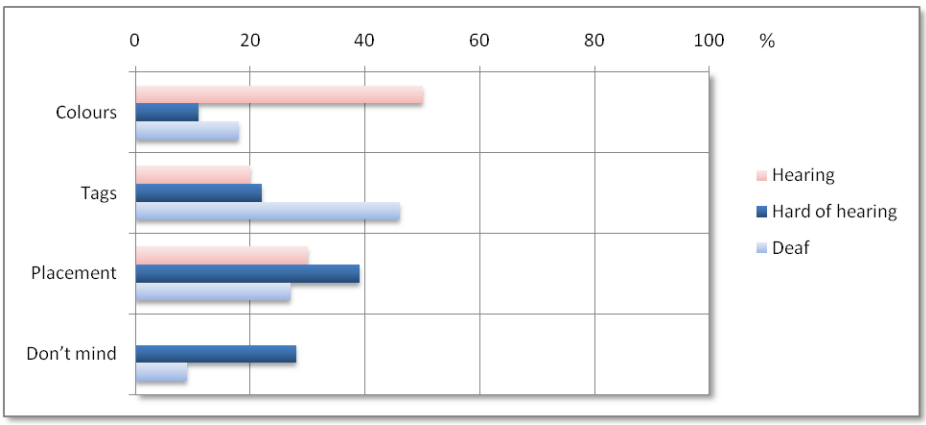
3.3.3 Preferences

In the pre-test preference questionnaire, 46% respondents stated that their preferred method of character identification is the use of colours, followed by a combination of colours and speaker-dependent placement (26%), and speaker-dependent placement only (23%). Only 5% people declared they would like to have name tags to identify characters. Having watched the clips, many participants changed their minds about how they would like characters to be identified.

Perhaps somewhat surprisingly given the results of comprehension questions and pre-test preferences, subtitles with tags earned approval of a significant number of participants, particularly those who are deaf. For them, tags are the preferred option of character identification.

	Colours	Tags	Placement	Don't mind
Hearing	5	2	3	0
Hard of hearing	2	4	7	0
Deaf	2	5	3	1

Post-test preferences: Character ID (total N=39)



Post-test preferences: Character ID (%)

Participants who chose colours as the best method of character identification noted that colours make it easy to identify characters and, as opposed to tags they do not take up space unnecessarily. They criticised tags for taking up too much space and making it difficult to recognize where a tag ends and where a subtitle begins. Those who preferred tags stated that this method enabled them unambiguous identification of the speaker and criticised colours for having to remember the colours allocated to particular characters. This comment is particularly pertinent to the short clip (1 minute 2 seconds) we presented during the tests. Needless to say, colours are much easier to use to identify characters with longer stretches of audiovisual material, such as feature films or TV series.

Many hearing-impaired participants stated they preferred speaker-dependent placement as the most convenient and easy method of character identification. Those who opposed this method claimed it was inconvenient to chase subtitles appearing in different places on the screen. It needs to be pointed out that the clip selected for the test was particularly suited to this identification method owing to the fact that it was mostly composed of shots with two characters standing next to each other, positioned respectively to the left and to the right of the screen. In other types of shots, for instance

with many characters on screen, this method may not provide such clear-cut identification.

2. Subtitle style

Generally speaking, subtitles for hearing-impaired viewers in Poland conform to what we have termed here the 'standard' subtitling style. This means that the editing of subtitles consists mostly in removing what is thought to be unnecessary elements of oral discourse, such as repetitions, hesitations, false starts, etc. Owing to spatial (up to 38 characters) and temporal constraints (12 characters per second[11]), it is unavoidable that subtitles be edited in some cases. It needs to be noted that the subtitling editing policy of the Captioning Department at the Polish public TV has evolved from extensive editing, including high omission and simplification rates as well as longer display times in the early days to less editing and shorter display times at present. This happened as a result of viewers' complaints and more contacts with the hearing-impaired community, who are generally in favour of verbatim subtitling.

The following three variables were tested in this parameter:

1) Verbatim subtitles, which included every single word of the dialogue, even words which usually do not find their way to subtitles, such as repetitions, hesitations and other elements typical of spoken language. These subtitles had shortest display times.

2) Standard subtitles, which included most of the dialogue apart from a few minor repetitions and elements of spoken language which were not crucial to the plot.

3) Edited subtitles, where not only many elements of oral discourse disappeared, but utterances were simplified in terms of vocabulary and syntax. These subtitles had the longest display times to allow for comfortable reading.

3.3.4 Comprehension

Lowest comprehension rates in all respondents groups were obtained with edited subtitles. Verbatim subtitles, in contrast, had the highest comprehension rates.

[11] Such parameters are used in SDH on public TV in Poland.

	Verbatim	Standard	Edited
Hearing	79 %	64%	49 %
Hard of hearing	78 %	73 %	57 %
Deaf	82 %	67 %	63 %

Comprehension: Subtitle style (%)

3.3.5 Eyetracking data

3.3.5.1 Time to first fixation (ms)

	Verbatim	Standard	Edited
Hearing	267	191	241
Hard of hearing	258	264	207
Deaf	178	212	239

Time to first fixation measures differ for each group of participants: verbatim subtitles were noticed fastest by deaf people, standard subtitles by hearing people and edited subtitles by the hard of hearing. Hence, no significant patterns can be discerned.

Hearing people had the longest time to first fixation in verbatim and edited subtitle style clips. Presumably, hearing participants only looked at the subtitle when they heard dialogue, while hearing-impaired viewers were observing the screen with more caution, waiting for subtitles to appear.

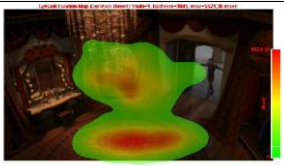
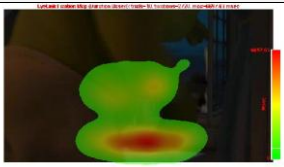
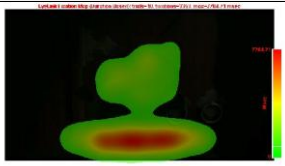
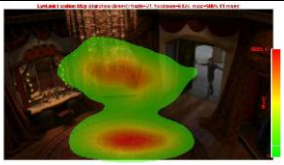
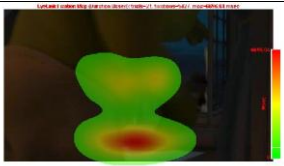
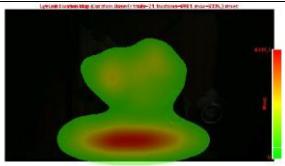
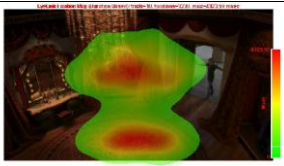
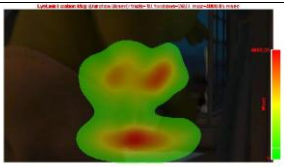
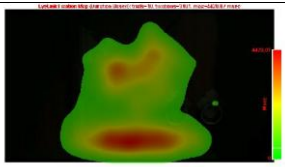
3.3.5.2 Mean reading time

	Verbatim	Standard	Edited
Hearing	45%	40%	35%
Hard of hearing	50%	44%	35%
Deaf	60%	51%	43%

Since verbatim subtitles include all the words from the dialogue, it is only natural that there were more subtitles in the verbatim version (32) than in the other two versions (24 in standard and 25 in edited). Not only were there more subtitles in the verbatim clip, but they were also longer (i.e. they consisted of more characters) and displayed for shorter periods of time. This inevitably contributed to the longer time spent by viewers in the subtitle AOI, thus lengthening the mean reading time in the verbatim version.

Eyetracking data also allowed us to see the reading patterns for the adapted, standard and verbatim styles of subtitling for the three groups of participants. Optimum subtitles, we thought, would allow viewers both to understand the dialogue and to give them time to focus on the image, too. After all, they were watching films which are a dynamic, polysemiotic medium, and not simply reading a motionless, monosemiotic text.

The results of the eyetracking data analysis revealed that people watching the clip with adapted subtitles had spent more time on watching the image than in the case of the clip with standard and verbatim subtitles (see the heatmaps below). In all the three clips, deaf participants spent more time on reading the subtitles than the hard of hearing and hearing participants. This tendency might indicate that deaf people are slower readers, but first and foremost it shows that hearing people have the comfort of not having to rely solely on subtitles to obtain information on the dialogue while for hearing-impaired people reading subtitles is a necessity, and if they stop reading them, they will lose track of the plot.

	Adapted	Standard	Verbatim
Deaf			
Hard of hearing			
Hearing			

Fixation duration-based heatmaps: Subtitling style

In a nutshell, adapted subtitles offer viewers a chance to follow the dialogue and to focus on the action on the screen whereas verbatim subtitles attract a lot of attention to themselves, thus leaving viewers less or no time to watch the image. Adapted subtitling, however, is the option that most hearing-impaired viewers are against, as shown in the pre- and post-test preferences questionnaires.

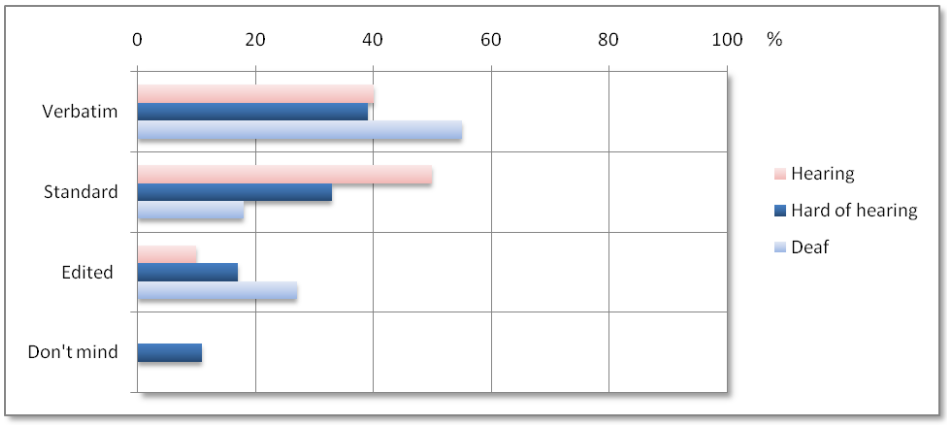
3.3.6 Preferences

It is a well-known fact that if you ask hearing-impaired people, especially those from the Deaf community, what type of subtitling they prefer, they will most likely answer: verbatim. This tendency is also confirmed in our preferences questionnaires.

In the pre-test questionnaire, 63% of respondents stated they want verbatim subtitles, while 38% declared they would prefer to have edited subtitles which are easier to read.

	Verbatim	Standard	Edited
Hearing	4	5	1
Hard of hearing	7	6	3
Deaf	6	2	3

Post-test preferences: Subtitle style (total N=39)



Post-test preference: Subtitle style (%)

The group most preferring verbatim subtitles were deaf participants. When justifying their choice, they said they use verbatim subtitles to learn Polish. Editing subtitles, they claimed, is a form of discrimination against them as they want to have all the information from the dialogue, just as hearing people do. "Why should somebody else decide for me what to include in subtitles and what to omit?" one person asked. Moreover, for those people who lip-read and those who have some residual hearing, verbatim subtitles were the preferred option as the visual and auditory signals they received conformed to the information in the subtitles.

Standard subtitles were the option preferred by hearing people. The reason for this may stem from the fact that they are accustomed to interlingual subtitles in cinemas being abbreviated and they treat reduction and condensation as an intrinsic characteristic of subtitling. Some people stressed it is not necessary for redundant elements of speech to make their way to subtitles. Others said standard subtitles were easier to follow than verbatim subtitles. According to those who chose this option, standard subtitles allow viewers both to read the subtitles and to watch the film, without having to spend too much time on reading. Interestingly, one hearing-impaired participant commented that in the pre-test questionnaire he selected the option "Verbatim", being convinced that it is this type of subtitling he prefers. Having watched all the three clips, however, he said he had changed his mind, stating that in fact it was standard subtitling that he liked most as it was the easiest to read and at the same time allowed him to focus on the film, too.

3. Position on the screen

Subtitles in Poland are usually positioned at the bottom of the screen. This applies both to interlingual subtitles for hearing viewers in cinemas and on DVDs as well as subtitles for the deaf and hard of hearing on television.

1) Bottom – the standard position of subtitles in Poland. Subtitles are sometimes moved to the top in order not to cover an important piece of information placed at the bottom, be it a caption, notice or other crucial element.

2) Mixed – a combination of subtitles containing dialogue placed at the bottom of the screen with a description of sounds placed at the top. Non-existent on the Polish audiovisual market, this style of subtitling was a complete novelty to Polish viewers.

Dialogue	Sound information
	

3) Top – all subtitles, including dialogue and sound description, placed at the top of the screen.

3.3.7 Comprehension

Perhaps somewhat surprisingly, top subtitles had highest comprehension scores. Mixed subtitles, non-existent on the Polish audiovisual scene, proved to be quite confusing to hearing-impaired participants, whose comprehension scores were significantly lower than those of hearing viewers.

	Bottom	Mixed	Top
Hearing	58 %	70 %	76 %
Hard of hearing	60 %	55 %	80 %
Deaf	59 %	41 %	78 %

Comprehension: Position

3.3.8 Eyetracking data

3.3.8.1 Time to first fixation (ms)

	Bottom	Mixed	Top
Hearing	295	342	192
Hard of hearing	262	332	335
Deaf	271	344	274

Bottom subtitles have generally faster reaction times, with the exception of top subtitles for hearing participants. The reason for this pattern seems quite obvious: the bottom position is the one where people expect subtitles to appear and this is where they look in search for subtitles.

3.3.8.2 Mean reading time

	Bottom	Mixed	Top
Hearing	32%	32%	28%
Hard of hearing	34%	33%	27%
Deaf	42%	38%	30%

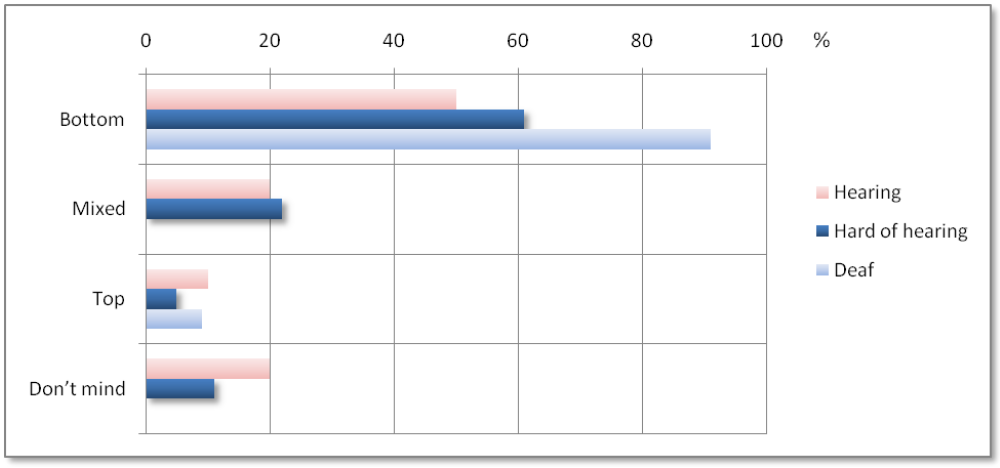
The three variables tested in this parameter do not differ significantly in terms of the time spent on reading subtitles. In each clip, viewers spent about one-third of the total time on reading subtitles. As in the case of other parameters, it was deaf viewers who took longest to read subtitles.

3.3.9 Preferences

In this parameter, comprehension and preference do not go together.

	Bottom	Mixed	Top	Don't mind
Hearing	5	2	1	2
Hard of hearing	11	4	1	2
Deaf	10	0	1	0

Post-test preferences: Position (total N=39)



Post-test preferences: Position (%)

In the pre-test questionnaire, when asked where they would like information about sounds to be placed, 82% people said they preferred the bottom position, while 16% chose the position next to the sound itself. Nobody was willing to see such subtitles at the top of the screen. The reluctance to place sound information at the top of the screen most probably results from habit and Polish standards, where all types of subtitles are placed at the bottom of the screen.

In the post-test preference questionnaire, the overwhelming majority of viewers opted for subtitles placed at the bottom of the screen, justifying their choice by the force of habit. Others stated it was better to have subtitles in one place on the screen exactly where you expect them, instead of having to follow two possible locations plus the image. The few participants who chose mixed subtitles, none of whom were deaf, said that separating dialogue and sound is an interesting idea, adding that if you do not need the information on sounds you can ignore it.

4. Emotions

Subtitles for the deaf and hard of hearing in Poland very rarely include information on emotions. Such information is included when it is difficult to understand an utterance based solely on the image, as is sometimes the case with some ironic remarks. The following variables were tested in this parameter:

1) Description – descriptions of emotions were written in capital letters and placed in brackets in order to distinguish them from dialogue.

2) Emoticons – utterances were preceded by an emoticon denoting mood of particular characters, such as :-(for sadness and :-0) for anger.

3) Nothing – emotions were not described in any way.

3.3.10 Comprehension

The best comprehension results were achieved in the case of subtitles with description of emotions, although many viewers expressed their objections towards any description of emotions on the grounds they are unnecessary (see below).

The clip with emoticons had lower than expected comprehension scores. As viewers did not expect to see subtitles with emoticons it must have taken them some time to process what they saw. It is also possible that emoticons worked as a distracter and hampered comprehension, diverting viewers' attention from the message to the unusual form of subtitles.

	Descriptions	Emoticons	Nothing
Hearing	82 %	79 %	85 %
Hard of hearing	73 %	70 %	60 %
Deaf	78 %	67 %	63 %

Comprehension in the emotions parameter

3.3.11 Eyetracking data

3.3.11.1 Time to first fixation (ms)

	Descriptions	Emoticons	Nothing
Hearing	238	337	205
Hard of hearing	246	322	204
Deaf	260	141	148

Subtitles with no description of emotions were the fastest to spot, which is also in line with the results of eyetracking tests in Spain and in Italy. Subtitles with no description of emotions are the standard option in Poland and this may be the reason for their being the quickest to notice as they contained no potentially distracting elements such as descriptions or emoticons.

The longest time to first fixation was attained by hearing and hard of hearing participants in the clip with emoticons. Deaf viewers, in contrast, noticed the subtitles with emoticons fastest.

3.3.11.2 Mean reading time

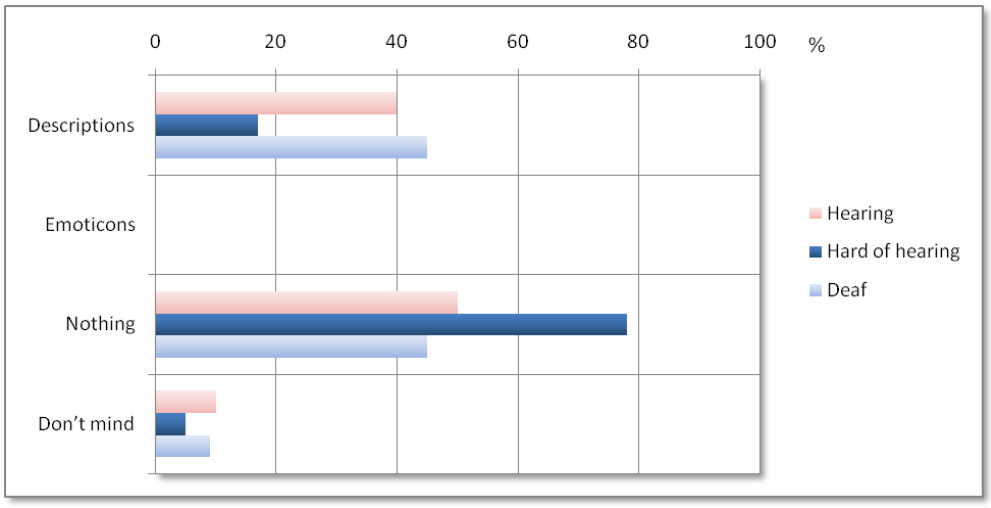
	Descriptions	Emoticons	Nothing
Hearing	40%	33%	42%
Hard of hearing	42%	37%	42%
Deaf	45%	42%	50%

Similarly to the results of the eyetracking tests in Spain, Polish participants spent less time on reading subtitles with emoticons than other types of subtitles. The differences between reading subtitles with and without any description were not significant.

3.3.12 Preferences

	Descriptions	Emoticons	Nothing	Don't mind
Hearing	4	0	5	1
Hard of hearing	3	0	14	1
Deaf	5	0	5	1

Post-test preferences: Emotions (total N=39)



Post-test preferences: Emotions (%)

Most people said they find any description of emotions unnecessary as it can be easily inferred from the image and context. One person noted that subtitles with descriptions of emotions are a simplification, depriving viewers of a chance to arrive at interpretations themselves.

In the pre-test questionnaire, 15% of participants stated they would like to see emoticons as a way to describe emotions in subtitles. The same number of people (15%) declared they preferred verbal description. As many as 69% said there was no need to describe emotions in subtitles. However, during the eyetracking tests, emoticons turned out to be the most severely criticised option in the post-eyetracking tests. One person admitted that he was quite open to the idea of using smileys in subtitles when he first

heard of it, but having seen the clip he changed his mind. On the whole, people opposed the use of emoticons in subtitles on the grounds that it hinders the reading and comprehension process. It is also difficult to interpret them, only a few of the emoticons can be easily understood, while others are not very intuitive. What is more, it is impossible to have an emoticon for every shade of meaning.

Some people suggested using graphic-based smileys [e.g. ☺], not text-based ones [e.g. ;)], as they would be easier to interpret. Others claimed that while smileys are common in Internet communication, they are inadequate for subtitles. Those people who were in favour of including descriptions of emotions in subtitles stated that they emphasise emotions and it is easier to understand the mood of the characters. All in all, however, the vast majority of people were against including any type of information concerning emotions in subtitling.

5. Sounds

Information on sounds in Polish television subtitles for hearing-impaired viewers is placed at the bottom of the screen in white capital letters against blue background:

On DVDs, such information is usually capitalised, but no colours are used. The variables tested in this parameter were:

1) Description – a word or phrase explaining the sound that could be heard, for example PURRING or SHOUTING.

2) Icons – a picture denoting a barking dog.

3) Nothing – sound information was not described in subtitles at all.

3.3.13 Comprehension

The best comprehension results were achieved in the clip which did not contain any description of sounds whatsoever. The worst results, quite surprisingly, were obtained in the case of verbal description of sounds: the option preferred by the vast majority of participants (see below).

	Description	Icon	Nothing
Hearing	79 %	85 %	85 %
Hard of hearing	45 %	73 %	82 %
Deaf	48 %	74 %	82 %

Comprehension: Sounds

3.3.14 Eyetracking data

3.3.14.1 Time to first fixation (ms)

	Description	Icon	Nothing
Hearing	294	988	219
Hard of hearing	271	1120	192
Deaf	338	973	234

The fastest to be noticed were subtitles containing no description of sounds.

The data on time to first fixation in the icon variable above relates not to a subtitle, but to the time which took the participants to notice the dog icon. As can be seen from the table above, it took people significantly longer to notice the icon. What is more, as we learned from post-test interviews, some participants did not notice the icon at all. This most probably stems from the fact that viewers are unaccustomed to seeing non-diegetic icons denoting sound in films. Thus, Polish viewers did not expect any icons in the top right-hand corner and were not looking there.

3.3.14.2 Mean reading time

	Description	Icon	Nothing
Hearing	28%	37%	25%
Hard of hearing	35%	40%	30%
Deaf	40%	45%	30%

In the clip with the dog icon, mean reading time was calculated based on two AOI: subtitles and the icon, whose total value was compared to the duration of the entire clip. Thus, it turns out that reading the subtitles and looking at the dog icon took viewers longer than simply reading subtitles at the bottom of the screen.

Since the clip with description included additional subtitles describing sounds, the total number of subtitles was higher than in the clip with no description (24 versus 18 respectively). This, in turn, resulted in longer times spent in the subtitle area and shorter times in the image. However, in certain cases the sound information needs to be prioritised over the image, so that hearing-impaired viewers can follow the story.

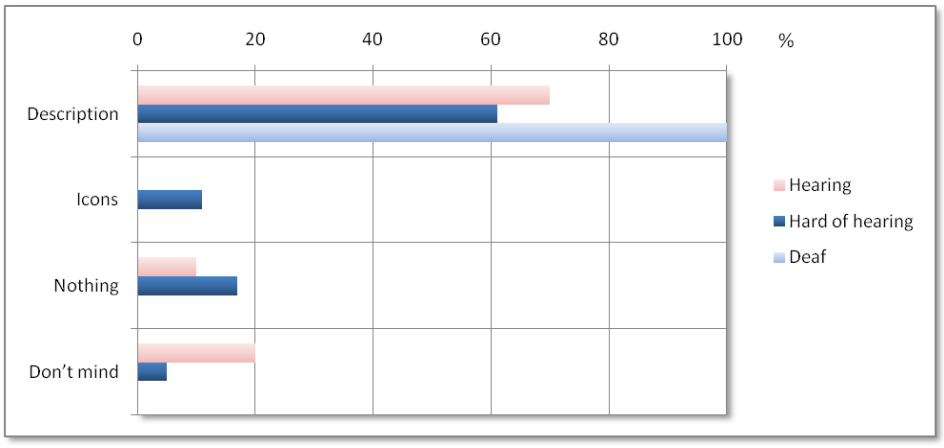
3.3.15 Preferences

In the pre-test questionnaire, 46% participants were in favour of including information on where the sound comes from, 11% were for using onomatopoeic expressions, 23% wanted to have a description of what the sound is like and 26% said that they would like to see sounds described as pictures. This tendency was not confirmed later on, after the eyetracking tests were carried out.

Before testing this parameter, we told the participants they would watch three clips with different methods of presenting sound information, without going into any details. As a result, some participants admitted that they had not noticed the icon with the barking dog. Those who did notice the dog found it rather distracting, while some participants took it to be part of the film.

	Description	Icons	Nothing	Don't mind
Hearing	7	0	1	2
Hard of hearing	11	2	3	2
Deaf	11	0	0	0

Post-test preferences: Sounds (total N=39)



Post-test preferences: Sounds (%)

One person had an interesting suggestion: icons, he said, could be used in a more dynamic way: instead of a motionless icon of a dog, the picture could be moving in the way dogs tend to move when they bark because sound is essentially a kind of movement.

All in all, the vast majority of people taking part in the study, including 100% of deaf participants, were in favour of including information on sounds in SDH in the form of verbal description, thus choosing what they were accustomed to.

6. Justification

Subtitles for hearing-impaired viewers on Polish television are usually placed in the centre of the screen, with the exception of news programmes (in which case they are always left-aligned) and fiction programmes where speaker-dependent placement is used (this method is only used in some films and TV series). In cinemas and on DVD, subtitles (not only SDH) are placed at the bottom of the screen in the centre. In this parameter, two variables were tested:

1) Left-aligned subtitles – moved to the left part of the screen at the bottom

2) Centred subtitles – placed in the centre of the screen at the bottom

3.3.16 Comprehension

Viewers watching clips with left-aligned subtitles achieved better comprehension results than when watching clips where subtitles were placed in their usual central position.

These results do not conform to the ones attained by respondents in Italy and Spain, so further research is necessary to provide conclusive results.

	Centre	Left
Hearing	67 %	88%
Hard of hearing	60 %	77 %
Deaf	56 %	63 %

Comprehension: Alignment

3.3.17 Eyetracking data

3.3.17.1 Time to first fixation (ms)

	Centre	Left
Hearing	165	263
Hard of hearing	252	215
Deaf	150	236

Centred subtitles were spotted faster by deaf and hearing participants, while hard of hearing viewers reacted faster to left-aligned subtitles. The fact that centred subtitles were generally the first to be noticed probably stems from the fact that they appeared in the place where users were expecting them, since the central position is the standard one.

3.3.17.2 Mean reading time

	Centre	Left
Hearing	41%	35%
Hard of hearing	44%	37%
Deaf	51%	48%

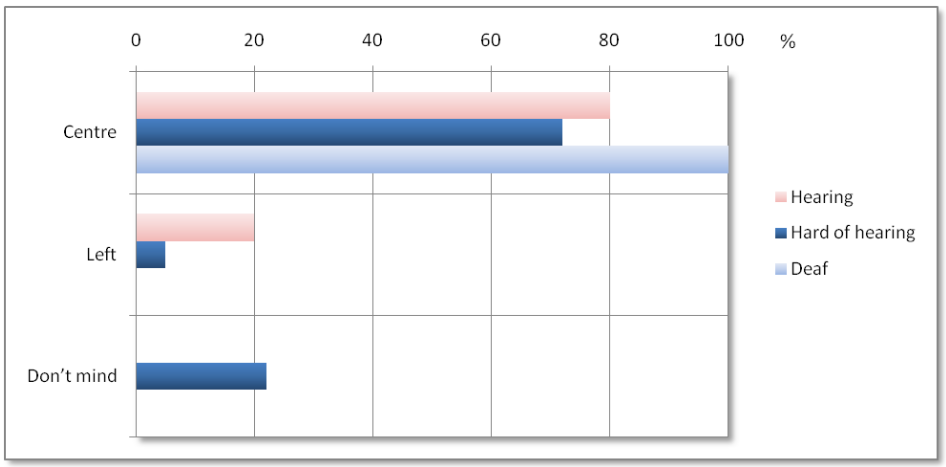
Although quicker to spot, centred subtitles are read slightly longer than left-aligned subtitles. The differences are not significant, hence they may be put down to the length, duration and the number of subtitles in particular clips. For instance, in the clip with left-aligned subtitles there were 15 subtitles, whereas in the clip with centred subtitles there were as many as 22 subtitles. This calls for further research, where the length, duration and the number of subtitles would be more comparable.

3.3.18 Preferences

In this parameter, comprehension and mean reading time results cannot be considered a good indication of viewers' preferences. Judging by the better results obtained with left-aligned subtitles, one would expect viewers to choose this variable as their preferred option. Nevertheless, viewers prefer to tread a well-known path and chose what they are accustomed to.

	Centre	Left	Don't mind
Hearing	8	2	0
Hard of hearing	13	1	4
Deaf	11	0	0

Post-test preference: Alignment (total N=39)



Post-test preference: Alignment (%)

The bottom position is where most viewers expect subtitles to appear. Not surprisingly, therefore, the overwhelming majority of participants (incl. 100% of deaf participants) declared they prefer subtitles to be placed in the centre of the screen, instead of being left-aligned. Many participants stressed it was due to habit that they prefer the centre option. Others pointed out that subtitles in the centre of the screen allow for more comfortable viewing as you can take in the entire screen, which is impossible when watching left-aligned subtitles.

7. Borders

Television subtitles for hearing-impaired viewers in Poland are displayed using analogue teletext, which means they have a black background and no borders.
In this parameter, we have tested the following two variables:

(1) Borders – each subtitle is surrounded by a black border. Using the border is meant to improve legibility, especially in scenes with a white or light background.

(2) No borders – white subtitles are displayed without any borders

3.3.19 Comprehension

Significant differences could be observed in terms of overall comprehension between the clip containing borders and the one without borders, see the table below. Viewers watching the clip without borders had better comprehension results than when watching the clip with borders.

This is visible particularly in the case of deaf viewers, whose comprehension of the textual elements was only 22% for the clip with borders (compared to 54% for hearing and 55% for hard of hearing viewers) and 66% for the clip without borders (compared to 90% for hearing and 75% for hard of hearing viewers).

	Border	**No border**
Hearing	61 %	91 %
Hard of hearing	62 %	70 %
Deaf	44 %	78 %

Comprehension in borders parameter

It seems that further research is necessary in order to exclude the possibility that the results were due to the variable difficulty of particular questions.

3.3.20 Eyetracking data

3.3.20.1 Time to first fixation (ms)

	Border	**No border**
Hearing	346	231
Hard of hearing	377	251
Deaf	347	284

Subtitles without borders were faster to spot than subtitles with borders which is contrary to our expectations, as it is subtitles with borders that are generally considered to have better visibility. This, however, does not necessarily translate into faster reaction times.

3.3.20.2 Mean reading time

	Border	No border
Hearing	42%	45%
Hard of hearing	48%	46%
Deaf	50%	50%

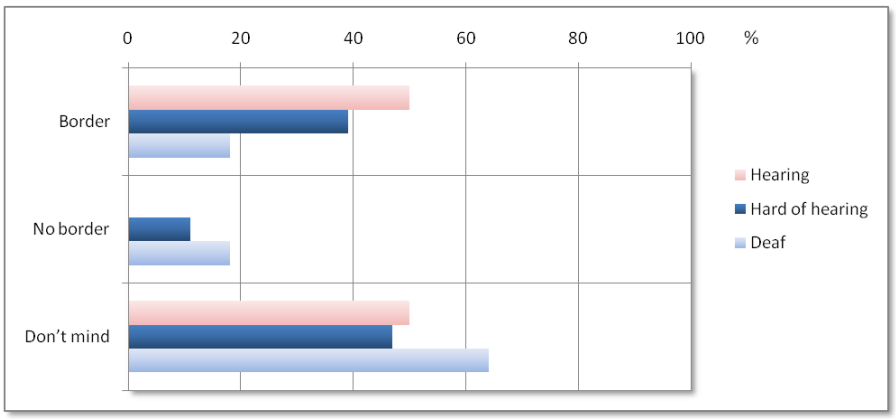
In terms of mean reading time, there seem to be no significant differences between subtitles with and without borders. This is in line with post-test preferences expressed by viewers most of whom failed to notice any differences between the two clips.

3.3.21 Preferences

Despite a marked difference in comprehension results, viewers did not reveal a marked preference for subtitles without borders.

	Border	No border	Don't mind
Hearing	5	0	5
Hard of hearing	7	2	9
Deaf	2	2	7

Post-test preference: Borders (total N=39)



Post-test preference: Borders (%)

Most people watching the two clips failed to notice the presence or absence of the borders. Hence, most of them said it did not make any difference to them. Some participants opted to have borders as a permanent option, stating that in this case subtitles have better legibility and that are clearly visible on any type of background.

8. Box

As stated above, SDH on Polish television are broadcast via analogue teletext, which means that all television subtitles appear on the screen against black background (with the exception of sound information, in which case the background is blue). On DVDs and in cinemas, however, this convention is not used.

The following two variables were tested in this parameter:

1) Black box

2) No black box

3.3.22 Comprehension

All groups of participants achieved better overall comprehension results after watching clips with the subtitles in a black box.

	Box	**No box**
Hearing	79 %	67 %
Hard of hearing	73 %	55 %
Deaf	63 %	52 %

Comprehension in the box parameter

When looking at the comprehension results for textual elements only there is no difference in the comprehension of the clip with the box and without the box in the case of deaf viewers, who attained a 66% comprehension rate in both clips. Significant differences were observed, however, in the case of hard of hearing viewers, who achieved textual comprehension rates of 55% for the box variable and 85% for the no box variable.

3.3.23 Eyetracking data

3.3.23.1Time to first fixation (ms)

As both clips begin with a one-line subtitle describing music, the subtitles examined here are identical in terms of length and content: MUZYKA ('music'). This makes them perfect for comparison (see below).

Box	No box
	

Eyetracking data reveals that hearing and hard of hearing participants noticed subtitles without the box faster, while deaf participants spotted subtitles in the black box faster.

	Box	No box
Hearing	267	209
Hard of hearing	373	275
Deaf	249	304

3.3.22.2 Mean reading time

	Box	No box
Hearing	42%	48%
Hard of hearing	39%	50%
Deaf	48%	58%

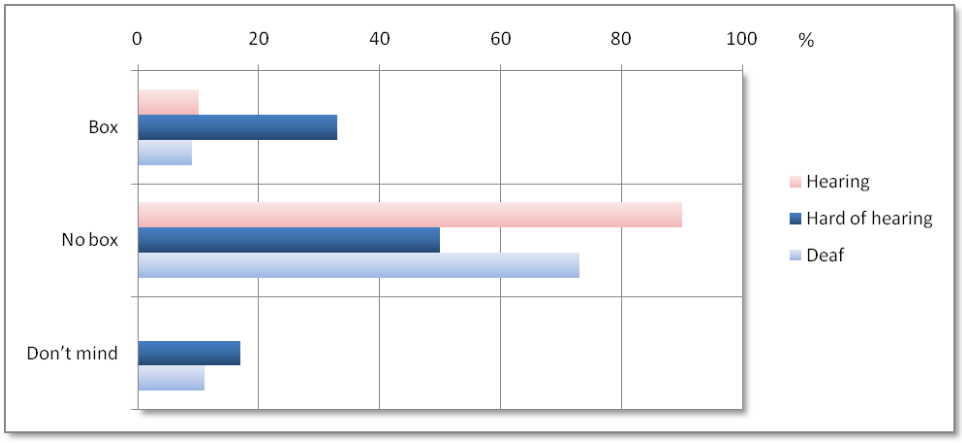
In our tests, people spent more time reading subtitles without the box than subtitles in the box. Again, the reason for this may be the fact that the number of subtitles in the clip with subtitles in the box was 20 and the number of subtitles in the clip without the box was 27, which must have increased mean reading times. The difference between the subtitles in the box versus subtitles without the box was most marked in the case of hearing-impaired viewers.

3.3.23 Preferences

Contrary to the comprehension results, most viewers were not in favour of the black background.

	Box	No box	Don't mind
Hearing	1	9	0
Hard of hearing	6	9	3
Deaf	1	8	2

Post-test preference: Box (total N=39)



Post-test preference: Box (%)

Most participants were opposed to placing subtitles on a black background, noting that it unnecessarily covers the image and makes it more difficult to read subtitles. Those who were in favour of the black background stated that in this case, the subtitles have better legibility and visibility.

9. Shadows

Shadows are not used in subtitles for the deaf and the hard of hearing on Polish public television. Similarly to borders, shadows are meant to improve subtitle legibility when no black background is used. In this parameter, we have tested the following two variables:

1) Shadows

2) No shadows

3.3.24 Comprehension

When watching subtitles with shadows, all respondent groups achieved better overall comprehension results.

	Shadows	No shadows
Hearing	73 %	70 %
Hard of hearing	78 %	68 %
Deaf	89 %	52 %

Overall comprehension in shadows parameter

The analysis of the textual elements of comprehension only shows that subtitles with shadows yielded much better comprehension results: 81% for hearing, 90% for hard of hearing and 100% for deaf viewers than subtitles without shadows: 54% for hearing, 55% for hard of hearing and 56% for deaf viewers.

3.3.25 Eyetracking data

3.3.25.1 Time to first fixation (ms)

	Shadows	No shadows
Hearing	355	261
Hard of hearing	333	277
Deaf	302	256

Participants from all groups needed more time to notice subtitles with shadows than subtitles without shadows, which is not confirmed by the DTV4All Italian tests, where the opposite tendency was observed. This calls for further research to be carried out on subtitle perception.

3.3.25.2 Mean reading time

	Shadows	No shadows
Hearing	63%	55%
Hard of hearing	55%	47%
Deaf	62%	56%

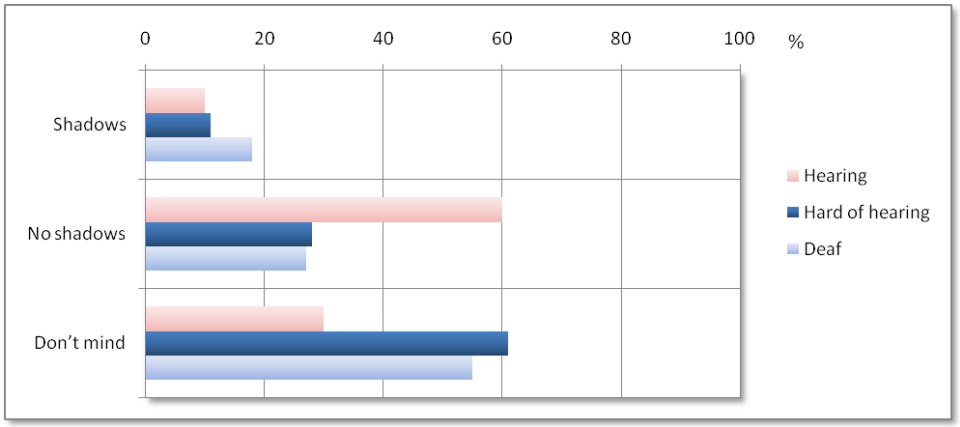
Although it took longer for all groups of participants to read subtitles with shadows, the corresponding comprehension results confirm that it was worth the effort.

3.3.26 Preferences

As in the case of borders, shadows were an option that was hardly noticed by the participants.

	Shadows	No shadows	Don’t mind
Hearing	1	6	3
Hard of hearing	2	5	11
Deaf	2	3	6

Post-test preference: Shadows (total N=39)



Post-test preferences: Shadows (%)

Most people stated that shadows did not matter to them since they did not even notice them. Others claimed it depended on the film, as shadows may be helpful in certain scenes and contexts.

Conclusions

The current standards used in SDH on Polish television are:

Character ID	Subtitling style	Position	Emotions	Sounds	Justification	Borders	Boxes	Shadows
Colours + displacement for fiction	Standard/ edited	Bottom	No description	description	Centre for fiction, left-aligned for news	No borders	Black box	No shadows

Overall, the standards currently in use largely overlap with the preferred options in a number of parameters, with the exception of subtitling style and border/box/shadow.

In some cases, such as borders, boxes and shadows it is difficult to draw definitive conclusions as information from eyetracking data, comprehension scores and preference tests is contradictory. For example, viewers tend to prefer options other than those which gave the highest comprehension scores or the best eyetracking results.

For **character identification**, it seems that colours and speaker-dependent placement can be combined. In certain shots, for instance with off-screen voices coming from characters for whom no colours were allocated, this method can be supplemented with name tags.

Subtitling style is the most controversial parameter in the study. While most people prefer verbatim subtitles, comprehension and eyetracking tests do not confirm the usefulness of this subtitling style. Therefore, perhaps the best option would be to use standard subtitling, i.e. restrict subtitle editing to the minimum and use it only with fast-paced dialogue, especially in close-ups.

While top subtitles had the best comprehension scores, they were not warmly received by the participants, who stated they prefer bottom subtitles. Since eyetracking tests did

not show significant differences between top, mixed and bottom subtitles, it is advisable that bottom subtitles be selected as the most preferred **position**.

With lowest comprehension scores and longest time to first fixation, subtitles with emoticons, fiercely criticised by participants, are the least likely candidates to be used to describe **emotions** in films. The winning variable in this parameter seems to be no description at all.

When it comes to ways of describing **sounds** in SDH, the idea of using icons did not gain much approval from the people taking part in the study. Although the best comprehension and eyetracking scores were achieved by the clip with no description of sound whatsoever, it seems that nevertheless such verbal description needs to be added as it is often necessary for the understanding of a clip and it is also the option most preferred by viewers.

In terms of **justification**, although left-aligned subtitles had better comprehension scores, the preference test results and time to first fixation scores point to centred subtitles as the best option in this parameter.

Based on the data we have researched in the **borders** parameter, subtitles without borders drew better results when it comes to comprehension and eyetracking tests. However, it seems that some kind of visual marking needs to be added to white subtitles, especially appearing on a light background, as otherwise they could be illegible.

As for viewers' preferences regarding the **box** option, it is subtitles without box that are preferred by a vast majority of people. This is also supported by results of comprehension questions, while eyetracking tests do not provide any conclusive results here.

As far as **shadows** are concerned, despite the fact that they took slightly longer to be noticed and had slightly longer mean reading time, subtitles with shadows had significantly better comprehension scores.

4 Eye-Tracking Test Report: Germany

4.1 Test Background

The eye-tracking test in Germany analysed the parameters 'Emoticons', 'Icons' and 'Speed/Subtitle Type'. The other parameters could not be considered, as the tests were realised as part of a diploma thesis. Apart from that, the study followed the test design established by the DTV4All project leaders.

The test was conducted at the Humboldt University of Berlin. In total, 21 subjects took part; that is 7 hearing, 7 hard of hearing (from here onwards: HoH) and 7 deaf participants. The hearing participants were aged between 25 and 34, the HoH between 28 and 52, and the deaf participants 25 to 41. Two of the tests failed and the data had to be rejected; i.e. one in the group of the deaf and one in the group of the hearing participants. In some cases individual trials (video clips) failed, which were then excluded from analysis, too. However, the minimum number of 5 participants in each group was achieved for all trials, for some of the trials data were obtained from 6 or 7 subjects. The data gathered from the questionnaires before and during the test were taken from all participants, also in those cases where the eye-tracking tests failed.

The deaf participants were highly interested in the subject and motivated to take part, certainly not least because of the poor subtitling rate in Germany of roughly 10 percent only (as at April 2009, source: Sign Dialog). TV stations are not obliged to offer subtitles therefore subtitling in Germany is a matter of self-commitment, which especially at private stations has hardly been complied with. The Deaf community has been struggling for years to increase the subtitling rate. In addition, there are no national guidelines for subtitles binding for all stations. Efforts are being made currently to establish ones (cf. www.sign-dialog.de). All in all, the situation of subtitling in Germany makes it a matter of high interest for those who rely on subtitles.

Finding HoH-participants was, however, more difficult, as especially the younger ones are not that well organised in organisations and represented in the Internet as are the Deaf, and therefore harder to get in touch with. Depending on their residual hearing, they can also more or less follow a programme that is not subtitled. Therefore, the subject is certainly not as important for them as it is for the Deaf. Some of those who finally took part in the test did not seem to be very interested and so, for example, did not take too much care in filling in the questionnaires.

The participants would first fill in the pre-test preference questionnaire, which would usually take 10 minutes at the maximum, 30 minutes in one exceptional case for a deaf participant. Then participants would start the eye-tracking test. As mentioned before, three parameters were tested with three variables each. Thus nine video clips were shown, which would take about 20 minutes, filling in the comprehension questionnaires included. Taking into account that the Eye-tracker used in the tests was an EyeLink I, a head-mounted system that requires participants to wear a headband with cameras registering their eye movements, this was a quite reasonable amount of time. Otherwise, wearing the headband for a prolonged time might cause headaches. As in this short amount of time tiring effects were kept to a minimum, the order of the videos was not randomised, except the last three videos for the speed parameter. Following the model in England, the same clip was shown with different subtitles, which required randomisation.

4.2 Comprehension

The comprehension data was obtained from the six videos of the *Emoticons'* and *Icons'* parameter. The overall comprehension rate reached by Hearers was 1.63 out of 2; the comprehension rates of the HoH and the Deaf were nearly the same with 1.39 (HoH) and 1.4 (Deaf). Looking more in detail reveals that the Hearers' comprehension is about the same for all comprehension instances, the visual comprehension (referred to as 'Image' in the graphs) being slightly minor to the other comprehension instances. Compared to the other two groups, Hearers achieved the best results for all comprehension instances, with sense making the most pronounced difference, 1.67 in the case of Hearers as against 1.48 (HoH) and 1.24 (Deaf).

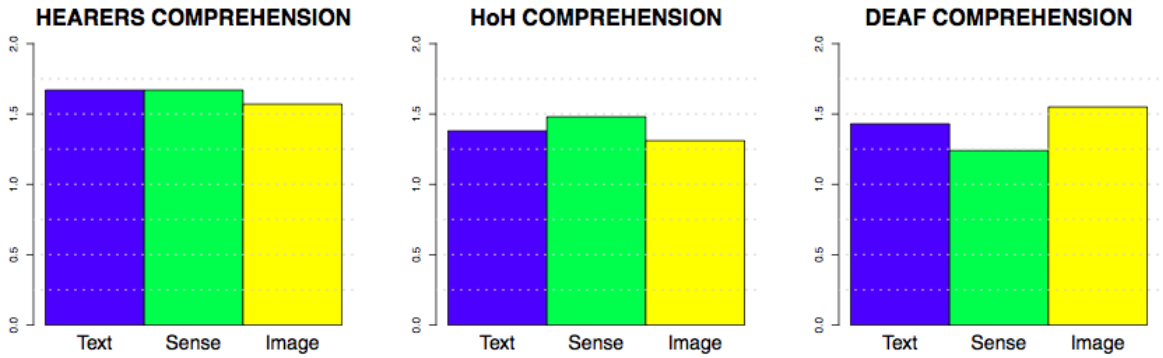


Figure 5.1: Comprehension instances across trials

The HoH participants obtained the lowest rates in both textual and visual comprehension (1.38 and 1.31); their general comprehension (referred to as 'Sense' in the graphs) was however better (1.48). Deaf participants, as expected, achieved the best results in visual understanding (1.55), thus being nearly at the same level as the hearing participants (1.57) and the lowest results in general understanding (1.24). Their textual understanding was better than that of the HoH, 1.48 (Deaf) as against 1.38 (HoH).

4.3 Information Processing

As mentioned before, there is no national standard concerning subtitling for the Deaf and Hard of Hearing in Germany. The main public TV broadcasting association in Germany, ARD, has set some standards, which are applied by most of its regional members. They are, however, quite general and not binding to other TV stations. Thus, the following information about German SDH refers to common practices rather than national standards, the latter being non-existent so far.

1. Emoticons

Description is the common way in German SDH to render the manner of speaking, where necessary. Emoticons are not used. To provide no information on emotions at all is not common either for SDH, and programmes with interlingual subtitles are quasi non-existent, as Germany is a dubbing country. Deaf viewers could, however, be used to interlingual subtitles from watching DVDs or subtitled films in cinema. The preferred option of the deaf participants was *description* (4 entries in the post-test questionnaire), *emoticons* were chosen by only one. The Hard of Hearing tended to prefer no additional information at all (3 entries); if provided, prefer description (2) to emoticons (1). The

hearing participants, however, liked the emoticons: three of them chose emoticons as against two each for description or no additional information.

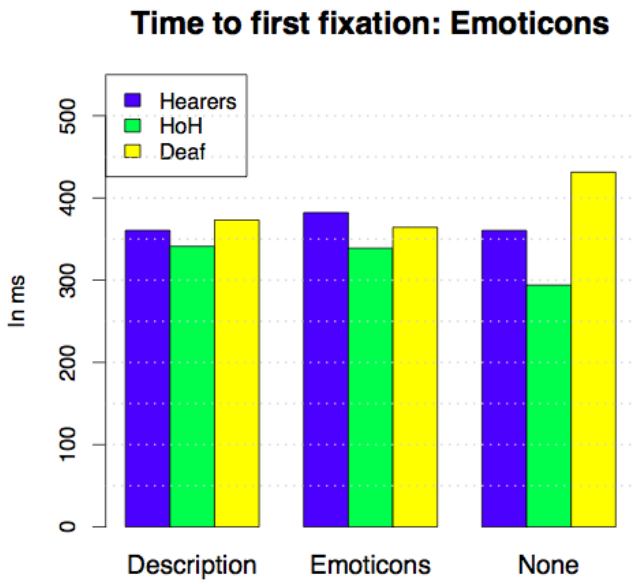


Figure 5.2: Converying emotion

The time elapsed from the previous fixation to the first fixation on the subtitle (from here onwards: reaction time) for *Description* and *Emoticons* is nearly the same, as Figure 5.2 shows. Only the reaction times for *Emoticon: None* differ significantly for the Deaf and HoH, although in an opposite direction: whereas the HoH are quicker to look at the subtitles, the Deaf take longer.

Table 1: Emoticons (n=number of participants from whom data could be obtained M=mean, SD=standard deviation)

		Emoticons:Description			Emoticons:Emoticons			Emoticons:None		
		N	M	SD	n	M	SD	n	M	SD
Time to first fixation (in ms)	Hearers	6	360.6	175.0	6	382.2	221.2	6	360.5	202.9
	HoH	6	341.2	193.0	6	338.9	179.8	7	294.0	133.6
	Deaf	5	373.0	201.0	6	364.3	202.0	6	431.3	234.3
First fixation length (in ms)	Hearers	6	200.6	73.9	6	228.3	121.8	6	196.1	87.4
	HoH	6	195.2	83.6	6	195.2	62.5	7	167.4	53.3
	Deaf	5	205.3	83.3	6	223.6	129.4	6	203.7	93.0
Mean reading time (in percent)	Hearers	6	45.9	22.7	6	43.9	23.0	6	52.7	23.3
	HoH	6	58.8	26.7	6	48.4	26.2	7	55.5	21.6
	Deaf	5	65.5	24.3	6	60.7	24.5	6	61.7	21.1

When it comes to first fixation length, there is hardly any difference between the three clips at all (cf. Fig. 5.3). The mean reading times (i.e. time spent on the subtitle in percent) however show distinguishable differences (cf. Fig. 5.4). In each group, the mean reading times for *Emoticons* were the shortest. Deaf and HoH viewers spent most time reading *Description* subtitles, whereas the hearing viewers spent most time on the

subtitles providing no further information at all. Emoticons were "seen", in the sense of directly fixated upon in 18% (Hearers), 20% (HoH), and 23% (Deaf) of instances.

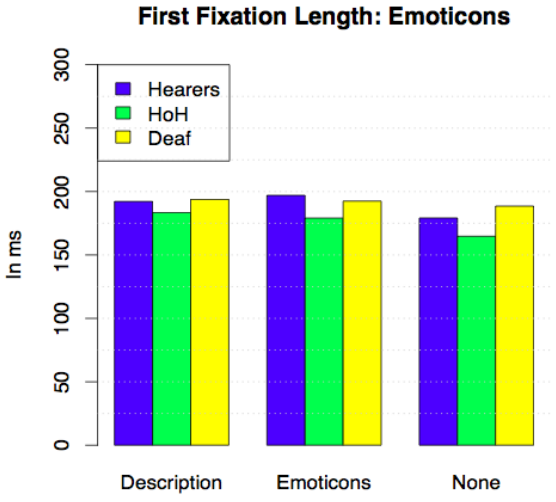


Figure 5.3

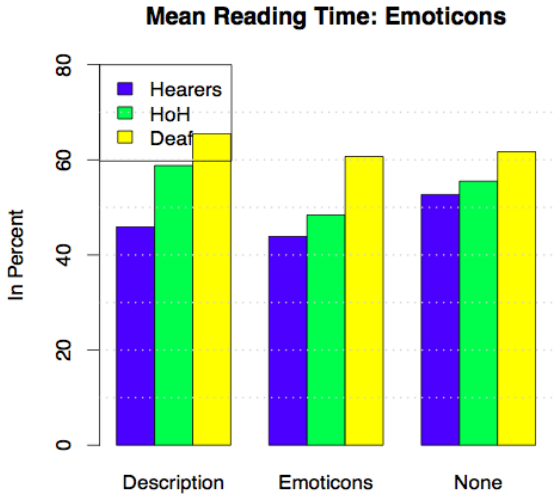


Figure 5.4

Looking at the comprehension instances, taking test, sense and image together, it is remarkable that *Description* yielded the poorest understanding for all groups (cf. Figure 5.5). Both Hearers and the Hard of Hearing reached highest comprehension rates with *Emoticons: None*, whereas the Deaf seem to have understood best the *Emoticons'* video.

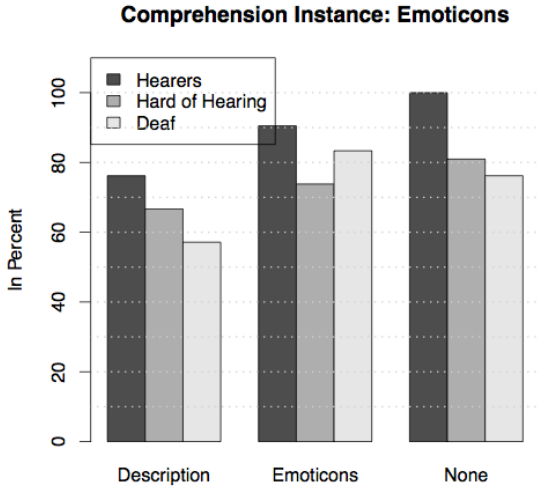


Figure 5.5

Looking at the comprehension instances individually in Figures 5.6 to 5.8 reveals the following: As to textual comprehension, the video with descriptions was least understood by all groups. Whereas hearers got best results with *Emoticons: None*, the Deaf and Hard of Hearing would achieve the same results both with *Emoticons* and *None* (cf. Figure 5.6).

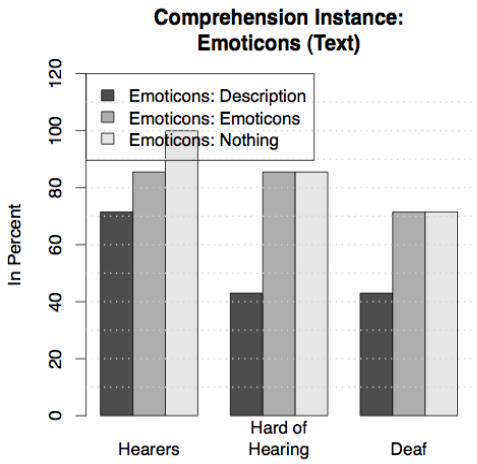


Figure 5.6

As to general comprehension (sense), *Emoticons: Description* returned the poorest comprehension rates for the Deaf and Hearing, too (cf. Figure 5.7). The Hard of Hearing, however, could grasp most the general sense best in this case. The Deaf got distinctly better general comprehension results with *Emoticons*.

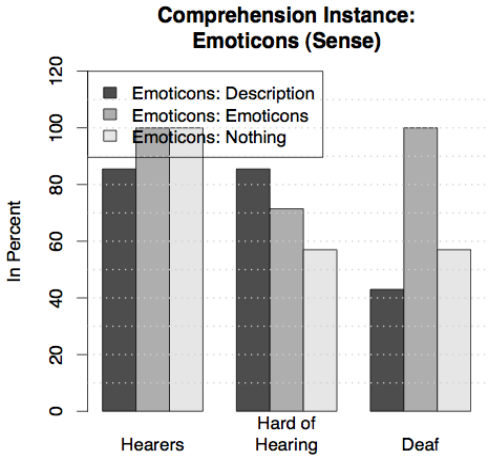


Figure 5.7

When it comes to visual comprehension, Figure 5.8 shows that the video without additional mood information ranks in first place with 100 % comprehension in all three groups. Both the Deaf and HoH got lowest results with *Emoticons: Emoticons* and slightly better results with *Emoticons: Description,* though the difference is negligible. Hearers, on the contrary, achieved lowest results with *Description* and better results with *Emoticons.*

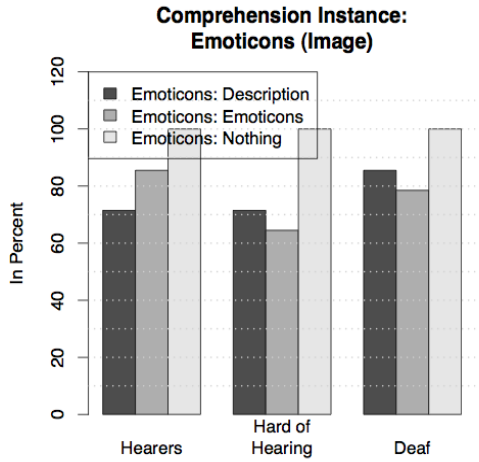


Figure 5.8

The following graphs compare the comprehension rates with the data obtained from the eye-tracker. In order to do so, the mean reading times in percent spent on the particular subtitle critical for answering the comprehension question were considered. As to textual comprehension, the figures below show that in the case of *Emoticons: None* and *Emoticons: Description* lower comprehension rates are accompanied with longer mean reading time (cf. Figures 5.9 and 5.10). This indeed suggests some difficulties in processing the information. However, the *Emoticons: Emoticons'* video does not show any such tendency (cf. Figure 5.11).

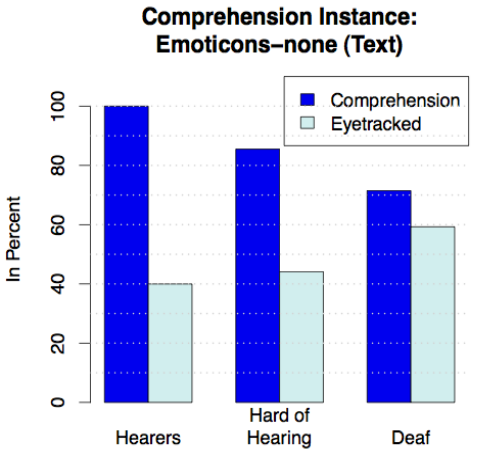


Figure 5.9

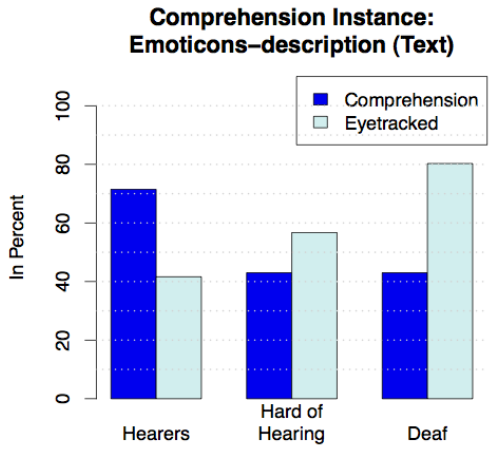


Figure 5.10

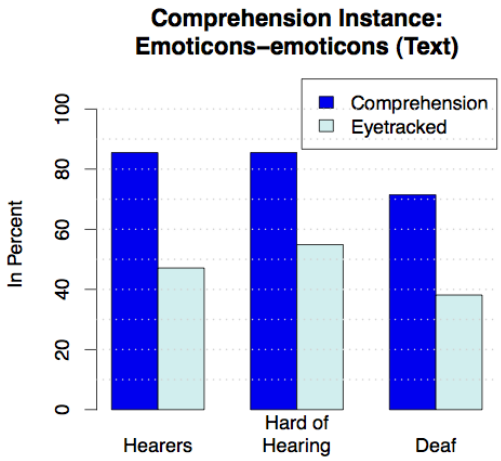


Figure 5.11

Regarding the general comprehension data (Figures 5.12-5.14), the only case where increased reading time and lower comprehension rates indicate any correlation is the video with *Emoticons* (cf. Figure 5.13). The other two don't show any obvious pattern, as it is not clear why the Deaf and HoH would get such different comprehension results with nearly the same amount of time spent on the critical subtitles in the *Emoticons: Description'* video (cf. Figure 5.12), nor why hearers' comprehension for *Emoticons: None* is so much higher than for the Deaf and HoH with mean reading times not considerably different from the other two groups (cf. Figure 5.14).

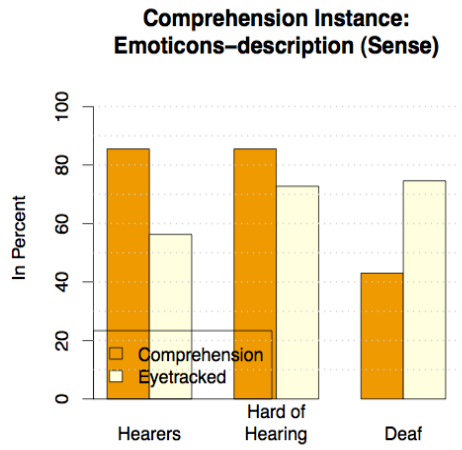


Figure 5.12

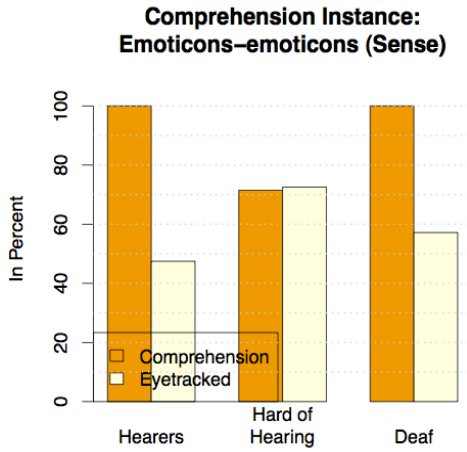


Figure 5.13

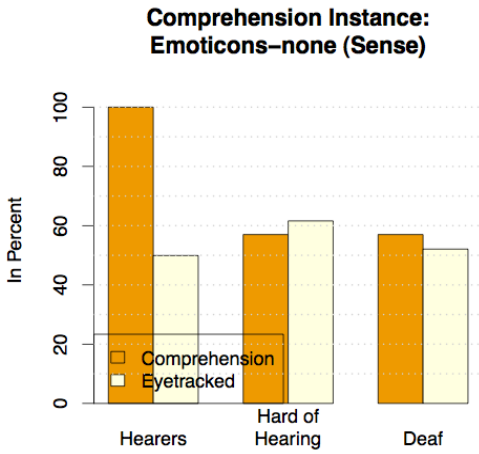


Figure 5.14

For the eye-tracking data of the following graphs on visual comprehension (Figures 5.15-5.17), total fixation time on the critical objects/regions in percent within the time the object was visible on screen was counted.[12] The graphs show that in the case of *Emoticons: Description* the Deafs' comprehension was higher as was the time they dwelled on the critical object (cf. Figure 5.15). In the case of *Emoticons: Emoticons*, however, they don't seem to indicate any correlation whatsoever between comprehension and eye-tracked data (cf. Figure 5.16). The comprehension rates for *Emoticons: None* are the same in all three groups, and the mean reading times don't show any noteworthy differences either (cf. Figure 5.17).

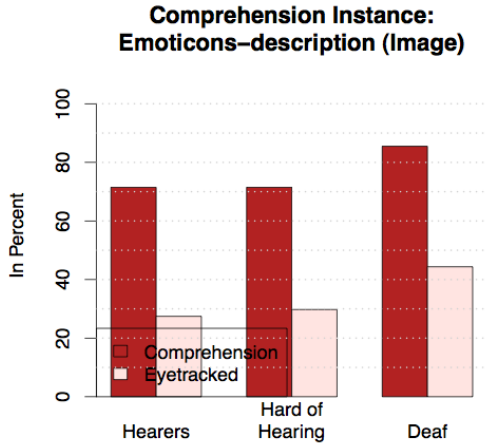


Figure 5.15

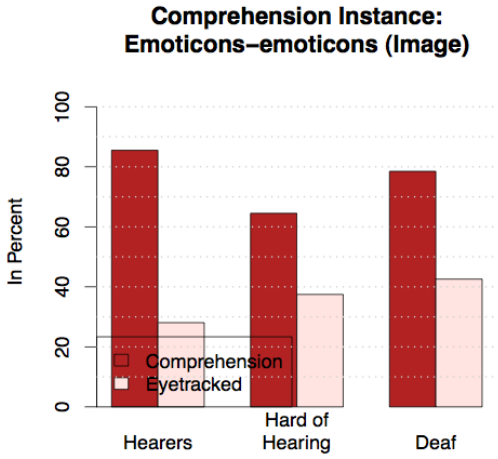


Figure 5.16

[12] The following interest areas were used: an elliptic interest area (in the following called IA) with the coordinates 379,213 (top-left) and 677,555 (right-bottom) for *Emoticons: Description*; two rectangular IAs with the coordinates 150,320 and 511,493 respectively 575,333 and 738,531 for *Emoticons: Emoticons*, and in the case of *Emoticons: None* a freehand IA.

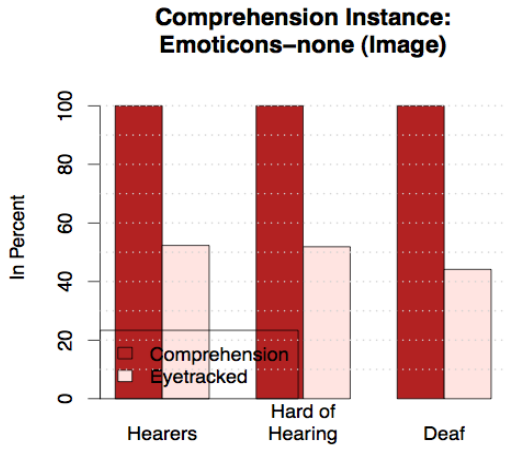


Figure 5.17

2. Icons

As with manner of speaking, sound information is rendered by description in German SDH. Icons have not been used at all so far. Subtitles without sound information might be familiar to deaf users from DVD or cinema. Here again, description was also the clearly preferred option of the test participants, chosen by five Hearers, five Deaf and three Hard of Hearing. Two of the deaf participants chose a combination of description and icons.

In all three groups, the reaction time is longest for *Description* (cf. Figure 5.18). Deaf viewers had the shortest reaction times with *Icons: None*, whereas Hearers and the Hard of Hearing reacted quicker with *Icons: Icons*.

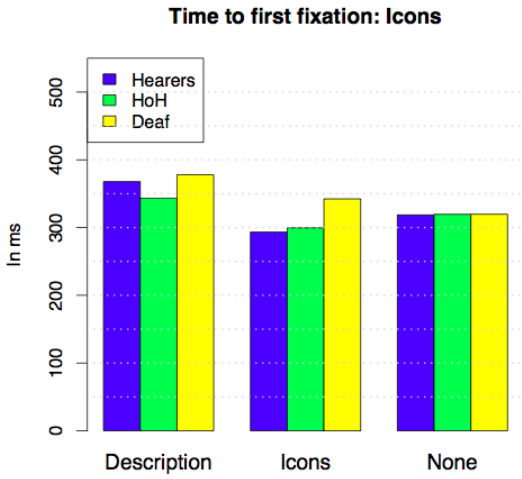


Figure 5.18

Table 5.2: Icons

		N	M	SD	n	M	SD	n	M	SD
Time to first fixation (in ms)	Hearers	6	367.8	194.6	6	293.3	148.9	6	318.6	181.1
	HoH	6	343.4	191.6	6	299.6	147.8	6	319.6	173.8
	Deaf	6	377.9	198.0	7	342.3	205.6	6	319.7	178.8
First fixation length (in ms)	Hearers	6	223.7	142.4	6	213.8	106.0	6	200.3	91.9
	HoH	6	194.0	67.9	6	162.0	55.9	6	172.4	65.4
	Deaf	6	198.0	81.5	7	160.3	84.9	6	199.8	82.0
Mean reading time (in percent)	Hearers	6	44.6	21.8	6	51.8	19.7	6	43.3	19.5
	HoH	6	49.2	26.6	6	60.9	22.6	6	52.9	23.8
	Deaf	6	51.2	25.8	7	71.0	20.2	6	50.2	23.1

	Icons:Description	Icons:Icons	Icons:None

The first fixation lengths of the Deaf and Hard of Hearing are somewhat shorter for the *Icons'* video, most noticeably in the case of the Deaf (cf. Figure 5.19). The first fixation lengths of Hearers show hardly any difference for the three videos. The mean reading times in Figure 5.20 show the opposite effect, they are longer in the *Icons'* video for all three groups, most remarkably in the group of the Deaf. The icon itself was fixated in about 26% (Hearers), 37% (HoH), and 48% (Deaf), respectively, of instances.

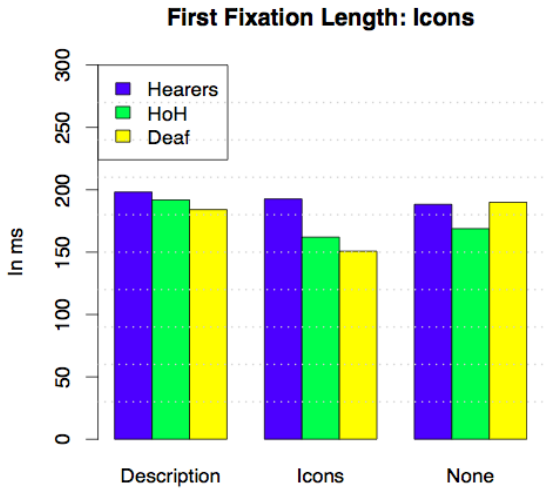


Figure 5.19

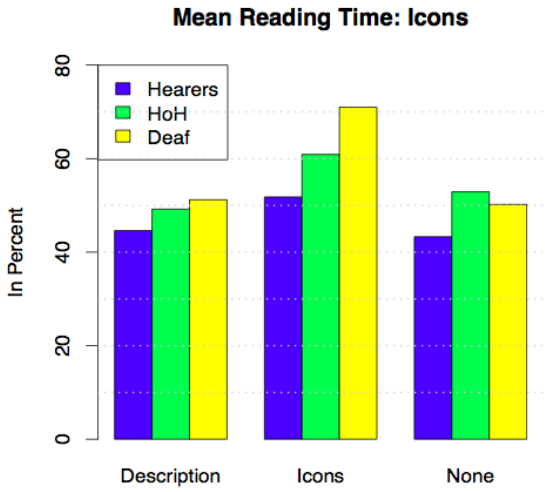


Figure 5.20

As far as comprehension is concerned, Hearers and Hard of Hearing achieved best results with *Description* subtitles. The Deafs' comprehension is the same for *Description* and *None*, but slightly less for *Icons* (cf. Figure 5.21).

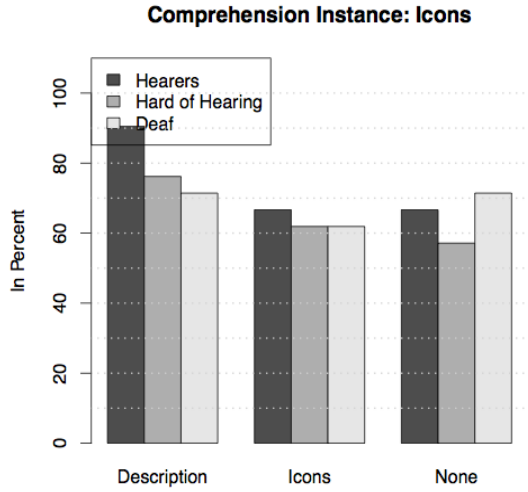


Figure 5.21

Looking at the comprehension instances separately in the figures below it is seen that textual comprehension tends to return lowest results in all three groups for *Icons: None*. Hearers achieved best results both for *Icons: Description* and *Icons: Icons*. The Deaf reached the highest rates with *Icons: Description*, whereas the HoH got the best results with *Icons: Icons* (cf. Figure 5. 22).

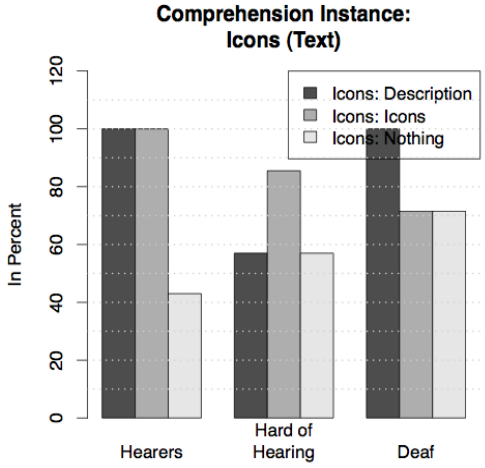


Figure 5.22

As to the other comprehension instances the video with *Icons* tends to return the lowest rates both for general and visual comprehension. The Hard of Hearing obtained best results in overall and visual comprehension with *Description* (cf. Figures 5.23 and 5.24). The Deaf obtained their highest results in general comprehension with *Icons: None*, and their highest results in visual comprehension both with *Icons: Description* and *Icons: None* (cf. Figures 5.23 and 5.24).

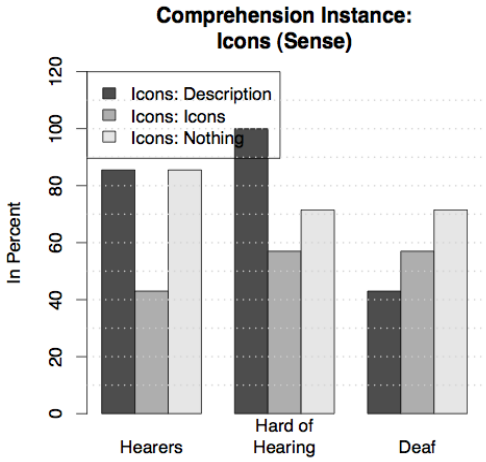


Figure 5.23

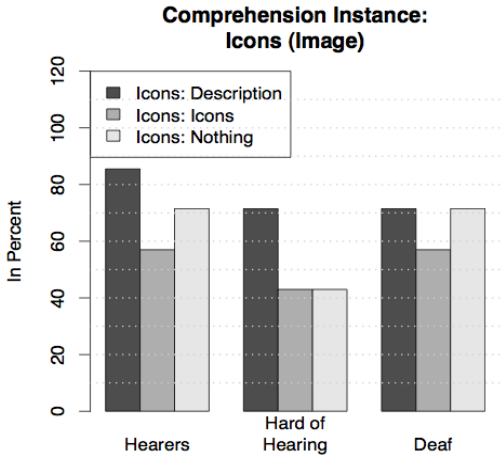


Figure 5.24

Comparing these data with the eye-tracking data below suggests a correlation between textual comprehension and mean reading time for the *Icons: Icons* video (cf. Figure 5.26). The eye-tracking data can however not explain why the comprehension rates of the Hard of Hearing with *Icons: Description* and the one of the Hearers with *Icons:None* were so poor (cf. Figure 5.25 and Figure 5.27).

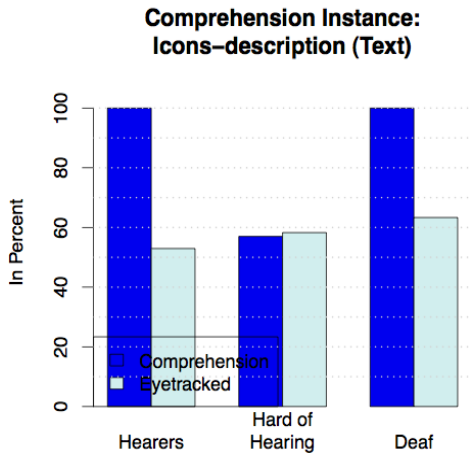


Figure 5.25

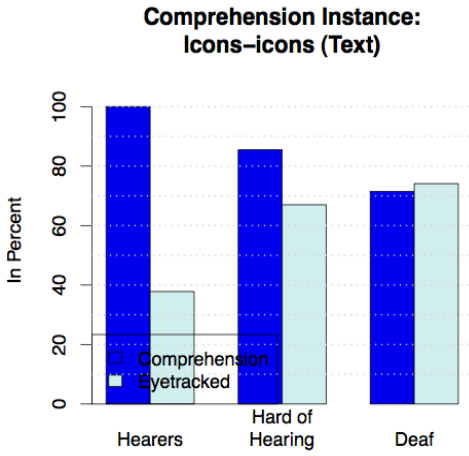


Figure 5.26

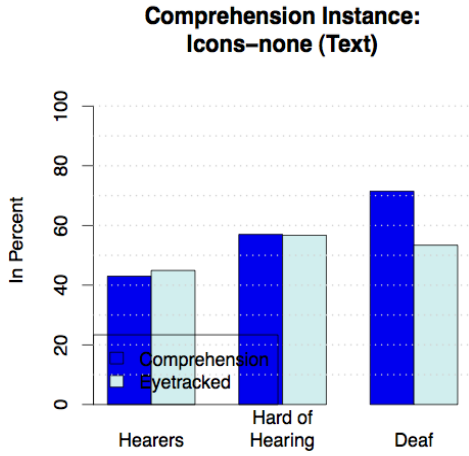


Figure 5.27

Looking at the overall comprehension in Figures 5.28 to 5.30 reveals that in the case of *Icons: Description*, where the Deaf obtained the poorest results, they also dwelled longer on the critical information (cf. Figure 5.28), yet again, the eye-tracking data cannot explain the poor comprehension results hearers achieved in the *Icons'* video (cf. Figure 5.29).

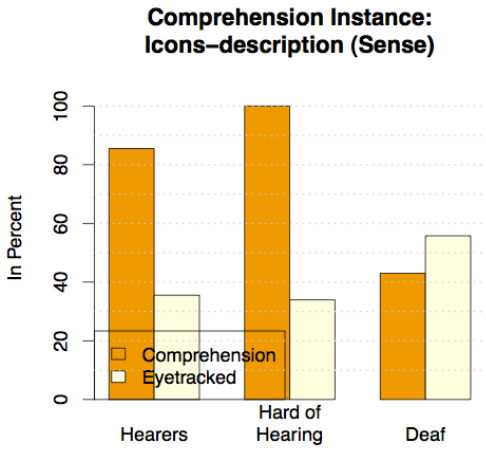


Figure 5.28

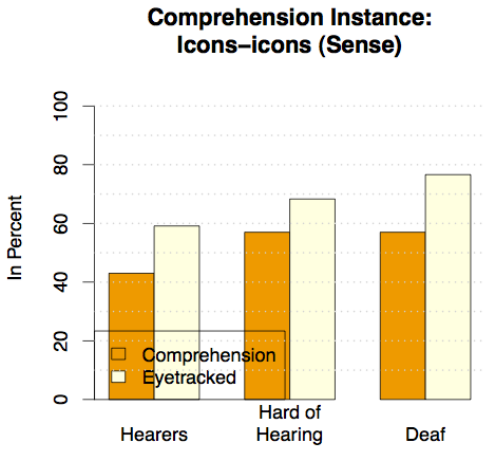


Figure 5.29

In the case of *Icons:None*, the critical information to answer the comprehension question was to be drawn from the image rather than from the subtitles. Thus, not reading time was the measure, but dwell time on the critical region of the image[13]. The data suggest that a longer dwell time on the critical objects leads to higher comprehension (cf. fig. 30), though in this case the differences between the data of the individual groups are too small to prove this.

[13] Elliptic IA with the coordinates 470,258 and 698,537

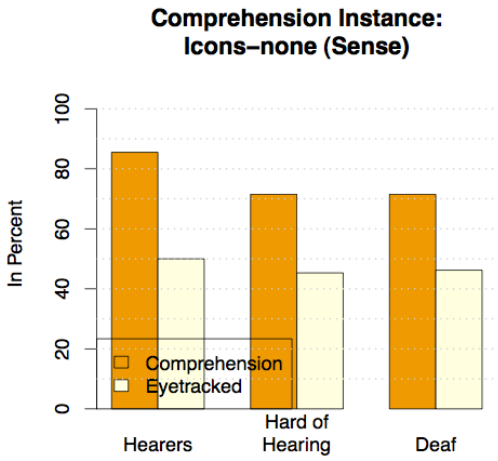


Figure 5.30

Eye-tracking data for the following graphs were obtained from the whole image excluding subtitles for *Icons: Description*, respectively from a rectangular IA[14] for *Icons: Icons*, and a freehand IA in the case of *Icons: None*. The differences in the eye-tracking data for the three groups of users in the *Descriptions'* and *Icons'* videos are, however, too small to say anything meaningful about them (cf. Figures 5.31 and 5.32). Only in the case of *Icons: None* was the dwell time visibly shorter with lower comprehension results at the same time (cf. Figure 5.33).

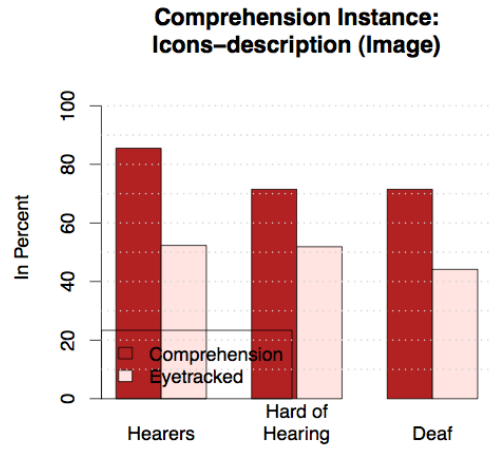


Figure 5.31

[14] Coordinates 433,276 and 608,480

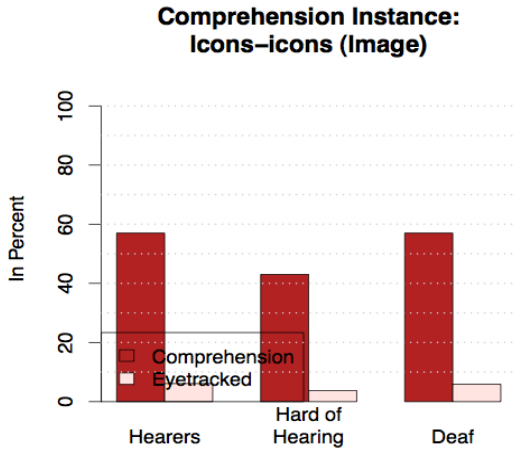


Figure 5.32

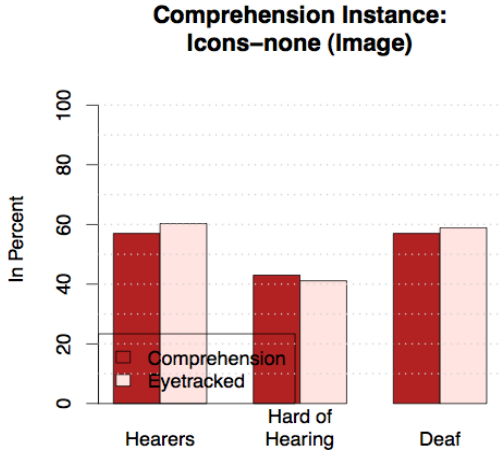


Figure 5.33

3. Speed

The standard reading speed of German SDH allows 12 to 13 characters per second with a number of characters per line of usually up to 36. The TV station ARTE allows up to 40 characters per line. As the *Verbatim* video required a line length of 42 characters, because otherwise the complete dialogue would not have fitted in, this maximum length was exceptionally kept for the *Standard* video, too. For the video of *Edited* subtitle speed, a norm of 9 characters per seconds was used, which is the standard for children's SDH of the subtitling company Titelbild, where the subtitles were created. All of the deaf participants and about 70% of the deaf respondents of the pre-test preference questionnaire chose verbatim subtitles as their preferred option.

The time to first fixation (cf. Figure 5.34) shows more or less the same pattern for all three videos. Reaction times for the Deaf and Hard of Hearing were slightly shorter with the *Standard* subtitling speed, whereas Hearers had shorter reaction times for the *Verbatim* subtitles.

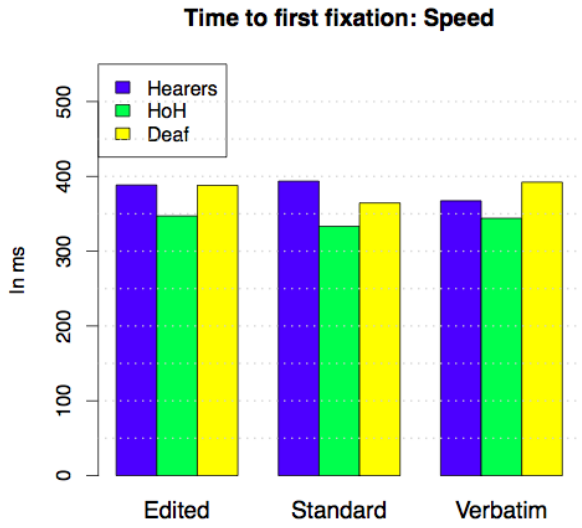


Figure 5.34

Table 5.3: Speed

		Speed:Edited			Speed:Standard			Speed:Verbatim		
		n	M	SD	n	M	SD	n	M	SD
Time to first fixation (in ms)	Hearers	5	388.5	219.8	5	393.6	210.3	6	367.4	207.4
	HoH	7	347.1	194.2	5	333.2	184.3	5	344.0	186.6
	Deaf	6	388.0	234.5	6	364.5	204.6	5	392.3	208.5
First fixation length (in ms)	Hearers	5	197.2	67.3	5	202.8	121.3	6	195.1	66.2
	HoH	7	174.0	56.8	5	176.9	59.2	5	179.0	80.6
	Deaf	6	185.8	80.2	6	183.0	59.9	5	194.7	153.1
Mean reading time (in percent)	Hearers	5	54.6	24.7	5	57.9	20.5	6	60.9	24.4
	HoH	7	60.3	26.7	5	69.4	22.8	5	74.5	22.7
	Deaf	6	61.9	25.3	6	70.1	25.7	5	76.8	22.0

The first fixation lengths in figure 5.35 don't show any noteworthy differences, either. However, mean reading times (cf. Figure 5.36) become clearly longer with increased subtitling speed, most remarkably in the case of the Deaf and the Hard of Hearing.

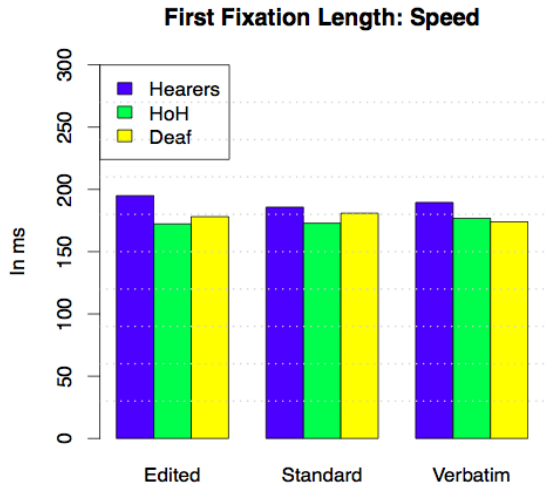


Figure 5.35

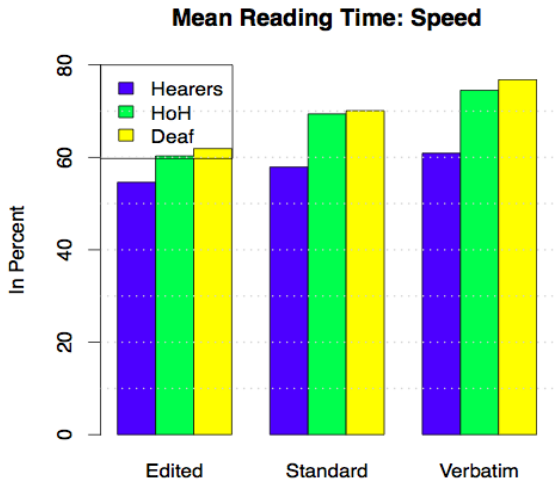


Figure 5.36

Comprehension results could not be retrieved from the *Speed* clips, as the same clip was shown three times, but the standard video, together with the videos from the other two parameters, was used to calculate the mean number of characters per fixation.

Hearing viewers grasped more information with one fixation (7.1 characters/fixation) than did deaf and HoH viewers (cf. Figure 5.37), which is not surprising as reading times were also shorter. It is maybe more surprising that Hard of Hearing went through slightly less characters per fixation (5.5) than did the Deaf (6) with mean reading times usually somewhat shorter. That means, that the hard of hearing participants fixated the subtitle more often, but at the same time the fixations were shorter than those of the other groups, a fact that the first fixations lengths and times to first fixation suggest, too. However, the difference is so minor that it can be neglected.

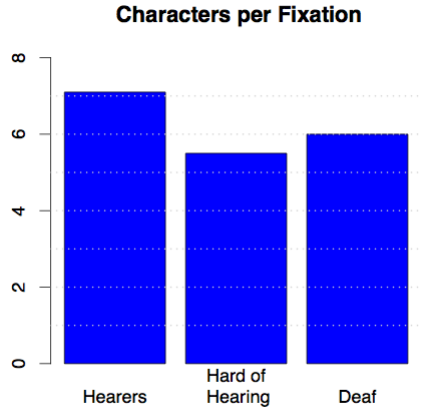


Figure 5.37

4.4 Conclusions

Emoticons: Mean reading time was slightly reduced with the *Emoticons: Emoticons'* video, and in terms of comprehension it returned roughly the same results as *Emoticons: None* for the hearing impaired. General comprehension was even at a rate of 100% for the Deaf. Therefore, using emoticons is worth considering in future subtitling. It has yet to be proved whether these differences are mere coincidence or can be confirmed in further tests. In both pre-test preference and post-test preference questionnaires the majority of the hearing impaired chose description rather than emoticons.

Icons: Both mean reading times and comprehension instances strongly indicate that description in this case is the best option. Here again, further investigation would be necessary to prove this assumption.

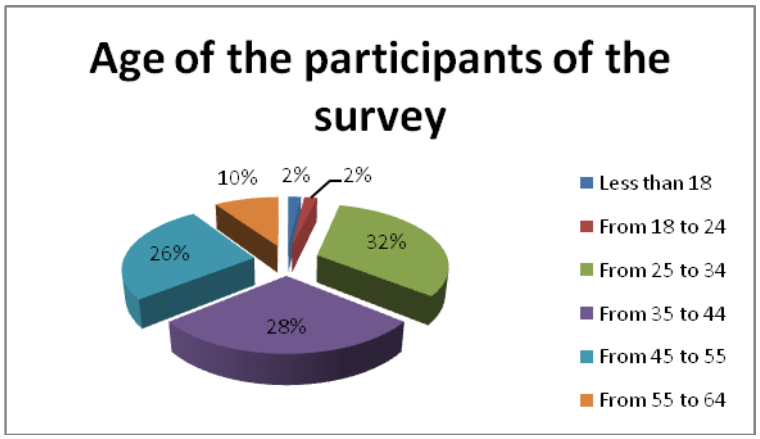
Speed: As expected, the dwell time on the subtitles increases proportionally to the subtitling speed. Five of the Deaf participants liked the verbatim video best; the other two did not opt for any of the presented videos. The Hard of Hearing, however, quite liked the simpler subtitles that would only give the key information. Given that in the pre-test preference questionnaires the majority stated that the speed of TV subtitles is usually about right, there is, however, no need for reducing the speed further. It seems thus that the speed in current subtitling is the ideal speed.

5 TVC Audio Description Evaluation Report

5.1 TVC (TV3) Mature Services Evaluated

This report analyses the results obtained after an on-line survey conducted by TVC, and which was offered at TVC's web page for several weeks in the beginning of 2010. The target of the survey was people who would benefit from TVC's Audio Description (AD) mature service: people with visual impairments and who understand the Catalan language. Their opinion will help us better know the customer insights for this service and, therefore, have the information that is needed to perform the right improvements on the service.

PARTICIPANTS' INFORMATION
A total of 53 people participated in the survey, 19 of whom where women; the remaining 34 participants were men. The age distribution of the participants is shown in the figure below:



The participants were from different cities of Catalonia (Barcelona, Sant Feliu de Llobregat, Manresa, Collbató, Cabrera de Mar, Reus, etc), but also from other parts of Spain (Vinaròs, València, Santurtzi, A Coruña, Murcia).

Nearly all of the participants stated they understand and speak Catalan perfectly well, and 3 of them said they understand Catalan but with certain problems. Regarding their degree of blindness, 47% of them answered they had the official recognition of total blindness, 47% of participants answered they were partially sighted, and 2% of the participants answered they have no visual impairment. That information was confirmed and detailed through question 6.

Among the people who answered they had the official recognition of total blindness:

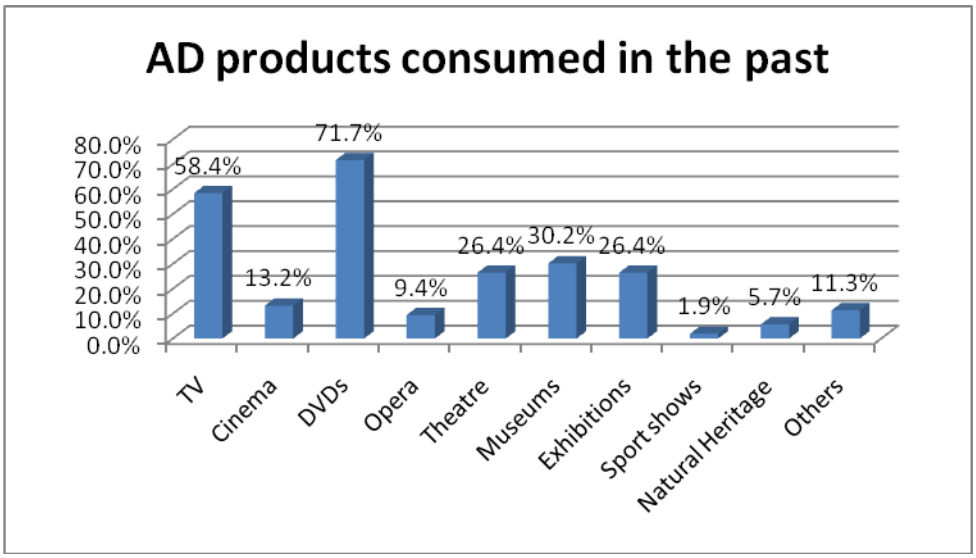
- 40% of the participants chose only the sentence: "*In a room during daytime, I can tell by the light where the windows are*"
- 4% said they could "*see the shape of the furniture in a room*"
- 24% participants chose the sentence: "*I cannot see anything at all*"
- A 40% did not answer the question.

The participants who answered they were partially sighted chose all options but "*I prefer not to answer*" and "*I can see well enough to recognise a friend across the road*." That is:

- 44% stated *"I can see well enough to recognise a friend across a room"*
- 40% stated *"I can see well enough to recognise a friend who is at arms length away"*
- 48% stated *"I can see well enough to recognise a friend if you get close to his or her face."*
- 40% stated *"I can see the shapes of the furniture in a room"*
- 40% stated *"In a room during daytime, I can tell by the light where the windows are"*
- 28% stated "*I cannot see anything at all"*

AD CONSUMING HABITS

Regarding AD consuming habits, the following graphic shows the percentage of participants who had consumed different products with AD at the moment of the survey.



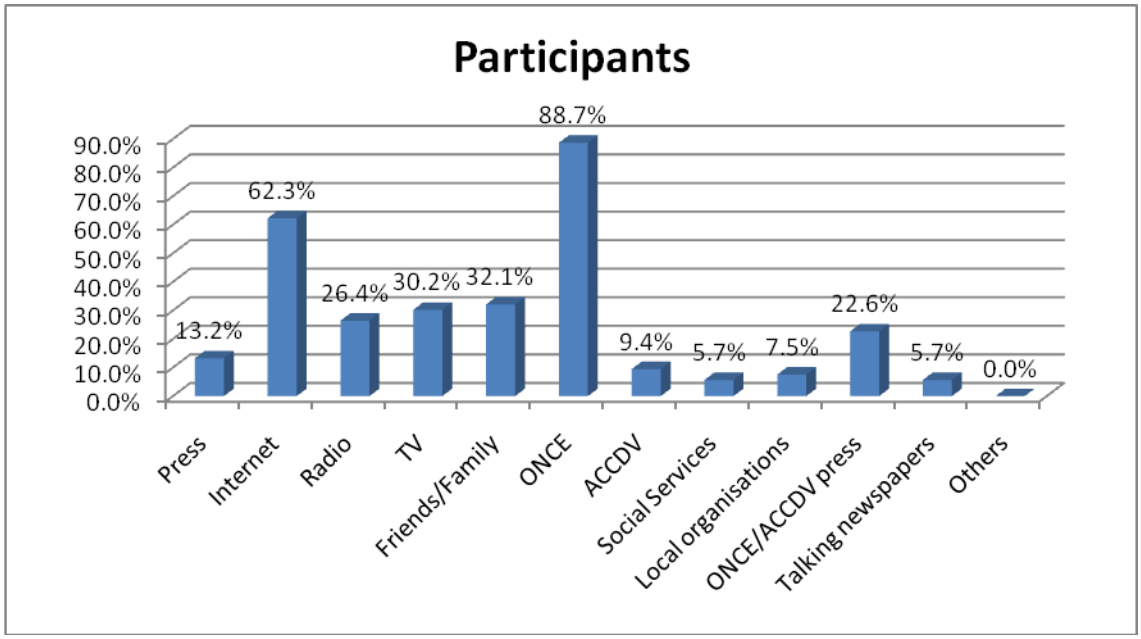
A sports event, which was another possible option, was chosen only by 1 person out of 53. Probably that is because nearly no sports events are offered with AD. The Natural Heritage ratio is also very low, less than 6%. DVDs and TVs are by far the most popular ways of consuming AD enriched contents, while only about one tenth of the participants have consumed audio description in Cinema, Opera, and other different products. Approximately, one out of every 3 participants has had a previous AD consuming experience related to theatre, museums and exhibitions. Provided all of these products are based on a business model that can supposedly support the costs of AD, the gap can then be explained because of two considerations: on one side, the bigger the audience the bigger the pressure on the content provider side to provide AD contents; on the other hand, some of the products are not frequently consumed, and that clearly affects the results of the survey. For example, a Spanish study of 2002[15] shows that in Spain in the

[15] Ministerio de Cultura, the numbers of the culture in Spain in 2002: http://www.mcu.es/estadisticas/MC/CCE/DescargaDatos.html

year 2000 99.5% of the homes had at least one TV, that 104 million DVDs or VHS were hired or bought that year (population was little more than 40 million in 2002), that the attendance at cinemas was 3,400 viewings for every 1,000 inhabitants, that the attendance ratio for the theatre was 314 / 1,000, that more than 5 million people went to a Spanish museum that year (some of whom were of course tourists).

SOURCES OF INFORMATION FOR PRODUCTS WITH AD

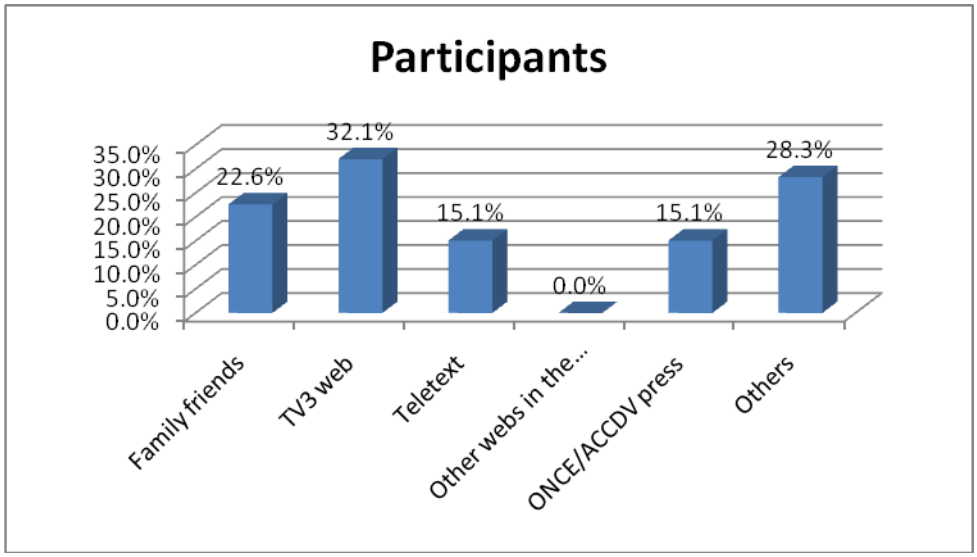
Regarding sources of information for products specially designed for blind and partially sighted people, the following graphic shows the percentage of participants using each kind of source.



Internet and ONCE are indeed the most popular sources of information for visually impaired inhabitants, while radio, TV, friends and family, and the specialized press is also used by a notable 22% to 32% of this segment of the population.

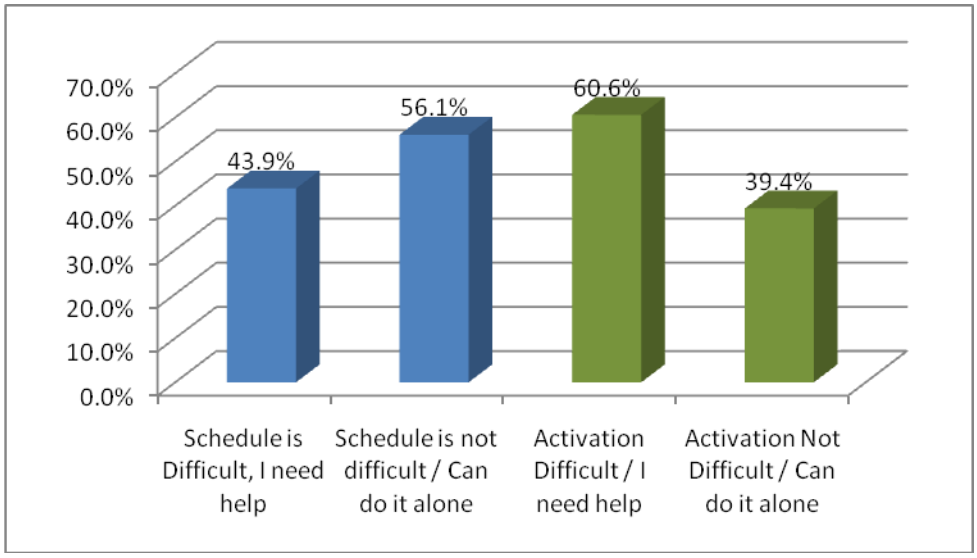
SOURCES OF INFORMATION FOR TV3 AD AND ACTIVATION THROUGH DTT

Regarding sources of information for AD in TV3, the following graphic shows the percentage of participants using each kind of source.



Here, most people only have one source of information. As an average, people either look at the web page of TVC on their own, or ask family and friends, or use an unspecified source of information (probably, ONCE) to know which programs are broadcasted with AD. As can be inferred from the graphic, there is clearly not a predefined and preferred channel to use to know which programs will be available in AD format in the near future.

With regards to the access to the TV3 AD schedule and the activation of the AD through the Digital Terrestrial Television (DTT), 28% of the participants did not answer the question about the difficulty of the accessibility of the schedule, while 38% did not answer the question about the difficulty of the activation of the AD channel. The next figure shows the percentages that were observed for the answers that were given.



COMMENTS

Some qualitative questions were included in the survey in order to gather information that cannot be obtained by statistical means. The more relevant inputs can be found in the following section.

Which are your proposals to improve the access to the AD schedule in TVC?

The answers to this question were the following:

1. "The AD channel should not deactivate automatically. That way, whenever there is AD it would come on by default."

2. "The most important problem is that DTT receivers are not accessible. I think the publicity offered by TV3 on AD is quite good, since it allows you to know, just before the programme starts, whether it will have AD."

3. "More information on TV3's own programmes, sending text messages or e-mails on the programmes with AD to all those people who ask for it."

4. "Accessible DTT receivers. An option in TV3's web page allowing users to subscribe to a service through which information on the AD programmes offered would be sent on a weekly basis."

5. "Accessible DTT receivers in order to interact without difficulties."

6. "ONCE should be the negotiator with TV3 as the representative of blind people in Catalonia. ONCE should have in advance the AD schedule in order to spread it among the blind and the visually impaired."

7. "Keep informing on TV3, as has been done until now, about the programmes offering AD. Also informing through the radio and other media."

8. "The biggest problem is that DTT receivers are not accessible. I believe the promotion that TV3 makes on AD is more than good enough. It allows us to know, just before the broadcasted program, that it uses audio description"

9. "DTT decoders which are accessible. An option on the Web in TV3 to allow mail subscription to receive information on the programming with audio description on a weekly basis"

10. "DTT devices that are accessible so that we can interact seamlessly."

11. "Spoken menus and interactivity"

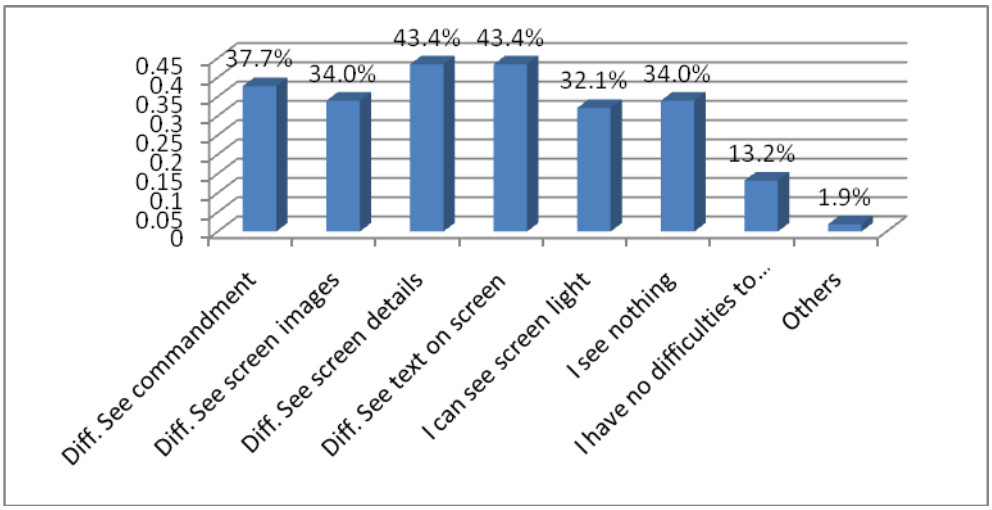
12. "Pay per view content operators such as Imagenio, should also broadcast programs with audio description."

Many of the suggestions of the users (points 1, 2, 4, 5, 8, 9, 10) refer to possible improvements to the electronic receivers, i.e., set top boxes and televisions, not so much by the broadcaster itself. Users perceive that the DTT equipment manufacturers could have a higher sensitivity to their needs: they should include by default accessibility options.

Other considerations (3, 6, 9) mainly refer to the need of the participants to know in advance the programs with AD. Besides a pair of participants have indicated that they would like other channels to follow TVC's example and include AD in their priorities.

MORE ACCESS DATA

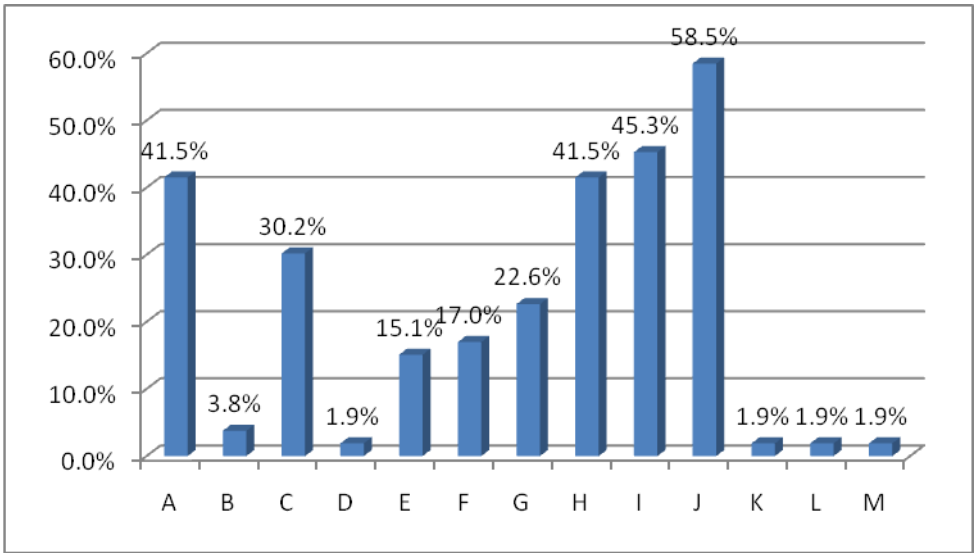
Regarding the ability to see different gadgets and information related to the fact of watching TV, the following graphic shows the percentage of participants for each answer.



The available answers were:

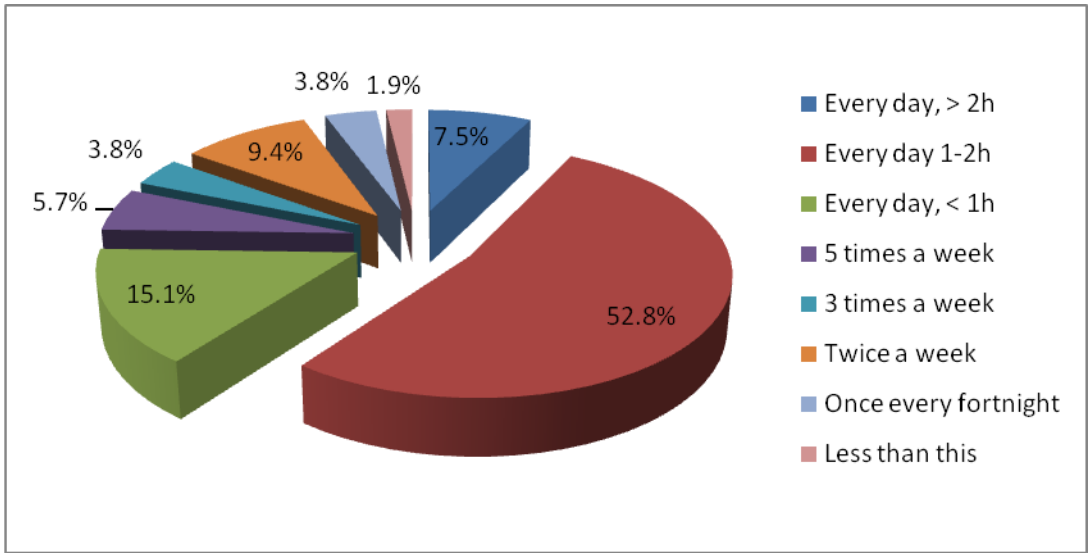
A. I have difficulty seeing the buttons on the remote control
B. I have difficulty seeing the picture on the TV screen
C. I have difficulty seeing the fine detail on the TV screen
D. I have difficulty seeing text on the TV screen
E. I am able to see the light of the TV screen
F. I cannot see anything on the TV screen
G. I do not find that I have any difficulty following what is going on the screen
H. Other...

Regarding the assistance they need while watching TV, the answers given were the following.



A. I use my residual sight to watch
B. I wear special stronger glasses
C. I get closer to the TV screen
D. I use a magnifier
E. I adjust the screen settings
F. I adjust the lighting in the room
G. I use a large screen TV
H. I ask my friends or family members to assist me by explaining what happens on the screen
I. I just try to pick up as much as I can from the sound of the film or programme
J. I use audio description to explain to me what happens on the screen
K. I make none of these adjustments
L. I never watch TV/ DVD(s)
M. Other

When consuming content, most of the polled people, declared that they used AD and asked for some help from non-visually impaired people, or simply accepted they would be losing information even though they would pay extreme attention to the audio, use their residual sight or move closer to the screen. A small percentage admitted that they modified the physical settings (light, use of glasses or lens, screen settings etc.). Doing so can sometimes affect the quality of experience of the other members of the family while watching TV. Finally, it should be noticed that "*I use AD to explain me what happens on the screen*" is the most highly valued option. That clearly means that AD is considered by visually impaired people to be a powerful tool that increases their understanding of what is happening on screen.



How often do you watch TV? (One answer only)

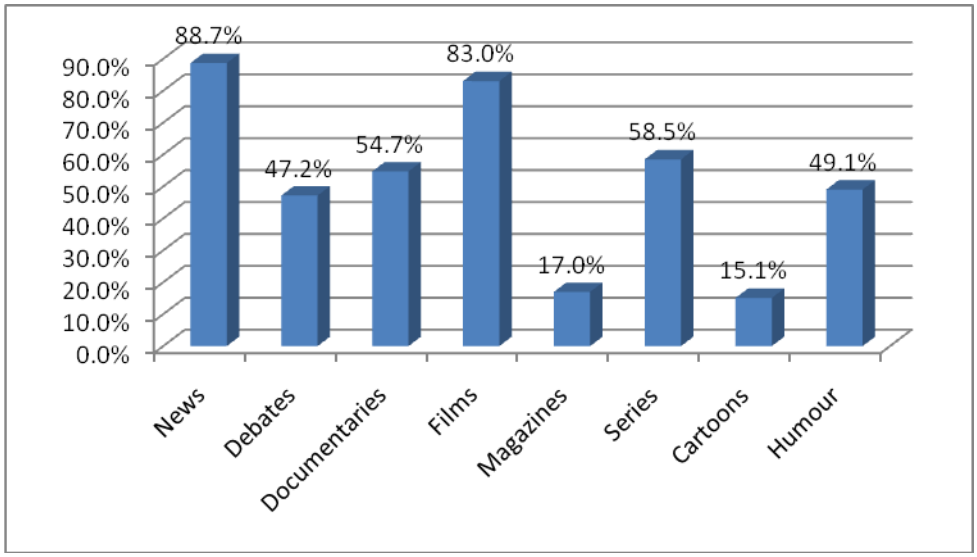
A. Every day, more than two hours.
B. Every day, from one to two hours.
C. Every day, less than one hour.
D. Five times a week.
E. Three times a week.
F. Twice a week.
G. Once every fortnight.
H. Less than that.

Three out of four people polled (in fact, 75.5%) watch TV on a daily basis, and 94.3% do it at least twice a week. This is a percentage that perfectly matches with the statistics of the overall population: 73.4%, or 32.3 million out of 44 million people in Spain, watched TV daily in Spain in 2006[16].

What type of programmes do you watch on TV? (Choose as many options as you want)

a. News
b. Debates
c. Documentaries
d. Films
e. Magazines
f. Series
g. Cartoons
h. Humour programmes

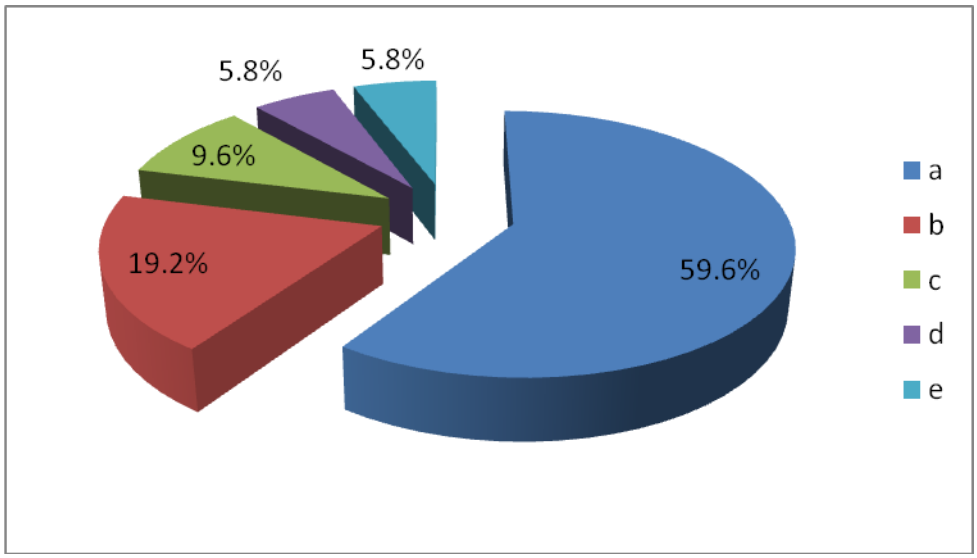
[16] http://www.televisiondigital.electronicafacil.net/Article6659.html



News is still the killer contents on TV, although films, series, documentaries, humour programs and debates have also a significant share. Again, these results match with those of the majority of the population. As was before mentioned, only 2% of the participants were under 18; as a consequence, it is logical that cartoons have only a 15% share amongst the participants.

Do you usually watch the programmes with AD that TV3 broadcasts? (One option only)

A. Yes.
B. No, I did not know TV3 broadcasts some AD programmes.
C. No, I know TV3 broadcasts some programmes with AD but I do not know how to activate the AD.
D. No, because I do not get DTT yet, but I would like to.
E. No, I do not need AD.



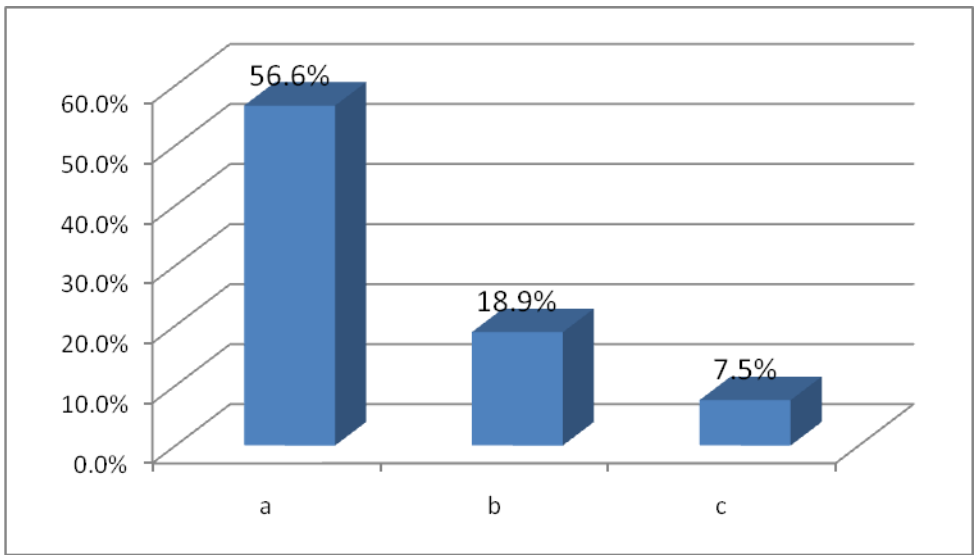
AD is used whenever it is available, viewers are aware the program is AD enriched, and their DTT is prepared for it. However, 20% of the participants did not know that AD

services are available at present. Besides, a 16% are not able to try AD, either because their DTT receiver does not allow it, or because they do not know how to activate this service, or because they received no DTT signal at home at the time the survey was done.

On the other hand, with regards to AD in TVC 94.2% of participants think that there should be more AD on TV, while 5.8% think the number of programmes with AD provided by TV3 is enough. This is quite logical, as it is quite a new service. AD has been broadcast since the 1st of January 2007 thanks to DTT broadcasting, as DTT can handle a single audio channel for AD. For example, every week "The "Great Movie" is broadcasted with AD every Friday and several children's series every week.

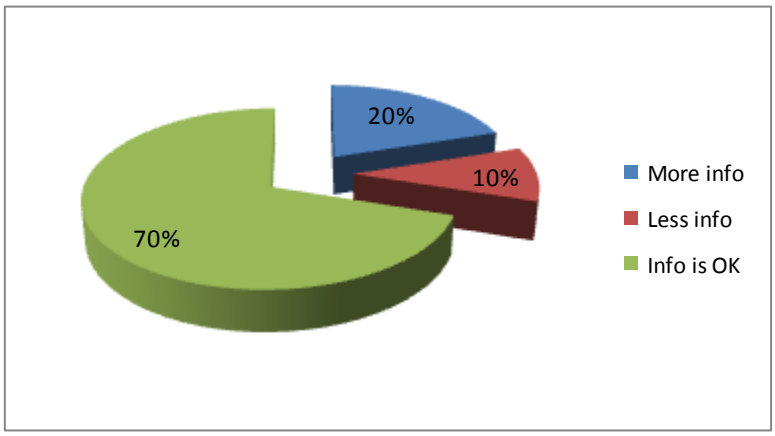
Which programmes with AD have you ever watched on TV3?

a. La Gran Pel•lícula.
b. Doraemon
c. Series for youngsters such as "Els desastres del rei Artús", "Em dic Eve" or "L'hotel zombi".



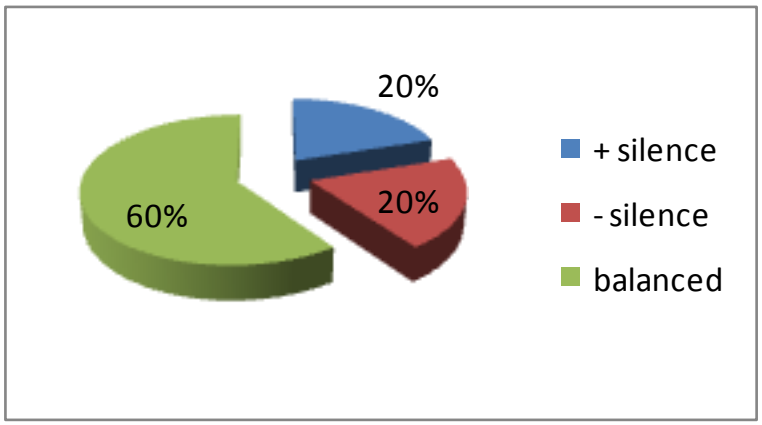
Please, choose one of the options to complete the following sentences based on your opinion (note: only 10 people answered these questions):

a. Taking into account that the space available is limited, audio descriptions...
   1. Should include more information.
   2. Should include less information.
   3. Usually include the adequate information.



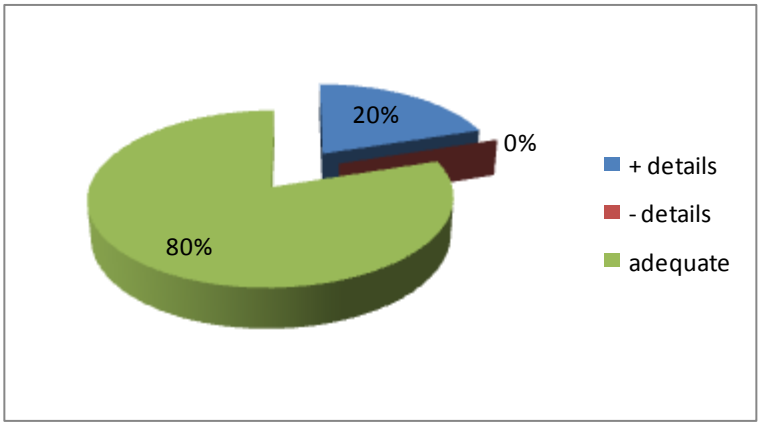
b. Audio descriptions...

1. Are usually overloaded, some silent moments are missing
2. Are usually too plain, there is too much silence.
3. Usually find a balance between silence and dialogue.



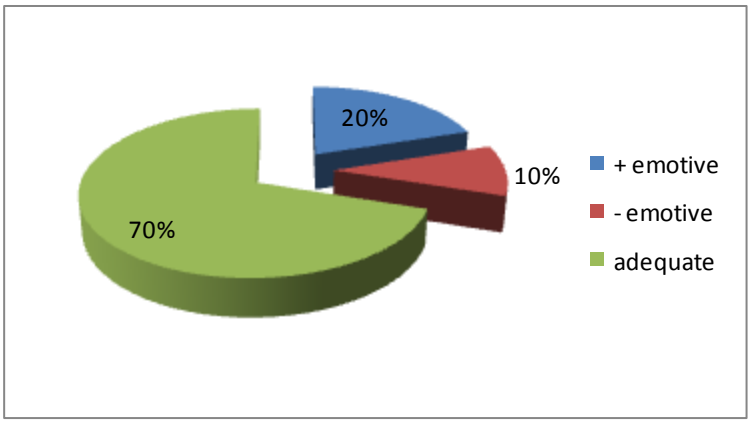
c. The information provided in the audio descriptions...

1. Should be more in detail.
2. Should be less in detail.
3. Is usually adequate.



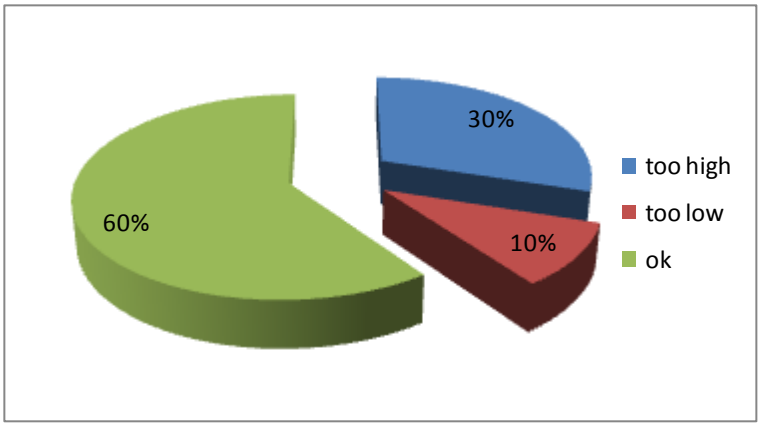
d. Narration in audio descriptions...

1. Should be more emotive.
2. Should be more neutral.
3. Is usually adequate.



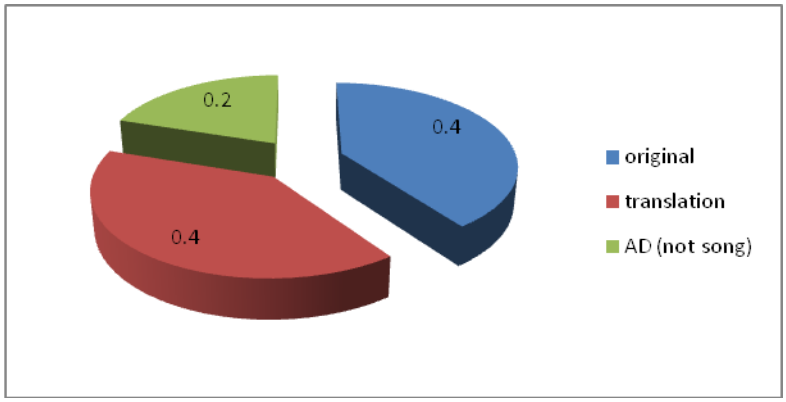
e. Audio descriptions...

1. Are usually too loud.
2. Are usually too quiet.
3. Are usually all right.



f. If there were spoken songs in other languages...

1. I would rather listen to them as they are.
2. I would rather listen to the translation, if there were subtitles.
3. I would rather listen to the details of the programme through the AD, although it was not essential to follow the argument.



The results in this section show that, technically speaking, the AD service is found to be adequate by its consumers. In other words, the information, overload, volume, and narrations of the AD services at TVC are well balanced and are fitted to their needs.

Which are your proposals to improve AD in TVC?

1. "I like it as it is, but I guess in the future they will offer more programmes with AD"

2. "When there is a song, do a quick summary if it is important for the argument and let the rest of the song be listened to in the original language. Find a balance among sounds so that none of them overlaps the other. The audio track should always be stereo."

3. "More information, more films, more AD"

4. "Sound quality should be improved (too much micro noise) and the voice volume should be lowered, it's too high with regards to the sound of films."

5. "More publicity should be done"

6. "It would be interesting to speak about technical cinematographic aspects such as lighting, focusing, shot, etc."

7. “To be careful so that everyone can understand it, to repeat names, to describe everything, and to consider first the needs of the blind people more than those of the visually impaired”

8. “The times that I've enjoyed it I think it is correct”.

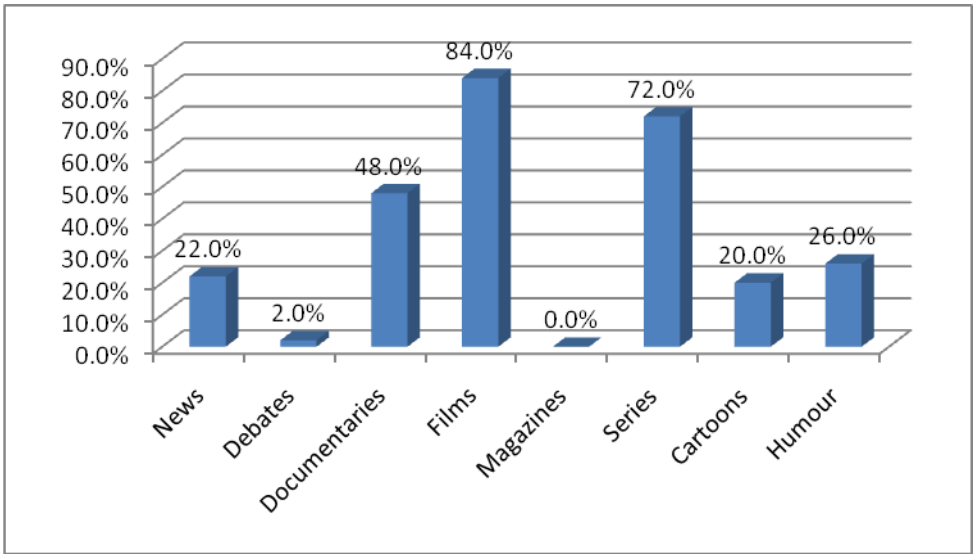
9. “Besides the movies, it should be extended to documentaries”.

10. “The AD volume should be monitored with respect to the film. Sometimes it is very strong or very weak”

Most of the proposals and suggestions that deal with AD have to do with their availability. Comments mainly have to do with the percentage of content with AD, as well as with the information that lets consumers anticipate when contents with AD are broadcast. Besides two participants suggest that the sound quality could be improved, another says that AD should not mask background music and a last one suggests that relevant photographic aspects (like lighting or shot description) might be also be audio described.

If TV3 had to widen its schedule with AD, which programmes would you prefer to be audio described? (Maximum of three)

a. News
b. Debates
c. Documentaries
d. More films
e. Magazines
f. Series
g. Cartoons
h. Humour programmes



Series films and documentaries are considered by the blind people the most valued to be audio described. Probably, cartoons are easier to follow, and humour relies quite a lot on the script, so that is why the perception of the need for AD seems lower for these kinds of contents.

Of the whole schedule of programmes in Televisió de Catalunya, which programmes would you like to have audio described:

"Damages that will start in January"

"Films and series" (several answers)

"Ventdelplà, 30 minutes, more films"

"Most of them"

"Especially films and series, but I cannot specify since I don't know the schedule very well"

“Polònia, Crackòvia, Més Dinamita, Ventdelplà.”

“Quizzes like Bocamoll, and el Gran Dictat, as we would then be able to participate more at home”

“Films and documentaries”

“Polònia, Crakòvia”

Any other comments:

Finally, the last question of the survey was an open question: “*Any other comments?*” Here, we find a two pieces of technical advice: we should put in an effort to audio describe any text on screen, and to foster spoken menus technology, so that blind people can be autonomous while watching TV.

Besides, the survey also showed that the community of blind and partially sighted people are aware of the technical and economical efforts of the AD service. And, more important, they are aware that TV3 is pioneering this service in Catalonia (and Spain), and they had some kind words for encouraging the work that is being carried out by TV3.

"Currently, I can listen to the AD autonomously, but it has been really hard to manage and remember the process. It is difficult to automate the process since one does not do it usually. The use of remote controls should be simpler and easier to control. I think it's very positive that, after the proposal we made to the Associació Catalana per a la Integració del Cec, audio navigation was included in the TV products edited in DVD, although we asked for it just before the "Les Veus del Pamano" DVD came out. Good job! Without audio navigation, AD cannot be accessed autonomously."

"Congratulations for all you've done until now. I hope the AD schedule will increase soon"

"It would also be very interesting to have an audio track with the AD in Catalan and the film in its original language."

“It would be very interesting to have one audio channel with AD and the film in One Shot Recorder (OSR) and mix it”.

"If you do not already have it, you should have contact with ONCE, which is the organization that represents blind people. They know rules and know how to make audio description in a suitable way, and what the programs are that interest blind people.”

“Congratulations you're the only ones who do audio description."

”I value very positively that audio navigation has been introduced to TV products that have been edited in DVD format, as has already been done with "*The Voices of Pamano*". If there is no audio navigation, then we cannot access the audio description”

"Very important: in the normal movies, when there is text on screen it should be audio described. We the visually impaired lose a lot of information there. Thank you, thank you, thank you, for everything you do Rosa V. who promptly sends me the schedule, is very charming and attentive"

“The audio description is a very good initiative to remove barriers that affect people, and to let communication arrive to all citizens under equal conditions. The next step would be adapting the DTT menus with voice, which would give us freedom to autonomously choose the programs we want to watch”

“Just thanks to this study, I hope it will have good repercussions on the quantity and quality of audio descriptions in TV3, and perhaps also on other broadcasters”.

“Congratulations for the work done so far, hoping the program will expand.”

Conclusions

As a conclusion, this survey showed that most blind and visually impaired users of AD approve of the way TV3 currently carries out AD with regards to the information provided and the way it is delivered. Most users only have one source of information on the AD schedule, either on the web page of TVC, through family and friends, or through an unspecified source of information (probably, ONCE), but most of them do not find it difficult to get it. However, it has been spotted that around half of AD users find it difficult to activate AD on their own. This is due to the great variety of DTT receivers and to the fact that they do not provide audio navigation through the menus. To conclude, it is important to note that most users acknowledge the effort TV3 is doing to provide AD, although they also demand that more programmes should be offered with this service, especially more movies and series. This demand should be taken into consideration when planning the future expansion of AD.

6 Audio Subtitling or Spoken Subtitles in Europe

Audio subtitling is the media accessible mode of reading subtitles aloud by a human, or voicing subtitles through speech synthesis software programmes.

With the exception of the European dubbing countries, Italy, France, Germany and Spain, the rest of Europe subtitles foreign productions for all but small children's programmes. Accessibility for those with sight problems should be adapted for material in a foreign language subtitled in the local language.

To add to those with sight problems, it has been estimated that 40 to 60 per cent of elderly people encounter problems when reading subtitles, and experience difficulties with the size of the subtitle letters and the contrast between the colour of the subtitle letters and the screen background.

The cost of audio description is an important issue as explained by Verboom et al (2002: 297)[17] AD is "a very expensive approach and it is not financially feasible for most broadcasting companies to make significant portion of their programmes accessible" hence alternative solutions were researched.

Background

The project "Spoken Subtitles" was developed and evaluated with a high degree of acceptance, and on 14th December 2001 the audio subtitling service was officially opened by the Dutch Secretary of State and the president of the NOS. Audio subtitling has been implemented as a permanent service on Dutch TV[18]. The software used is ScanSoft.[19]

Technological possibilities

Two possible avenues for broadcasting audio subtitles are:

- Broadcaster mix, when a separate audio channel is sent in the multiplex, and the user can make the choice to accept/reject the extra information.
    - Broadcasters needed a speech-synthesis computer which is fitted with speech-synthesis software which converts the text into speech. This output is then converted into a signal and broadcast without disturbing the programme.
    - This system for a mono channel has an approximate cost of 30,000 euros per year, but there are instances where it can be provided for 7,000 euros.
- Receiver mix, when the user has a decoding system inbuilt in their set-top-box.

[17] Maarten Verboom, David Crombie, Evelien Dijk and Mildred Theunisz, "Spoken Subtitles: Making Subtitled TV Programmes Accessible", http://www.springerlink.com/content/6jugv491muwbknc2/

[18] Mildred Theunisz, 2002, "Audiosubtitling: A new service in Netherlands making subtitling programmes accessible", http://www.springerlink.com/content/6jugv491muwbknc2/

[19] http://www.encyclopedia.com/doc/1G1-83705170.html

Criticism

Negative

Audio subtitling or Spoken subtitling has met with a negative r**e**ception by user associations across Europe. Their main concern is the rejection of a synthetic voice, favouring human intonation.

Positive

While the preference for a human voice has been clearly stated, it is true that:

- Blind people across Europe are using DAISY: an acronym which stands for Digital Accessible Information System. It is now a recognised worldwide standard for audio developed by the DAISY Consortium. DAISY is a digital reading format that can combine audio, text and graphical information in one production, making it accessible to a wide range of people with print disabilities. DAISY can be played on a standalone DAISY player, or by using the DAISY software player on a computer. People across the world are used to listening to DAISY speech synthesis. You can hear the voice heard from DAISY at:

  http://www.youtube.com/watch?v=HZicWnCfHlw&feature=related

- The synthetic voice system have improved enormously in recent years, you can listen to a couple of demonstrators and here you have a simulator which can read the text of your choice, so you can appreciate the new generation of Speech Synthesisers:

  http://www.loquendo.com/en/demos/demo_tts.htm

  or

  http://212.8.184.250/tts/demo_login.jsp

- Time and financial backing are needed to create human audio description.
    - With audio subtitling a high number of hours of subtitled programmes in a foreign language could be made available in the user's language, since it is instant and almost cost free.