This is the accepted version of the article.

To cite this article: Mangiron, C. (2016). Reception of game subtitles: an empirical study. The Translator, 22 (1), 72-93.

https://www.tandfonline.com/doi/abs/10.1080/13556509.2015.1110000

Reception of game subtitles: An empirical study

Carmen Mangiron

Departament of Translation and Interpreting & Centre for Accessibility and Ambient Intelligence in Catalonia, Universitat Autònoma de Barcelona

Facultat de Traducció i Interpretació, Edifici K, Campus de la UAB, 08193 Bellaterra, Barcelona, Spain

carme.mangiron@uab.cat

Acknowledgements

This research is supported by a Post Doc Scheme Ref. 2008 BP B 00074 by the Catalan Government and by the Spanish Ministry of Finance and Competivity project no. FFI2012-39056-C02-01 *Subtitling for the deaf and hard of hearing and audio description: new formats*, as well as the Catalan Government funds 2014SGR27, and the European project Hbb4All from the FP7 CIP-ICT PSP.2013.5.1 # 621014.

Reception of game subtitles: An empirical exploratory study

Abstract: Over the last few years accessibility to the media has been gathering the attention of scholars, particularly subtitling for the deaf and hard of hearing (SDH) and audiodescription (AD) for the blind, due to the transition from analogue to digital TV that took place in Europe in 2012. There is a wide array of academic studies focussing on subtitling and SDH in different media, such as TV, cinema, and DVD. However, despite the fact that many video games contain cinematic scenes, which are subtitled intralingually, interlingually or both, subtitling practices in game localization remain unexplored, and the existing standards widely applied to subtitling for TV, DVD and cinema are not applied. There is a need for standardisation of game subtitling practices, which will ultimately lead to an enhanced gameplay experience for all users. This paper presents a small-scale exploratory study about the reception of subtitles in video games by means of user tests through a questionnaire and eye tracking technology in order to determine what kind of subtitles users prefer, focusing on parameters such as presentation, position, character identification, and depiction of sound effects. The final objective is to contribute to the development of best practices and standards in subtitling for this emerging digital medium, which will enhance game accessibility not only for deaf and hard of hearing players but also for all players.

Keywords: video game subtitles, accessibility, subtitle reception, eye tracking

Introduction

Video games have become a pervasive and ubiquitous form of entertainment in our society, increasingly being used beyond entertainment, for education and therapeutic purposes. Since the introduction of sixth generation consoles in the late 1990s —such as the PlayStation 2, Nintendo's GameCube and Microsoft Xbox—, the capacity of which allowed for the first time the use of voiced-over[1] dialogues, game technology has improved dramatically, allowing games to display high-quality graphics to accompany their complex story lines and scripts. Nowadays, many original games include intralingual subtitles for cinematic scenes. Localised versions of games can be dubbed into the target language and subtitled intralingually or can be only be subtitled into the target languages. In the case of Japanese games, they are usually dubbed into English and this version becomes the pivot version for the French, Italian, German and Spanish (FIGS) versions, which maintain the English soundtrack and provide interlingual subtitles.

Despite the widespread use of subtitles in games, current subtitling practices often do not follow commonly accepted audiovisual translation (AVT) industry standards. For example, console platform holders require that intralingual game subtitles are verbatim, while intralingual subtitles for TV and DVD are usually condensed, in order to allow users, often deaf and hard of hearing, to have enough time to read them. In addition, subtitles in games tend to have a higher amount of characters, with many games ranging from 50 to 70 characters per line and even higher, and some games using subtitles with three or more lines (Mangiron, 2013). While a considerable number of games include intralingual subtitles, this is not always the case, and usually there are no subtitles for sound effects, raising accessibility barriers for deaf and hard of hearing players.

Although accessibility to the media, in particular subtitling for the deaf and hard of hearing and audiodescription for the blind and visually impaired, has gained prominence in recent years, accessibility to video games remains a pending issue both for the game industry

[1] In the game industry, the term *voice-over* is used to refer to the recording of the script of a game by voice actors.

and academia. This situation may be partially attributed to the relative young nature of the game industry and its lack of standarisation. In addition, the interactive nature of games, which require a response from the player to the game stimuli, as well as the wide spectrum of users with different type of disabilities —sensorial (visual and hearing), motor, and cognitive— pose great challenges to the design of universally accessible games, as barriers found by users are different. For example, deaf and hard of hearing players are likely to miss any information that is only conveyed by means of audio, such as dialogue in a cinematic scene that is not subtitled, sound effects and music. Despite the fact that such players will still be able to play, they may have a "reduced gaming experience" due to the information loss (ibid., 5).

Providing good quality subtitles in games is crucial, due to their interactive nature. Players may need to perform an action in the game depending on the information in the subtitles, so readability is paramount to ensure comprehension and progress in the game. This is not only the case for deaf and hard of hearing players, but also for gamers who speak a different language, language learners, people with mild cognitive or learning impairments, people who like playing a game in the original game, people playing in noisy environments or any player who simply does not want to miss out on any information provided by the audio. The lack of standardisation of subtitling practices in the game industry, which in turn poses an accessibility barrier for all players, triggered this small exploratory research project. Its objective is to analyse the reception of game subtitles by both hearing and deaf users, combining subjective, qualitative data (a questionnaire) with objective, quantitative data (obtained with an eye tracker). The paper starts with a brief review of the use of eye tracking in AVT, followed by the description of the experiment and the obtained results. It concludes with a reflection on the findings and the limitations of the study.

The use of eye tracking technology in AVT

Eye tracking consists of measuring the movement of the eyes, usually in response to visual stimuli. The earliest eye trackers date from the late 1800s, although eye tracking research became more important as of 1950 and experienced an enormous growth since the 1990s (Holmqvist et al., 2011, p. 10-11). Eye tracking technology is used in different fields, such as psychology, cognitive linguistics, medical research marketing and advertising, to name but a few[2], in order to obtain information about how we perceive and process information from different stimuli. A number of psychologists have used eye tracking technology to study how subtitles are processed cognitively. As early as 1984, Belgian psychologist Géry d'Ydewalle carried out a study about the reception of subtitled films, focussing on the cognitive processes and the different factors involved in reading subtitles. In later studies d'Ydewalle and his team continued to explore the reception of subtitles, analysing issues such as: automatic reading behaviour when watching television (d'Ydewalle et al., 1991); incidental foreignlanguage acquisition by children watching subtitled television programmes (d'Ydewalle and Van de Poel, 1999); the effects of lengthening the presentation time of subtitles on children's reading time (Koolstra et al., 1999), and more recently a study about eye movements of children and adults while reading television subtitles (2007). Jensema et al. (2000) also performed tests with hearing and deaf users in order to analyse the impact of subtitle speed in the viewing experience and Jensema, Sarma Danturthi & Burch (2000) studied the amount of time spent viewing subtitles (2000). Bisson et al. (2012) used an eye tracker to analyse the processing of foreign and native language subtitles and the incidental acquisition of foreign language vocabulary.

---

[2] For a detailed account of the different applications of eye tracking research, see Duchowski (2007).

While the application of eye tracking for research in Psychology is well established, its use in in AVT for carrying out reception studies is more recent. The use of eye tracking technology has been introduced into AVT because it can provide researchers with objective, quantitative data that complement the information obtained by questionnaires, interviews and think aloud protocols (TAP), which provide qualitative and more subjective data. To date, eye tracking research has been mainly applied to the areas of subtitling and audiodescription (AD). For example, Caffrey (2008, 2012) analyzed the effects of experimental subtitling procedures on viewers perception of subtitled Japanese anime content. Perego (2010), working with an interdisciplinary team made of a psychologist and two computer engineers, studied the effectiveness of subtitle processing and concluded that the processing of subtitled films is cognitively effective, regardless of the quality of line segmentation. In 2012 Perego also edited a volume gathering several contributions on the use of eye-tracking in AVT, dealing with topics such as the impact of translation strategies on subtitle reading (Ghia, 2012); the effect of linguistic variation on subtitle reception (Moran, 2012), and the use of live subtitles on portable devices (Miquel-Iriarte et al., 2012). Künzli and Ehrensberger-Dow (2011) carried out an experimental study aimed at comparing the reception of movies subtitled with standard subtitles versus innovative subtitles and found no significant difference in terms of reception. Kruger et al. (2013) used eye tracking technology, as well as electroencephalography, self-reported ratings of participants, and comprehension tests in order to measure the impact of subtitles on viewers' cognitive load. In another study, Kruger and Steyn (2013) state that reading subtitles has a positive impact on academic performance, after carrying out eye-tracking and comprehension tests to a group of students. Under the EU funded project DTV4ALL[3], which ran from 2008 to 2011 with the aim of facilitating the

[3] For more information about the project, see the project website http://www.psp-dtv4all.org/.

provision of access services on digital television across the European Union, tests were carried out with eye tracking technology, in order to analyze the reception of SDH, —see, for example, Arnáiz (2008 & forthcoming) and Szarkowska et al. (2013) for SDH subtitles and Romero-Fresco (2009, 2010, 2012) for respoken subtitles.

The DTV4ALL project also focused on audiodescription. Tests with eye tracking with sighted users were carried out in order to establish what information is perceived and is more important for sighted users, with a view to determine what should be described, and thus, improve accessibility for blind and visually impaired users. For example, Vilaró et al. (2012) tested the effect of varying audio stimuli on visual attention distribution, while Orero and Vilaró (2012) analysed the perception of minor details in films. There is also a dedicated issue in the journal *Perspectives* (Mazur and Kruger, 2012) presenting the results of the Pear Tree Project, a subproject within the DTV4ALL project, the objective of which was to develop common European guidelines on AD. Other authors have also explored the use of eye tracking technology for research in AD, such as Igareda and Maiche (2009), who looked at the description of emotions. More recently, Di Giovanni (2014) confirmed the applicability of eye tracking research for AD through an experimental study, in which blind participants were offered two kinds of AD: traditionally drafted AD and AD made after determining with the eye tracker the visual priorities of sighted users. The DTV4ALL project has been followed by the EU funded Hbb4ALL, which started in 2014 with the aim of making hybrid broadband broadcast TV accessible for all users by customizing accessibility services through options for personal preferences[4].

Eye tracking technology has also been applied to Game Studies in order to perform usability studies —see for example Alkan and Cagiltay (2007) regarding how novice players

[4] For more information about the project, see the project website http://www.hbb4all.eu/.

learn how to play a computer game— and to use eye tracking as an alternative input interaction device —see, for example, Kaufman et. al. (1993); Smith and Graham (2006); Nacke et al. (2009). Eye tracking is also used to measure players' gameplay experience, often in conjunction with psychophysiological methods, such as electroencephalography, electrocardiography, galvanic skin —see, for example, Nacke and Lindley (2008); Anders et al. (2010); Nacke et al. (2011). Undoubtedly, eye tracking is a useful tool for reception studies and can also be applied to the reception of subtitles in games, similarly to what has been done in AVT. However, no studies to date had been carried out regarding the reception of subtitled content in a video game, a fact that triggered this exploratory study.

Experiment

The test consisted of a reception study of the subtitles and the depiction of sound effects in the demo *Hounted House*, purposedly developed by [removed for peer review][5] for this research. The test combined qualitative and quantitative methodologies: a questionnaire —to establish which type of subtitles and method for depiction of sound effects and character identification participants preferred— and eye tracking technology for recording the eye movemement of the participants and obtain information about their visual attention and reading speed. The parameters that were tested in the experiment were:

- Subtitle presentation: a) in a text box; b) in a speech box pointing to the mouth of the character speaking; c) directly onto the screen.
- Subtitle position: centered or aligned to the left
- 1 line subtitles versus 2 line subtitles

[5] [removed for peer review]

- Efficiency of the use of speaker portraits next to the subtitles
- Depiction of sound effects: a) with a pop-up star transcribing the sound; b) a blue text box on the top right hand side of the screen

Participants

The subjects of the experiment consisted of hearing and deaf users: 12 hearing users and 13 deaf users: 11 prelocutive and 2 postlocutive deaf participants[6]. The difficulty of gathering subjects to take part in the experiment should be highlighted, especially in the case of deaf participants. In order to seek for volunteers contact was made with the main associations for deaf people in Barcelona and with a sign language interpreter who helped recruit volunteers. Before the test, a small pilot was conducted with two hearing users to validate the experiment design. The results of their test are included in the hearing group. It should be mentioned that originally 30 subjects were tested, 15 for each group, but due to the poor quality of the recording by the eye tracker for five users, they could not be included in the eye tracking analysis and had to be discarded. The average age of participants of the hearing group was 24.6 and for the deaf group 22. In the hearing group; 66.6% of participants were female and 33.3% male, while in the deaf group the ratio was inverted, with 61.5% male participants and 38.5% female. Regarding their gaming habits, in the hearing group, 16.6% of respondents were regular players, 33.3% played occasionally, 16.6% played seldom, and 33.3% never played. In the deaf group there was a higher proportion of gamers, with 30.5% regular gamers and 46% of respondents who played occasionally. The remaining 23.3% seldom played. Participants from both groups were mostly students. A more balanced distribution of gender for both groups would have been preferable, as it was observed that all males were regular players or played often, while females played seldom or never. However, this was not

[6] [removed for peer review]

possible due to the difficulty obtaining volunteers, but it is a factor that should be taken into consideration in future studies.

Material

For the purpose of the experiment, it was necessary to be able to manipulate the subtitles and the depiction of sound effects, in order to be able to present different alternatives to users and compare the results. However, as games are a software programme, in which the subtitles are embedded, access to the subtitles of an existing game would not have been possible without performing what is known as *rom-hacking*, that is, manipulating the original game code, which is considered an illegal practice. For this reason, collaboration was established with the [removed for peer review]. As their final project, a group of four students in this course developed a demo called *Casa Encantada* ("Haunted House"), under the supervision of lecturers, who are also professionals working in the industry. It is a 10 minute demo in which there are two cinematic scenes of approximately two minutes each. There are three characters: the House (the player) and a young couple who have just bought the house, David and Emma. The objective of the game is to scare the couple and make them leave the house, without scaring them to death, which results in game over. The game is voiced-over in Spanish and Catalan and includes intralingual subtitles, which are verbatim, as this is the convention in the game industry, despite the fact that it contravenes the established practice of condensation in standard subtitling. In the settings screen, the player can choose the language. For the purpose of the study, the first cinematic scene, of approximately 2 minutes duration, was chosen in the Spanish language, with Spanish subtitles.

There are 17 subtitles, 8 subtitles for the house, 5 for David and 3 for Emma. The maximum of character in one line was 56, which is common in game subtitles but not in standard TV subtitling, which uses between 37 and 40 character per line. Subtitles for each character were given different format, in order to check users' preferences. All the sounds in the audio track were depicted, such as the sound of footsteps or the shouts of the characters when they were scared. In order to test the different parameters, two conditions were designed, as explained in the next section.

Design and procedure

Before starting the experiment, participants were given simple and clear instructions about what they had to do, and according to the established protocol when doing research with human subjects, they were asked to sign a consent form and were told they could leave the experiment at any time. The tests were carried out individually. Participants were not told about the objectives of the study, to prevent them paying too much attention to the subtitles, following previous reception studies about reading subtitles (see d'Ydewalle et al., 1991). The questionnaire included questions about personal data (age, sex, hearing capacity), gaming habits, their opinion about current subtitling practices in games, and whether they activate subtitles in games when they play.

After the first set of questions, participants were asked to sit before the eye tracker monitor, at a distance of approximately 65 cm. Calibration of their eye movements was then performed. A T-60 Tobii eye-tracker was used, with embedded cameras and without physical contact with the participant. After successful calibration, users played the game for approximately 5 minutes. This involved viewing a two minute cinematic scene, followed by some free play, for about three more minutes. Then they were asked to answer more questions about the subtitles and sound effects in the game in order to find out about their preferences. While participants were playing the game, an eye tracker was recording their eye movements. This was done in order to find out the more efficient way of presenting subtitles and sound effects, based not only on users' preferences, but also analysing objectively their viewing and reading behaviour. Due to the complexity of analysing the raw data provided by the eye-tracker, it was decided to use only one variable, "fixation lenght", which provides information about the amount of time participants spent looking at a particular area in the screen. The data obtained with the eye tracker was used to analyse the following parameters:

- Participants' reading speed.
- Whether it took longer to read a long one line subtitle or two shorter line subtitles.
- The most effective way to depict sound effects.
- The effectiveness of a speaker portrait for character identification.

In order to be able to test the reception of the different parameters, two different conditions were designed for the eye tracking exploratory study. They were applied to half the group of hearing participants and half the group of deaf participants. The conditions were the following:

Table 1. Different conditions for the experiment

Parameters	Condition 1	Condition 2
Character identification	Label with character's name	Label with character's name & speaker portrait
Depiction of sound effects	Pop-up star with onomatopoeya	Blue text box on top right corner of screen describing the sound
1 vs. 2 line subtitles	House: 1 line David: 2 lines Emma: 1 line	House: 2 lines David: 1 line Emma: 2 lines

For both conditions, the font, size, colour, alignment, and presentation of the subtitles of each character remained unaltered, and the following parameters were used: **Table 2.** Fixed parameters for subtitle presentation

Parameter	House subtitles	David's subtitles	Emma's subtitles
Font	Titania	Arial	Slap
Size	22 points	32 points	32 points

Color	Black	White	Black
Alignment	Center	Left	Left
Presentation	Light brown text box with artistic design on the corners	Directly on the screen	White speech text box

Next, there are some screenshots of the different format of subtitles, followed by screenshots depicting the two different types of sound effects deployed in the game.

Figure 1. Different types of subtitles for condition 1

Figure 2. Different types of subtitles for condition 2

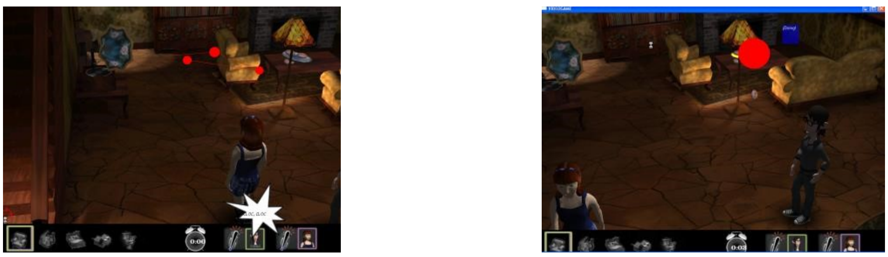

Figure 3. Depiction of sound effects in condition 1 (left) and 2 (right)

Results

In this section, the results of the exploratory study are presented. First, there is a summary of the data gathered from the questionnaire, followed by the data obtained with the eye tracker.

Data gathered from the questionnaire

The data compiled from the questionnaire is presented according to the following topics:
General opinion about subtitles in video games

75% of hearing participants did not activate the subtitles when playing games, while 19% did and 6% did not answer the question. On the other hand, 86% of deaf players thought that subtitles were very important, because they allow them to understand the game, the plot and what the characters say. One deaf user also said that sound effects should also be subtitled in all games. Interestingly, 14% of deaf users thought it was not necessary to activate subtitles. This can be explained due to the fact that they stated they mainly play platform and simulation games, which do not usually contain much voiced-over dialogue.

Opinion about the subtitles in the game Casa Encantada

This section gathers the information about the subtitles in the game used for the test, according to the answers from the questionnaires.

Preferred subtitles

The first question users had to answer after playing the game was what type of subtitle they liked best overall. There was a clear difference between the hearing and the deaf group, as illustrated in the graphics below.

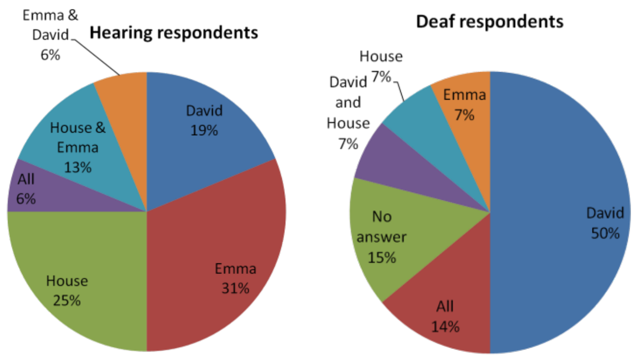

Figure 4. Preferred subtitles

As can be seen, 31% of hearing participants preferred Emma's subtitles, with a black Slap font, 32 points, on a white speech box, followed closely by the House, with 25%, the subtitles of which were a black Titania font, 22 points, on a brown box with ornaments. David's subtitles, Arial, 32 points, white, projected directly onto the screen, were third with 19%. Some users indicated two types of subtitles, and if we add those percentages to each category, Emma's subtitles would reach 50%. However, deaf users preferred David's subtitles, with 50%, plus an extra 7% who preferred his subtitles and the House's subtitles best. They claimed that this subtitles interfered less with the image. This is the most common presentation format for subtitles in TV and DVD, so they are probably used to this format. It is also interesting to note that 14% percent of deaf users did not answer this question, which may mean that they did not feel the differences amongst the different subtitles mattered.

Preference between one or two line subtitles

In this case, the majority of participants in both groups indicated a preference for two-line subtitles as opposed to long one-liners (88% of users in the hearing group and 64% in the deaf group).

Alignment

Again, both groups of users were in agreement and preferred centred subtitles (88% of participants of the hearing group and 86% of deaf users).

Subtitle presentation

For this parameter a major variation can be detected amongst participants, due to the different aesthetic preferences and habits regarding the use of subtitles. In the hearing group, 50% of respondents stated they preferred subtitles to be presented in a box, while 44% preferred them to be projected directly onto the screen and 6% opted for the speech bubble. On the other hand, 50% of deaf users preferred subtitles to be displayed directly onto the screen, while 29% preferred a text box and 21% a speech bubble. These data are aligned with those obtained in question 1, about user's overall preference for subtitles. It seems that deaf users prefer subtitles to be projected directly onto the screen, as it interferes less with the image.

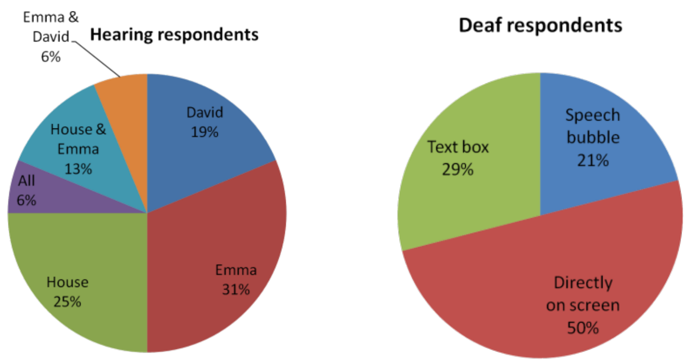

Figure 5. Preferred subtitle presentation

Character identification

This is another parameter where participants' answers differed considerably, as illustrated in the graphics below. Hearing respondents favoured equally the use of an avatar or speaker portrait and the use of a name tag (both 31%) for character identification, followed by the use of colours (25%), and the use of name and colour (13%). Deaf participants preferred the use of colours (36%), followed by the use of name tags and colours (29%). 21% preferred speaker portraits and 14% the use of a label with the name of the character. The use of different colours to identify characters is common in SDH for TV and is thus the method for character identification to which deaf users are more habituated.

It is interesting to notice that despite the fact that after playing the game an important number of participants indicated that the speaker portrait was their preferred choice for character identification, the eye tracking results show that most of them did not notice the speaker portrait on the screen, as will be explained in more detail later.

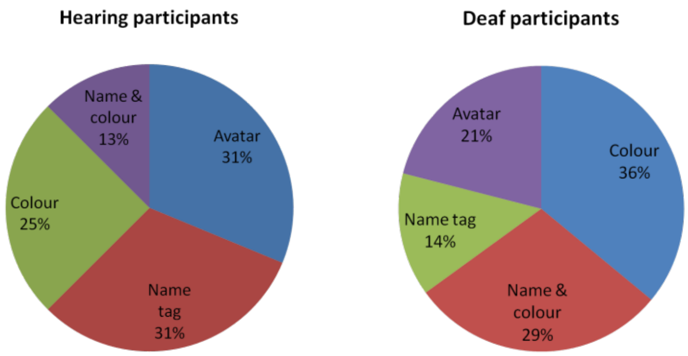

Figure 6. Preferred method for character identification

Depiction of sound effects

Both groups expressed a preference for comic style pop-up onomatopoeia in order to depict sound effects, with 69% of hearing users and 67% of deaf users choosing this method. The eye tracking data confirms that comic style pop-up sound effects were perceived by most users, as will be explained in more detail later. It is also interesting to point out that 12% of hearing users stated that they preferred for sound effects not to be depicted, which indicates lack of awareness about accessibility issues. On the other hand, all deaf users agreed they should be depicted.

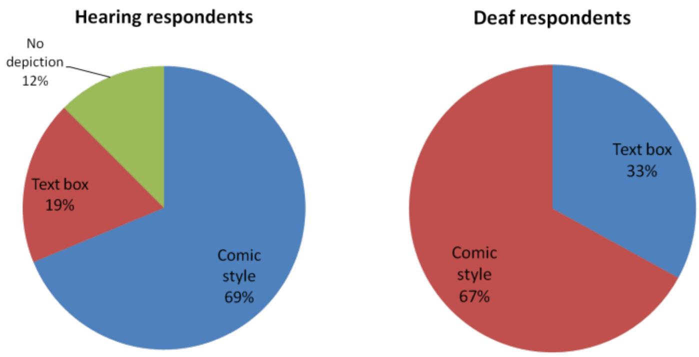

Figure 7. Preferred method for depiction of sound effects

Additional comments

In the additional comments section, 4 hearing participants (25%) commented on the lack of synchrony between the game graphics, the audio, and the subtitles, since in occasions a character's mouth keeps moving and the subtitle remains on the screen, although the character is no longer speaking.

Data obtained with the eye tracker

Given the subjective component inherent to qualitative reception analysis, one of the ideas behind this exploratory study was to contrast the qualitative data obtained from the questionnaire with objective, quantitative data gathered using eye tracking technology. It should be noted, however, that analysing video games reception with eye-tracker is complicated due to the dynamic and interactive nature of games and the fact that a game is a type of software that needs a powerful computer to run it. Due to the complexity of analysing the reception of dynamic objects with the eye-tracker, it was decided to focus the study on one single variable, *fixation length*, which allowed to obtain data about the time participants spent reading subtitles or looking at other elements in the screen. This variable also allowed to check empirically whether there was any difference regarding the time it takes to read a long one line subtitle as opposed to a long subtitle split in two lines. Finally, the eye tracking data also allowed verifying which method was more effective to depict sound effects and to establish whether participants had looked at the speaker portrait, as it recorded the areas in the screen participants had looked at, marked with red dots, as can be seen in the screenshot below:

Figure 8. Raw data showing where a participant was looking

In order to be able to analyse the recordings of the different participants, each recording had to be segmented in different scenes, one per subtitle. Usually the eye tracking software allows segmenting recordings into scenes automatically, but since this was a game, i. e., a software programme, all recordings had to be segmented manually. Once the different scenes have been created, the eye tracker can provide different types of raw data, such as gaze plots, displaying all the different points in the screen participants looked at, as well as heat maps showing the points in the screen to which participants paid more attention, i. e. spent more time looking.

Figure 9. Exemple of a gaze plot (left) and the corresponding heat map (right)

In order to gather numeric data regarding subtitle reading speed —obtained from the amount of time participants fixed their gaze in the subtitles— areas of interest (AOI) were manually created for each scene, containing one subtitle. In the case of the cinematic scene, for condition 1, two AOI were created: the main scene and the subtitles. For condition 2, which included a speaker portrait, three AOI were established: the scene, the speaker portrait and the subtitle, as shown in the image below.

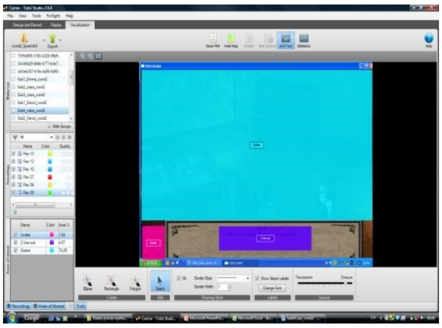

Figure 10. AOI used to analyse condition 2, with a speaker portrait

For the scenes of free game play, where participants could click on objects and interact with them, which produced sound effects that were described, different areas were set in relation to the areas where the sound effects description appeared on screen.

From the AOI, the eye-tracker analysis software provides information about different variables, albeit for this study only *fixation length* was selected, as explained previously. Data can be obtained in a graphic or a numeric format. The eye-tracker also lists the time participants spent looking at places that were outside the defined AOI. **Table 3.** Fixation length per second

	Not on AO	Avatar	2 line sub	Scene
Rec 05	0	0	3.778513	1.915365
Rec 06	0.599475	0	1.399029	3.177205
Rec 07	0.049994	0	1.782116	3.909946
Rec 10	0.71381	0	2.698192	2.641874
Rec 12	1.98143	0	1.848576	1.99846
Rec 13	0.568225	0	2.131804	3.543661
Min	0	0	1.399029	1.915365
Max	1.98143	0	3.778513	3.909946
Sum	3.912934	0	13.63823	17.18651
Mean	0.652156	0	2.273038	2.864419
Median	0.58385	0	1.99019	2.90954
Stddev	0.716728	0	0.85466	0.818834

Next, a summary of the main data obtained with the eye-tracker is provided, regarding the following parameters: a) subtitle reading speed; b) reading speed for long one-line subtitles and their equivalent in two lines; c) perception of sound effects, and d) speaker portrait perception.

Subtitle reading speed

For all the scenes, the heat maps show that the places where participants looked at longer were subtitles. This confirms previous studies stating that when there are subtitles available on the screen, viewers tend to read them, even if they understand the audio track (see, for example, d'Ydewalle et al., 1991). Secondly, participants focussed on characters' faces, another fact proven by several eye tracking studies (eg. Birmingham, Bischof and Kinstone, 2008). For the hearing group, the average reading speed was 16.4 characters per second, while for the deaf group reading speed was slightly slower, with 11 characters per second. This can be accounted to the fact that deaf users usually read at a slower rate (see for example, Báez & Fernández, 2010). In addition, hearing users do not need to look at subtitles in so much detail, as the information they provide is redundant to that provided in the audio. Deaf users, however, need to read the subtitle in order to understand what the characters are saying.

The following table presents some illustrative examples of the reading speed of hearing and deaf participants. Since there were no significant differences between participants of the same group with different conditions, the table just distinguishes between hearing and deaf users.

Table 4. Examples of average reading speed for both groups

Participants	Characters	Reading speed in seconds
Hearing	40	3.2
Deaf	40	3.5
Hearing	55	3.3
Deaf	55	5.4
Hearing	83	4.4
Deaf	83	8.5

From the table it can be appreciated that the longer the subtitle, the bigger the difference between the reading speed of both groups, as it took deaf participants more time to read long subtitles. That said, in general the reading speed for both groups is higher than the recommended speed for television subtitles, which varies between 37 and 40 characters in six seconds. Therefore, more research would be useful to confirm whether reading speed for game subtitles projected on different screens such as a PC is faster (perhaps because the user is sitting closer to the display screen) or if reading speed in general has increased for all kind of users, which would mean that current subtitling standards for TV, DVD and cinema should be reviewed and updated to accommodate the new reality.

Reading speed for long one-line subtitles and their equivalent in two lines

88% of hearing participants and 64% of deaf participants stated in the questionnaire that they preferred long subtitles to be split over two lines. With the eye-tracker, however, there is no significant difference between reading speed for the same subtitle when it is in one line or in two lines for the two different conditions.

Perception of sound effects

Comic-style pop-up onomatopoeia, such as *Argh!*, were more effective to describe sound effects. In the free play element of the test, where users had to move around the house and interact with objects clicking in them, pop-up sound effects were perceived by all users in both groups in condition 1, with a 100% perception rate. On the other hand, the blue text box describing the sound effect placed on the top right part of the screen, the method commonly used for subtitling sound effects in Spanish television, was less effective. Only 42% of participants in the hearing group perceived it, while 50% of participants in the deaf group noticed it. Given the interactive nature of games, which require user's attention in order to be able to respond to all the different game stimuli, it seems appropriate to depict sound effects in comic style, provided that this method fits with the genre and aesthetics of the game. A number of games already use this method, such as *Freedom Force vs. the Third Reich* (2005) and *XIII* (2003) (Van Tol, 2006), which fits in with the visually rich and playful nature of video games.

Figure 11. Depiction of sound effects with a pop-up onomatopoeia (left) and a text box (right)

Perception of speaker portraits

In the questionnaire, 31% of respondents in the hearing group stated that a speaker portrait was their preferred method for identifying characters, as did 21% of deaf participants. In the test, speaker portraits were only included in condition 2, but the data recorded by the eyetracker show that they were not very effective, as in most cases participants did not see them. Speaker portraits were perceived by hearing users in a proportion of almost 7%, calculated taking into account the amount of scenes with a portrait (17) and the amount of users who perceived it in each scene. Interestingly, deaf participants perceived the speaker portraits more often than hearing users, amounting to almost 22%. This seems to indicate that deaf participants perceived the scenes more globally than hearing users and paid more attention to different details, although further research with larger scale studies would be necessary to confirm this hypothesis.

Summary of the data obtained with the eye tracker

To conclude this section, the main findings are summarised in relation to the parameters tested in the experiment:

- Subtitle presentation: there was a difference between the hearing and the deaf group, as the former preferred mainly subtitles in a black font on a text box and the latter a white font projected directly onto the screen. Since games provide the option of setting up different parameters, the possibility of including different format for the subtitles would be desirable if possible.
- Subtitle position: most users preferred centred subtitles.
- 1 line subtitles versus 2 line subtitles: most users indicated a preference for shorter two line subtitles, although data provided by the eye-tracker do not show a significant difference in terms of reading speed.

- Efficiency of the use of speaker portraits: in this particular occasion speaker portraits did not prove efficient, as they were not perceived. It would be interesting to perform more research in this topic. Also, the fact that all subtitles included the name of the character speaking made the portraits redundant, and therefore not needed. Further tests should include either the label with the name of the character or the portrait, to be able to assess more accurately both methods.
- Depiction of sound effects: pop-up onomatopeya style depictions were more effective and seem to be more appropriate for video games, given their more ludologic nature, provided that they suit the game genere and style.

Discussion

This exploratory study combined qualitative data obtained by means of questionnaires with quantitative data provided by an eye tracker. In terms of the sample, a more balanced distribution between genders, as well as gaming preferences, would have been desirable, but not possible in this occasion due to the difficulty of recruiting volunteers to take part in the subject, particularly within the deaf community. In addition, despite having carried out calibration for all 30 participants, the quality of the data obtained with the eye tracker was not good enough for 5 users, and their data had to be discarded, despite the fact that calibration had been performed correctly and clear instructions had been given to subjects about body posture during the experiment. A portable, not fixed eye tracker may have been more appropriate, as well as a more powerful computer, as the game took some time to start.

Moreover, the fact that the stimuli was a demo developed by master students, as opposed to a mainstream video game developed by professionals, also had an impact on the study, as there were synchrony issues with the subtitles, the audio, and the movements of the character's mouths, which hearing users noticed and commented on. A collaboration agreement with an industry developer which granted researchers access to the code of a game and allowed them to manipulate different subtitling parameters would be most beneficial for this kind of test.

It should also be noted that the information compiled from the questionnaires is subjective and some inconsistencies were noticed between participants' answers and the eye tracker data in the case hearing users regarding speaker portraits. While 21% stated that speaker portraits were their preferred method for character identification, only 7% noticed them. Also, it was detected that participants in the hearing group chose as their preferred subtitle the subtitle belonging to the character they liked best in the game due to the interpretation of the voice over actor, because they included in their answers comments such as "I prefer David's subtitles because his voice is beautiful and soothing" or "I prefer the House's subtitles; both her voice and the presentation of the subtitles are funny" (author's translation). This issue was not detected for deaf users, who judged subtitles only by their format.

Finally, as this was a small scale exploratory study, the data obtained are not sufficient to establish subtitling norms for video game. Nevertheless, some interesting data about user's preferences and the effectivity of certain parameters were obtained that should be further explored in the future. Issues that should be further analysed are the reading speed in relation to the lenght of subtitles; the difference in reading speed between hearing and deaf participants in longer subtitles; the use of comic style onomatopoeia for depicting sound effects, and the global perception of scenes and subtitles by hearing and deaf users. In addition, in this study it was not checked whether participants went back and read the subtitles again, which may have also provided useful information. Comprehension was not measured either, but in future studies it would be interesting to test whether participants could perform certain tasks in the game based on the information provided in the subtitles, as video game users have to respond to several game stimuli to be able to progress in a game.

Conclusions

The main objective of this exploratory pilot study was to analyse the reception of game subtitles, an area that has been to date unexplored in academia and that can help contribute to standardise subtitling practices in games, which in turn, would improve accessibility for all players. In order to be able to gather both qualitative and quantitative data, both questionnaires and eye tracking technology were used, with a reduced sample of participants. Larger scale reception studies with more participants and using mainstream games would be necessary to obtain more information on the quality of game subtitles and their reception, with a view to establishing industry wide standards and improve the gaming experience of both hearing and deaf and hard of hearing players.

References

Alkan, S., & Cagiltay, K. (2007). Studying computer game learning experience through eye tracking. *British Journal of Educational Technology*, 38(3), 538-542.

Arnáiz, V. (2008) La Objetividad en el Subtitulado. Justificación de los Parámetros Formales Mediante EyeTracking. In: Á. Pérez-Ugena, & R. Vizcaíno (Eds.). *ULISES: Hacia el desarrollo de tecnologías comunicativas para la igualdad de oportunidades. Retos y perspectivas para sordos signantes* (pp. 73-85). Madrid: Observatorio de las Realidades Sociales y de la Comunicación.

Arnáiz, V. (forthcoming) Viewers' Perception of SDH in Spain: An Eye-Tracking Study. In P. Romero-Fresco (Ed.) *The Reception of Subtitles for the Deaf and Hard-of-Hearing* in Europe, Bern: Peter Lang.

Báez, I. & Fernández, A. M. (2010). Spanish deaf people as recipients of close captioning. In Matamala, Anna & Orero, Pilar. *Listening to subtitles: Subtitling for the deaf and hard of hearing* (pp. 25-44). Bern: Peter Lang.

Birmingham, E., Bischof, W.F. & Kinstone, A. (2008). Social Attention and Real–World Scenes: The Roles of Action, Competition and Social Content. *The Quarterly Journal of Experimental Psychology*, 61, 986–998.

Bisson, M. J., Van heuven, W. J. B., Conklin, K. & Tunney, R. J. (2012). Processing of native and foreign language subtitles in films: An eye tracking study. *Applied Psycholinguistics*, FirstView Article pp 1-20, available on CJO2012. doi:10.1017/S0142716412000434. Retrieved from http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=8722810

Caffrey, C. (2008). Viewer perception of visual nonverbal cues in subtitled TV Anime. *European Journal of English Studies,* 12 (2), 163-178.

Caffrey, C. (2012). Using an eye–tracking tool to measurethe effects of experimental subtitling procedures on viewer perception of subtitled AV content. In E. Perego (Ed.). *Eye tracking in audiovisual translation* (pp. 223-258). Roma: Aracne.

Di Giovanni, E. (2014) Visual and narrative priorities of the blind and non-blind: eye tracking and audio description. *Perspectives: Studies in Translatology*, 22(1), 136-153.

Drachen, A., Nacke, L., Yannakakis, G. & Pedersen, A. L. (2010). Psychophysiological Correlations with Gameplay Experience Dimensions. In *CHI 2010 Workshop: Brain, Body, and Bytes* (pp. 1-4). Retrieved from http://arxiv.org/abs/1004.0243.

Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice*. 2nd edition. London: Springer.

d'Ydewalle, G. (1984). Processing TV information and eye movements research. In Veer, G.C. van der, Tauber, M.J., Green, T.R.G. & P. Gorny. (Eds.). *Readings on Cognitive Ergonomics: Mind and Computers* (pp. 200-204). Heidelberg: Springer Berlin. d'Ydewalle, G., Praet, C., Verfaillie, K. & Van Rensbergen, J. (1991). Watching subtitled television - automatic reading behavior. *Communication research*, 18(5), 650-666. d'Ydewalle, G. & Van de Poel, M. (1999). Incidental foreign-language acquisition by children watching subtitled television programs. *Journal of Psycholinguistic Research*, 28, 227-244.

d'Ydewalle, G. & De Bruycker, W. (2007). Eye movements of children and adults while reading television subtitles. *European Psychologist: the Journal for Psychology in Europe*, 12, 196-205.

Ghia, E. (2012). The impact of translation strategies on subtitle reading. In E. Perego (Ed.). *Eye tracking in audiovisual translation* (pp. 157-182). Roma: Aracne.

Holmqvist, K., Nyström, M., Andersson, R.; Dewhurst, R.; Halszka, J. & van de Weijer, J.(2011). *Eye Tracking: A comprehensive guide to methods and measures*. Oxford & New York: Oxford University Press.

Igareda, P.; Maiche, A. (2009). Audio Description of Emotions in Films using Eye tracking. In *Proceedings of the Symposium on Mental States, Proceedings of the Symposium on Mental States, Emotions and their Embodiment* (pp. 20-23). Edinburgh: The Society for the Study of Artificial Intelligence and Simulation of Behaviour.

Jensema, C. J. (1998). Viewer reaction to different television captioning speeds. *American Annals of the Deaf.* 143(4), 318-324.

Jensema, C. J., El Sharkawy, S., Sarma Danturthi, R., Burch, R. & Hsu, D. (2000). Eye movement patterns of captioned television viewers. *American Annals of the Deaf.* 145 (3), 275-285.

Jensema, C.J., Sarma Danturthi, R. & Burch, R. (2000). Time spent viewing captions on television programs. *American Annals of the Deaf.* 145 (5), 464-468.

Kaufman, A. E., Bandopadhay, A., & Shaviv, B. D. (1993). An eye tracking computer user interface. In *Proceedings IEEE 1993 Symposium on Research Frontiers in Virtual Reality* (pp. 120-121). Retrieved from labs.cs.sunysb.edu/labs/projects/eye/Papers/short.paper.ps

Koolstra, C.,Van der Voort, T. & d'Ydewalle, G. (1999). Lengthening the presentation time of subtitles on television: Effects on children's reading time and recognition. *Communications*, 24, 407-422.

Kruger, J.-L., Hefer, E. & Matthew, G. (2013). Measuring the impact of subtitles on cognitive load: eye tracking and dynamic audiovisual texts. *Proceedings of Eye Tracking South Africa*, pp. 29-31 August 2013, Cape Town. Retrieved from http://dl.acm.org/citation.cfm?id=2509331

Kruger, J.-L. & Steyn, F. (2013). Subtitles and Eye Tracking: Reading and Performance. *Reading Research Quarterly*, 49(1), 105–120. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/rrq.59/pdf

Künzli, A.; Ehrensberger-Dow, M. (2011). Innovative subtitling: A reception study. In C. Alvstad, A. Hild & E. Tiselius (Eds.). *Methods and Strategies of Process Research: Integrative approaches in Translation Studies* (pp. 187–200). Amsterdam & Philadelphia: John Benjamins.

Mangiron, C. (2012). Exploring new paths towards game accessibility. In A. Remael, P. Orero & M. Carroll (Eds.). *Audiovisual Translation and Media Accessibility at the Crossroads: Media for All 3* (pp. 43-59). Amsterdam & New York: Rodopi.

Mangiron, C. (2013). Subtitling in Game Localisation: A Descriptive Study. *Perspectives: Studies in Translatology*, 21(1), 42-56.

Mazur, I. & Kruger, J.-L. (Eds.) (2012). Pear Stories and Audio Description: Language, Perception and Cognition across Cultures. *Perspectives: Studies in Translatology*, 20(1).

Miquel–Iriarte, M., Vilaró, A., Orero, P., Serrano, J. & Delgado, H. (2012). Entitling. A way forward for accessibility. In E. Perego (Ed.). *Eye tracking in audiovisual translation* (pp. 259-276). Roma: Aracne.

Moran, S. (2012). The effect of linguistic variation on subtitle reception. In E. Perego (Ed.). *Eye tracking in audiovisual translation* (pp. 183-222). Roma: Aracne.

Nacke, L. & Lindley, C. (2008). Flow and immersion in first-person shooters: measuring the player's gameplay experience. In *Proceedings of the 2008 Conference on Future Play: Research, Play, Share* (Future Play '08), (pp. 81-88). New York: ACM.

Nacke, L., Stellmach, S., Sasse, D. & Lindley, C. A. (2009). Gameplay experience in a gaze interaction game. In *Proceedings of The 5th Conference on Communication by Gaze Interaction – COGAIN* (pp. 49-54) Retrieved from. http://arxiv.org/ftp/arxiv/papers/1004/1004.0259.pdf

Nacke, L., Kalyn, M., Lough, C. & Mandryk, R. L. (2011). Biofeedback game design: using direct and indirect physiological control to enhance game interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)* (pp. 103-112). New York: ACM. Retrieved from http://doi.acm.org/10.1145/1978942.1978958

Orero, P. & Vilaró, A. (2012) Eye tracking analysis of minor details in films for audio description. In R. Agost, E. Di Giovanni & P. Orero (Eds.), *Multidisciplinarity in audiovisual translation. MonTi*, 4, 295-319.

Perego, E. (ed.) (2012). *Eye tracking in audiovisual translation*. Roma: Aracne.

Perego, E., Del Missier, F., Porta, M. & Mosconi, M. (2010). The Cognitive Effectiveness of Subtitle Processing. *Media Psychology*, 13(3), 243-272.

Romero-Fresco, P. (2009). More haste less speed: Edited versus verbatim respoken subtitles. *Vigo International Journal of Applied Linguistics*, 6, 109-133. Retrieved from http://webs.uvigo.es/vialjournal/pdf/Vial-2009-Article6.pdf

Romero-Fresco, P. (2010). Standing on quicksand: hearing viewers' comprehension and reading patterns of respoken subtitles for the news. *Approaches to Translation Studies*, 32, 175-194.

Romero-Fresco, P. (2012). Quality in Live Subtitling: The Reception of Respoken Subtitles in the UK. *Approaches to Translation Studies*. 2012, 36, 111-131.

Smith, J. D., Graham, T. C. N. (2006). Use of eye movements for video game control. In *Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology* (ACE '06) (pp. 20-28). New York: ACM. Retrieved from http://www.cms.livjm.ac.uk/Library/AAA-GAMES-Conferences/ACM-ACE/ACE2006/Papers/S31_a86.pdf

Szarkowska, A., Krejtz, I., Krejtz, K. & Duchowski, A. (2013). Harnessing the potential of eyetracking for media accesibility. In S. Grucza, M. Pluzyczka, J. Zajac. (Eds). *Translation Studies and Eye-Tracking Analysis* (p. 153-183). Frankfurt et al.: Peter Lang.

Van Tol, R. (2006). The Sound Alternative. *Game Accessibility*. Retrieved from http://www.accessibility.nl/games/index.php?pagefile=soundalternative

Vilaró, A., Duchowski, A., Orero, P., Grindinger, T., Tetreault S., & Di Giovanni, E. (2012). How sound is The Pear Tree? Testing the effect of varying audio stimuli on visual attention distribution. *Perspectives. Studies in Translatology*, 20, 55-65.

Wieczorek, A. M., Kłyszejko, Z., Szarkowska, A., & Krejtz, I. (2011). Accessibility of subtitling for the hearing-impaired. In *Proceedings of the Conference: Interfejs użytkownika - Kansei w praktyce* (pp. 27–33). Warsaw: Wydawnictwo PJWSTK. Retrieved from https://repin.pjwstk.edu.pl/xmlui/handle/186319/106