Could integrated subtitles benefit young viewers? Children's reception of standard and integrated subtitles: a mixed methods approach using eye tracking

Sharon Black (sharon.black@uea.ac.uk)

University of East Anglia, Norwich Research Park, Norwich, UK. NR4 7TJ

Abstract

This paper investigates children's reception of AV content with standard and integrated interlingual subtitles. To this end, an experimental study was conducted with 17 children aged 8-9 years, to assess their reception of clips of an animated film in Spanish with standard and integrated subtitles in English. A mixed methods approach was adopted with the aim of obtaining a robust, comprehensive understanding of the children's reception of the subtitled AV content, using eye tracking, scene recognition tests, content comprehension tests, questionnaires, and interviews.

It was established that the children spent a significantly larger proportion of their viewing time looking at the images and also fixated more times on the images when the subtitles were in the integrated position. However, the hypothesis that participants would exert lower levels of cognitive effort when watching the AV clips with integrated subtitles was only partially confirmed. The integrated subtitles did not have detrimental effects on their viewing patterns, scene recognition performance, or comprehension of the clips. The majority approved highly of both subtitle positions and perceived both as easy to read and understand. These findings show that further research in this area is warranted.

Keywords: integrated subtitles, eye tracking, cognitive processing, children

1. Introduction

There is considerable evidence that watching subtitled audiovisual (AV) content can aid children's acquisition of other languages, enhance their literacy skills, and foster intercultural awareness (see, for example Koolstra and Beentjes, 1999; Linebarger, 2001; Bravo, 2008;

Kothari and Bandyopadhyay, 2014; and for extensive reviews of the existing research, see Gernsbacher, 2015; Black, 2017). Indeed, there is a need to raise awareness among broadcasters, video on demand providers, educators and parents of the value in using subtitles as a simple and motivating educational tool. However, while a sizeable number of studies have investigated the educational benefits of subtitling, there is a lack of research on children's attention to and processing of subtitled AV content, with a particular dearth of studies using eye tracking. Several eye tracking studies have found that adult viewers allocate a greater proportion of attention to the images of AV content with integrated subtitles[1] than with standard subtitles (Brooks and Armstrong, 2014; Brown, et al., 2015; Fox, 2018; Kruger, et al., 2018), although no such study seems to have been undertaken with children. Therefore, the key question addressed in this study is whether integrated subtitles might help children process AV content more effectively, and thus enhance their educational benefits.

2. Eye tracking research on children's processing of subtitled AV content

Over the last decade, a considerable number of studies have used eye tracking to investigate adult viewers' reception of subtitled AV content, focusing on a wide and growing range of subtitling issues. Szarkowska and Gerber-Morón (2019) provide a useful overview, and works such as Bisson, et al. (2014) and Romero-Fresco (2015) are particularly relevant as they examine attention distribution between subtitles and images. In contrast to the growing body of research focusing on adult viewers, there is a notable dearth of research with children as participants. While several empirical studies have focused on subtitles for children who are deaf or hard of hearing over the past few decades (see, for example Ward, et al., 2007; Tyler, et al., 2009; Lorenzo, 2010; Zárate, 2014; Zárate and Eliahoo, 2014; Tamayo, 2015; Tamayo and Chaume, 2017)[2] only a few studies have examined children's eye movements (Jensema, 2003; Cambra, Leal and Silvestre, 2013; Cambra, et al., 2014).

There is a similar paucity of research that has used eye tracking to test children's attention to interlingual subtitles. It seems that only four such studies have been conducted to

[1] A subtitle placement strategy which aims for the subtitles to fit in with viewers' natural gaze patterns, to reduce the length of the gaze path between the subtitles and the characters or salient elements of the image, and to maintain viewers' engagement with the characters and action.
[2] Also, see Fresno (2018) and Romero-Fresco (2018) for useful reviews of research in this area.

date, and have provided promising insights into children's subtitle reading behaviour compared to that of adults (d'Ydewalle and Van Rensbergen, 1989; d'Ydewalle and De Bruycker, 2007; Muñoz, 2017) and children's processing of subtitles presented at different speeds (Koolstra, van der Voort and d'Ydewalle, 1999). However, the range of eye tracking metrics tested in d'Ydewalle and Van Rensbergen (1989) and Koolstra, van der Voort and d'Ydewalle (1999) was limited, the only study to combine eye tracking with consecutive performance testing was Koolstra, van der Voort and d'Ydewalle (1999), and none collected self-reported data from the children. Clearly, therefore, there is a need for further research which includes a comprehensive range of eye tracking metrics, combines eye tracking with consecutive performance tests to assess the effectiveness of children's processing of subtitled AV content, and complements these with socio-cultural and attitudinal data, to provide a rigorous, comprehensive account of children's reception of subtitled AV content. The present study aims to contribute to filling these gaps.

3. Integrated subtitling

With mainstream subtitling practice having changed little since the 1930s, it is only recently that a small but growing number of researchers have begun to explore non-conventional subtitling forms, focusing on topics such as fansubbing and the growing use of creative text in arthouse cinema, commercial film and television, and theatre productions. McClarty (2013) argues in support of creative subtitling and sheds light on various potential benefits that creative subtitles, including integrated subtitle positioning, could bring to AV content and audiences. McClarty posits that "placing subtitles as close as possible to the face of the speaker would reduce the distance by which the viewer has to move their gaze between the subtitles and the intended point of interest" (2013, p.146) and suggests that this subtitling placement strategy maintains the viewer's engagement with the characters on screen. While McClarty's creative subtitles differ from conventional subtitles in several aspects other than their placement, such as colour, font, animation and other special effects, a small number of experimental studies have isolated the variable of subtitle position in their experiments, testing the reception of subtitles which differ from standard subtitles only in their placement.

Brooks and Armstrong (2014), Brown, et al. (2015) and Fox (2018) used eye tracking with adult participants to determine whether placing subtitles closer to the characters speaking

or other areas of interest on the screen would cause viewers to spend less time reading the subtitles and more time looking at the images, due to the reduced distance between the subtitles and points of interest on the screen. They posit that integrated subtitles could provide a more natural, immersive viewing experience, and function in a way that is more in keeping with the image composition and aesthetics of AV content than standard subtitles. Other possible advantages are that they may make processing subtitled AV content less effortful and speaker identification easier. All three found that integrated subtitling increased the amount of time participants spent viewing the images in AV content, and their findings indicate that integrated subtitles allow viewers to watch AV content in a more natural way. Kruger, et al. (2018) used eye tracking, electroencephalography and transportation subjective measures to measure cognitive load generated by viewing subtitles in standard and integrated positions, and concluded that the integrated subtitles may have allowed for a "deeper, uninterrupted processing of the content" (p.289). Moreover, the aforementioned studies which asked participants for their views on integrated subtitles received generally positive responses.

The findings discussed here reveal that integrated subtitling shows promise and is deserving of further scholarly attention. However, it appears that children's reception of integrated subtitles has yet to be investigated. Thus, given the positive effects observed in adult viewers' gaze behaviour when watching integrated subtitles, the present study has identified a need to investigate whether similar effects will occur in child viewers.

4. Methods

Gambier's model of the *three Rs* (2012) was adopted to investigate children's reception of standard and integrated subtitles: *response,* tested using eye tracking data to measure visual attention allocation and cognitive effort; *reaction*, measured via the performance variables of scene recognition and content comprehension; and *repercussion*, evaluated using questionnaires and interviews to assess the children's perceptions and views of the subtitle positions.

It is expected that since the integrated subtitle positions reduce saccade lengths (the distance that eyes travel) between the subtitles and the salient elements in the images, participants will thus have more time to allocate cognitive resources to processing the images and to understanding the AV content. Therefore, it is anticipated that participants will fixate

more and for longer on the images when viewing the integrated subtitles than when watching the standard subtitles. Moreover, it is predicted that, as a result, participants will expend less cognitive effort when processing the AV clips with the integrated subtitles. The variables of percentage of fixation time, percentage of fixation count and mean fixation duration are used to measure cognitive effort, since a larger number of fixations or longer fixation durations on the subtitle area are widely regarded as indicators of greater cognitive effort (d'Ydewalle and De Bruycker, 2007; Orrego-Carmona, 2016; Gerber-Morón, Szarkowska and Woll, 2018). Therefore, the hypotheses and sub-hypotheses formulated to test these assumptions are as follows:

H_1: *Participants will expend less cognitive effort when watching the AV content with integrated subtitles than when viewing the AV clips with standard subtitles.*

Therefore:

$H_{1.1}$: % fixation time on the images will be higher for the integrated subtitles condition.

$H_{1.2}$: % fixation count on the images will be higher for the integrated subtitles condition.

$H_{1.3}$: Mean fixation duration on the subtitles will be shorter for the integrated subtitles condition.

Moreover, it is expected that participants will process the images more effectively and will attain higher levels of comprehension of the AV content with integrated subtitles than with standard subtitles. Thus, the following hypotheses are proposed:

H_2: *Participants will attain higher scene recognition scores for the integrated subtitles condition.*

H_3: *Participants will attain higher content comprehension scores for the integrated subtitles condition.*

The additional eye tracking metrics of skipped subtitles, attention shift ratio and time to first fixation are also tested to assess whether the integrated subtitles have any disruptive effects on participants' reading patterns. Participants are also asked about their preferences and perceived levels of effort when viewing the standard and integrated subtitles.

4.1 Participants

17 participants aged 8-9 years took part in this study. However, 4 participants were excluded due to poor quality eye tracking data and the data of the remaining 13 participants (7 female, 6 male; mean age: 8.23; *SD* = 0.44) are analysed. A homogeneous sampling method (Coolican, 2014) was used to control subject variables.[3] Participants were recruited from the same geographical area (Northern Ireland), school and class. Moreover, their Spanish teacher confirmed they were all native English speakers with comparable knowledge of Spanish. Participants close in age were recruited to limit the developmental range of the group.

4.2 Eye tracking and other instruments

Eye tracking

Participants' eye movements were recorded using a Tobii T60, with a sampling rate of 60Hz. A nine-point calibration procedure was performed before each recording. The gaze data were recorded using Tobii Studio version 3.2.2. The I-VT fixation filter was selected and the minimum fixation duration threshold was set at 100 milliseconds (Perego, et al., 2010; Tobii Technology, 2012). To assess data quality raw gaze data were examined and processed manually to remove track losses and other poor quality data. The statistics for the analyses were calculated using Microsoft Excel and IBM SPSS Statistics version 24.

Scene recognition tests

The scene recognition tests (SRTs) are forced choice recognition memory tests. Five pairs of screenshots were presented to participants after each clip. One image in each pair appeared in the clip and one did not, but was from the same film. The image pairs were selected to ensure the composition, characters and setting were as similar as possible. The presentation order of the pairs was randomised.

Content comprehension tests

Participants completed two content comprehension tests (CCTs), one after each clip. Each CCT contained five multiple choice questions. Each question had one correct answer and two plausible distractors, the order of which was randomised, and a "Not sure" option. The

[3] Homogeneous sampling is a non-probability, purposive sampling method which involves selecting "settings, groups, and/or individuals based on similar or specific characteristics (Onwuegbuzie and Collins, 2007, p.285).

questions assessed literal and inference comprehension, and tested gap-filling, causal and emotional inferences (Kispal, 2008). Moreover, the CCTs were designed to assess participants' comprehension of the visual and auditory verbal and non-verbal information in the clips holistically. The level of difficulty of the vocabulary in the CCTs was analysed to verify that it was appropriate to the participants' stage of literacy education, following a similar procedure to Zárate (2014). Kress and Fry's list of the first 1000 words acquired by children (2015) and *The Children's Printed Word Database* (Masterson, et al., 2002) were consulted. Low frequency words were checked against the spelling lists in *English programmes of study: Key stages 1 and 2* (Department for Education, 2013). Finally, the CCTs were approved by an independent primary school teacher.

Questionnaires

Questionnaires can be used with children in middle childhood, but should be designed to be appropriate to their stage of development (de Leeuw, 2011). Thus, the questionnaires used in this study were written using vocabulary that is easy to understand and short, simple, clear sentences. Moreover, faces scales were used as they attract attention, are thought to be easy for children to understand (Chambers, et al., 1999) and to avoid satisficing due to disinterest or boredom (de Leeuw, 2011). The faces were paired with text items on a 5-level Likert scale. The number of questions was limited to seven, and response options were limited to a maximum of five. The level of difficulty of the vocabulary in the questionnaires was assessed using the same procedure as that adopted for the CCTs. The questionnaires measured self-reported effort using the ratings provided by participants in response to the following questions:

- *How hard or easy were the subtitles to read?*
- *How hard or easy was it to understand the story of the video?*

Moreover, the responses to the following questions were used to measure perception of subtitle speed and approval of subtitle position, respectively:

- *Were the subtitles too fast, too slow or about right?*
- *Did you like the position of the subtitles on the screen?*

Semi-structured interviews

Semi-structured interviews were held with participants in pairs. Semi-structured interviews are considered to be appropriate for participants in middle childhood and are thought to address the issue of developing literacy as children can express themselves in more complex, nuanced ways verbally than in writing (Scott, 2000). Only data relating to participants' views and perceptions regarding subtitle placement are analysed here.

4.3 Materials

Two clips from a Mexican animated film, *La Leyenda de la Llorona* [The Legend of the Weeping Woman] (2011), were used as stimuli. The film presents a well-known Mexican folktale and thus could be used as a resource for teaching Spanish and about Mexican folklore and culture. Both clips, coded WKM (*With Kika's Mother*) and LL (*La Llorona*) in Table 1, contain complete scenes and present coherent sections of the film's story arc. Thus, as illustrated in Table 1, the clips are of comparable rather than identical duration. Word count and rate of the source dialogues are also comparable. Comparability of subtitle density was also a key criterion of selection – the differences between the numbers of subtitles, two-line subtitles and characters per second are small. While the differences between the numbers of characters and percentage of time that the subtitles are on screen are somewhat larger, a degree of variation is to be expected, particularly since the aim was to create subtitles that are analogous to good quality commercial subtitles rather than to impose artificial restrictions for the aims of the experiment. Overall, the clips were deemed comparable and suitable for use as stimuli in the study.

Table 1. Density of source dialogue and subtitles in clips

		Source dialogue		**Subtitles**					
Clip	*Duration (seconds)*	*Word count*	*Words per second*	*No. subtitles*	*No. 2-line subtitles*	*% 2-line subtitles*	*No. characters*	*Characters /second*	*% time subtitles on screen*
WKM	106	169	1.59	40	6	15%	821	11.09	65.25%
LL	116	177	1.53	39	5	13%	768	10.63	57.1%

The subtitle speed conformed with the traditional six-second rule, as recommended by the BBC (Ford Williams, 2009) for its channel aimed at viewers in middle childhood, CBBC. No

subtitles appeared over shot changes as it would complicate the positioning of the integrated subtitles and since participants were young readers. The typeface used was Arial, following Zárate (2014) and Tamayo (2015).

4.4 Integrated subtitle placement

Participants viewed two clips, one with the subtitles in the standard position; i.e., centre-aligned and at the bottom of the screen, and one with individually placed, integrated subtitles. The criteria adopted for integrated subtitle placement were shaped by Fox's "modular guidelines" (2018) and McClarty's (2013) recommendations. They were also informed by eye tracking research which has found that viewers exhibit attentional synchrony when viewing AV content (Goldstein, Woods and Peli, 2007; Sawahata, et al., 2008). Moreover, viewers focus most on faces and other salient elements in static and moving images, making repeated cycles of fixations on focal points (Yarbus, 1967; DeAngelus and Pelz, 2009). Furthermore, Cambra, et al. (2014) observed that their child participants focused most on characters' faces, particularly their mouths.

Before deciding on which strategy to use for the placement of each integrated subtitle, the following elements were considered:

- focal points and their distribution;
- proportions of screen space occupied by characters/focal points and the background;
- the character(s) speaking, whether they are on- or off-screen, and the direction of speech;
- movement action and camera movement;
- subtitle size;
- whether the area considered for subtitle placement provides clear, even contrast.[4]

The subtitle positions used and the proportions of subtitles placed according to each strategy are shown in Table 2.

Table 2. Proportions of integrated subtitle positions used

Integrated subtitle position	**% subtitles**
In between 2+ characters on screen	35
Following direction of speech (character on screen)	32

[4] This list is informed by Fox's "modular guidelines" for individually placed subtitles (Fox, 2018).

Below speakers' faces	23
Following direction of speech (character off screen)	9
Standard subtitle position	1

4.5 Dependent variables

Six eye tracking metrics were tested, as shown in Table 3. Eye tracking data were collected from areas of interest (AOIs) placed on the subtitle and image areas of the clips. Other dependent variables are scene recognition and content comprehension, and the questionnaires measured self-reported effort, perception of subtitle speed and approval of subtitle position.

Table 3. Eye tracking metrics tested

Eye tracking metric	**Calculation**
% fixation time on subtitle/image areas	Dividing total fixation time on subtitle/image AOIs by total fixation time of segments of clips when subtitles on screen.
% fixation count on subtitle/image areas	Dividing fixation count on subtitle/image AOIs by total fixation count for segments of clips when subtitles on screen.
Mean fixation duration on subtitle/image areas (ms)	Dividing total fixation time on subtitle/image AOIs by fixation count on subtitle/image AOIs for each recording.
% skipped subtitles	Subtitles which the participants have not read; i.e., subtitle AOIs on which no fixations have been recorded. Calculated by dividing the number of skipped subtitles by the total number of subtitles in each recording.
Attention shift ratio	The average number of times the viewer's gaze shifts from the subtitle to the image AOIs and vice versa per subtitle. Calculated by counting the number of times a participant's fixation on the subtitle AOI is followed by a fixation on the image AOIs and vice versa while the subtitles are on the screen, and dividing the total number of gaze shifts by the number of subtitles on the clip.
Time to first fixation (ms)	The duration of time from the onset of a subtitle to the viewer's first fixation on the subtitle. Calculated by subtracting the recording time of the start of the first fixation from the onset time of each subtitle.

4.6 Procedure

This study took place at the children's primary school. Participants' eye movements were recorded while they watched two clips. After each clip, participants completed an SRT, CCT and questionnaire. Participants were tested individually and interviewed in pairs afterwards. The duration of testing was 25 minutes per participant, including 10 minutes for the interviews. This study was approved by the Ethics Committee of the School of Modern Languages at Queen's University Belfast. Permission was obtained from the school principal and information letters and Informed Participant Release Forms were sent to parents. The active, freely given, ongoing consent of the children was also sought. All information and instructions were written in clear, age-appropriate language and approved in advance by the children's teacher.

5. Results

5.1 Statistical testing and control variables

To assess normality of distribution and thus determine whether parametric or non-parametric tests should be employed to test the continuous dependent variables, the Shapiro-Wilk test was performed on the data for each variable, and Q-Q plots and skewness and kurtosis statistics were analysed. Most variables were found not to have a normal distribution for at least one of the subtitle conditions, apart from the CCT scores and time to first fixation data, which were observed to have a normal distribution for both subtitle conditions. Thus, and moreover since this study employs a repeated measures design, the paired *t*-test was used to test the CCT scores and time to first fixation data, and the Wilcoxon matched-pairs signed-ranks test was selected to test the other continuous dependent variables. Several control variables – order of clip presentation, order of presentation of subtitle position, and clip type – were included as independent variables in the statistical tests to assess whether they had any significant impact on the dependent variables. Moreover, to control for order effects the order of presentation of the subtitle positions and clips was counterbalanced by randomly assigning participants to one of four groups. The statistics from the tests to assess normality of distribution and the control variables can be consulted in Author (2017).

5.2 Eye tracking results

Table 4. Differences in eye tracking metrics between standard and integrated subtitle conditions

Eye tracking metric	Median (*IQR*)				
	Standard	**Integrated**	***Z***	***p***	**r**
% fixation time (subtitles)	69.5 (19.5)	63.7 (16.8)	2.55	.011*	0.50
% fixation time (images)	30.5 (19.5)	36.3 (16.8)			
% fixation count (subtitles)	74.3 (14.1)	68.7 (20.7)	2.48	.013*	0.49
% fixation count (images)	25.7 (14.1)	31.3 (20.7)			
Mean fixation duration (subtitles) (ms)	267 (49.8)	263 (37.5)	0.18	.86	-
Mean fixation duration (images) (ms)	319 (53.2)	317 (71.6)	0.59	.55	-
% skipped subtitles	0.0	2.5	0.91	.36	-
Attention shift ratio	1.18 (0.45)	1.38 (0.44)	0.98	.33	-
	Mean (*SD*)				
	Standard	**Integrated**	***t***	***df***	***p***
Time to first fixation (ms)	204 (88)	253 (75)	-1.61	12	.13

A summary of the results for the eye tracking metrics is presented in Table 4. As can be observed, the median of the percentages of fixation time on the integrated subtitles (*Mdn* = 63.7%) is lower than for the standard subtitles condition (*Mdn* = 69.5%), and correspondingly, the median of the percentage of fixation time on the images for the integrated subtitles condition (*Mdn* = 36.3%) is higher than for the standard subtitles condition (*Mdn* = 30.5%). Moreover, 10 of the 13 participants spent proportionally longer looking at the images when the subtitles were in the integrated position. The differences between the results for the two conditions were found to be significant (Z = 2.55, p = .011) and the effect size is large (r = 0.50).[5] When outliers are removed, the differences are still significant (Z = 2.67, p = .008) and the effect size is large (r = 0.81).

Likewise, the median of the percentage of fixation count on the integrated subtitles (*Mdn* = 68.7%) is lower than the median for the standard subtitles condition (*Mdn* = 74.3%), and the median of the percentage of the number of fixations on the images for the integrated subtitles condition (*Mdn* = 31.3%) is higher than for the standard subtitles condition (25.7%). Moreover, 11 of the 13 participants fixated proportionally more times on the images when the subtitles were in the integrated position. The differences between the results for the two conditions were found to be significant (Z = 2.48, p = .013), with a moderate to large effect

[5] Cohen's r has been used to calculate effect size.

size (r = 0.49). With outliers removed, the differences remain significant (Z = 2.22, p = .026) and the effect size is large (r = 0.67).

Table 4 shows that no statistically significant differences were found in the results for the other eye tracking metrics. However, while the result for time to first fixation was only somewhat close to statistical significance ($t(12) = -1.61$, p = .13), once outliers are excluded, the difference between the two conditions is statistically significant, with $t(10) = -2.42$, p = .036.

5.3 Scene recognition

The SRTs measured scene recognition according to the number of correct responses, from 0 to 5. The median values, range and interquartile range for scene recognition scores were found to be the same for both conditions (*Mdn* = 5.0; *R* = 2.0, min. = 2.00, max = 5.00; *IQR* = 1.00) and indeed, no statistically significant difference was found between the two (Z = 0.09, p = .93).

5.4 Content comprehension

The CCTs measured content comprehension according to the number of correct responses, from 0 to 5. Mean CCT scores were similar (standard subtitles condition: M = 3.1, SD = 1.50; integrated subtitles condition: M = 2.9, SD = 1.38) and the difference between the scores for the two conditions was found not to be statistically significant t(12) = 0.29, p = .78.

5.5 Self-reported effort

Self-reported effort is measured on a 5-point Likert scale, with scores from *very hard*= 5 to *very easy*= 1. The median self-reported effort scores for both conditions are identical (*Mdn* = 2.0) and the range is similar (standard subtitles condition: *R* = 3.0, min. = 1.0, max. = 4.0; integrated subtitles condition: *R* = 2.5, min. = 1.0, max. = 3.5). The interquartile range for the integrated subtitles condition (*IQR* = 2.5) is much larger than for the standard subtitles condition (IQR = 0.5). However, the differences in the two conditions were found not to be statistically significant (Z = 0.29, p = .77). Moreover, no clear pattern emerges from the scores given by each participant, as five reported equal scores for both conditions, four perceived the

clips with the standard subtitles as more difficult, and four rated the clips with integrated subtitles as more difficult.

5.6 Perception of subtitle speed

Table 5: Perception of subtitle speed

	Perception of subtitle speed		
	Too slow	*About right*	*Too fast*
Standard subtitles	-	11	2
Integrated subtitles	-	10	3

As is shown in Table 5, a large majority of participants responded that the speed of both the standard and the integrated subtitles was about right, with little difference between the two conditions.

5.7 Approval of subtitle position

Table 6: Approval of subtitle position

	Q5. Did you like the position of the subtitles on the screen?				
	Not at all	*Not much*	*Neither*	*Quite liked it*	*Yes, a lot*
Standard subtitles	-	1	-	6	6
Integrated subtitles	-	2	-	8	3

As is displayed in Table 6, almost all participants "quite liked" or liked the position of both the standard and integrated subtitles on the screen "a lot". The results for the standard subtitles appear more positive, and although the result of the Wilcoxon test for this variable was very close to significance ($Z = -1.89$, $p = .059$), it was beyond the threshold of 5%.

5.8 Interview results

Participants' generally positive approval ratings for both subtitle positions in the questionnaire results are broadly consistent with the interview findings, as most participants commented that

they liked both subtitle positions and did not overtly state a preference for either. A few participants specified that they liked viewing the integrated subtitles on different areas of the screen, while they also found the standard subtitles easy to read. For example, one remarked: "The first clip trained my eyes to look everywhere and the second clip made it really legible".

However, a few participants expressed the difficulty they felt when reading the integrated subtitles, due to not knowing where the next one would appear. One commented:

> I think they [the integrated subtitles] were a bit hard, 'cause you never knew when they were gonna pop up, you never knew where they were gonna go and stuff. […] The second one [the standard subtitles] you know where it's gonna pop up 'cause they were all in the middle.

Others expressed a preference for the integrated subtitles in their comments. One commented that they found the integrated subtitles easy to understand, and another thought that the different positions of the integrated subtitles on the screen helped them to read and understand them. Others reported that they felt the integrated subtitles provided the opportunity to look more at the images of the AV content. For example, one participant remarked:

> I would prefer them in different places because if you are looking at them in different places then normally I see the entire screen, instead of keeping my face there [indicates bottom of imaginary screen] so I don't get to see the pictures as well.

Another participant also felt that the standard subtitles limited their gaze to the bottom of the screen: "I liked it when it was, like, moving around all over the place so you had to move your eyes more. You just stared at the ones at the bottom instead of moving your eyes." Another participant felt that subtitles placed in the standard position are too far away from the central points of interest on the screen and are thus harder to read. Moreover, they identified contrast as an issue for subtitles in the standard position:

> When it's at the bottom, it's hard for me to read because if I want to look at the pic–, and there might be some white stuff at the bottom and the subtitles are white too so then you won't see […] Also, it's just harder to read in general, because it's out of the way.

6. Discussion

Table 7. Summary of hypothesis testing

Hypothesis	Variable	Confirmation	
H_1	cognitive effort	(✓)	This hypothesis is only partially confirmed, as while the percentages of fixation time and count on the images of the subtitled AV clips were found to be significantly higher for the integrated subtitles condition, no effect was found for mean fixation duration.
$H_{1.1}$	*% fixation time on images*	✓	*The % of fixation time on the images of the subtitled AV clips was higher for the integrated subtitles condition than for the standard subtitles condition.*
$H_{1.2}$	*% fixation count on images*	✓	*The % of fixation count on the images of the subtitled AV clips was higher for the integrated subtitles condition than for the standard subtitles condition.*
$H_{1.3}$	*mean fixation duration on the subtitles*	*X*	*No significant effect was found for subtitle position on the mean fixation durations on the subtitles.*
H_2	scene recognition	X	SRT scores were high overall but subtitle position was not found to have an effect on scene recognition.
H_3	content comprehension	X	The mean and distribution of the CCT scores were very similar for both conditions and subtitle position was not found to have an effect on content comprehension.

The results of the hypothesis testing for the independent variable of subtitle position are summarised in Table 7 and are discussed here. As expected, when participants watched the clips with integrated subtitles, they spent a significantly larger proportion of their viewing time looking at the images and also fixated more times on the images while the subtitles were on the screen, compared to when they watched the clips with the standard subtitles. Interestingly, this finding is consistent with those of existing studies on integrated subtitling conducted with adult viewers (Brooks and Armstrong, 2014; Brown, et al., 2015; Fox, 2018), as discussed in section 2. However, the hypothesis that the participants would exert lower levels of cognitive

effort when watching the AV clips with integrated subtitles (H_1) was only partially confirmed, as subtitle position was not found to affect mean fixation duration.

It is also noteworthy that the percentages of fixation time on the subtitles were much higher than those reported in d'Ydewalle and Van Rensbergen (1989), Koolstra, van der Voort and d'Ydewalle (1999) and d'Ydewalle and De Bruycker (2007) and the percentages of fixation count on the subtitles were also high overall. In the aforementioned studies, child participants looked at the subtitles for approximately one third to half (34%-52%) of the time they were on screen, whereas in this study 11 of the 13 participants allocated between 57-86% of their fixation time to the (standard) subtitles when they were on screen. This may be partly explained by the fact that while the children in the aforementioned studies live in countries where viewers are regularly exposed to subtitled AV content in other languages (Belgium and the Netherlands), the participants in this study have had little prior exposure to such content. Moreover, since the participants in d'Ydewalle and De Bruycker (2007) were slightly older (10-12 years) than those in the present study, age may have played a role in terms of more advanced readers spending less time fixating on the subtitles.

As yet, it seems that only one other study has measured children's mean fixation durations when viewing subtitles (d'Ydewalle and De Bruycker, 2007). Thus, the present study contributes new data to address this gap. The mean of the duration of all the participants' fixations on the standard subtitles was calculated: M = 266 ms, SD = 150, which is also very close to the median of 267 ms shown in Table 4. This figure is longer than those reported in existing studies conducted with adults which were found to range from 178-221 ms (d'Ydewalle and De Bruycker, 2007; Perego, et al., 2010, Orrego-Carmona, 2015; Perego, Orrego-Carmona and Bottiroli, 2016). Moreover, it supports the findings of d'Ydewalle and De Bruycker (2007), who reported longer fixation durations on subtitles in children than in adult viewers.

Furthermore, this finding is consistent with observations of reading researchers that children's fixations are longer than adults' when reading static text. Mean fixation durations of 256-285 ms have been reported for 7-11-year-olds (Blythe, et al., 2009; Joseph, et al., 2009), while a widely cited figure for adult readers is 225 ms (Rayner, 1998). Interestingly, the mean fixation duration reported here falls within the above-cited range, and is longer than those presented by d'Ydewalle and De Bruycker (2007),[6] which supports the hypothesis that the

[6] 243 ms for one line subtitles and 252 ms for two line subtitles (d'Ydewalle and De Bruycker, 2007, p.199).

duration of children's fixations when reading subtitles decreases with age as their reading skills develop, as appears to be the case when reading static text. Moreover, as shown in Table 4, the mean fixation durations on the images are longer than those on the subtitles for both conditions. This result is consistent with the findings of existing studies performed with adult viewers (Perego, et al., 2010; Orrego-Carmona, 2015; Perego, Orrego-Carmona and Bottiroli, 2016).

As shown in Table 4, the percentages of skipped subtitles were low overall, and no difference was found between the two subtitle conditions. This is an interesting finding given that one concern might have been that since the integrated subtitles appear in different areas on the screen participants may have missed more of them. However, this was not found to be the case. This may be due to their saliency; i.e., that the participants' attention was attracted to the subtitles due to their dynamic nature (Bisson, et al., 2014). Moreover, the low percentages of skipped subtitles overall are consistent with the findings of Muñoz (2017) for viewers aged 10-12. The figures for attention shift ratio were also low, which is likely since, overall, participants spent a large proportion of the time that the subtitles were on the screen looking at the subtitles. Moreover, subtitle position was not found to have an effect on attention shift ratio. As for time to first fixation, it is noteworthy that the difference between the two conditions was statistically significant once outliers were excluded (see section 5.2). Integrated subtitles may cause longer first fixation times as viewers cannot be certain where each subtitle will appear. A more conclusive result may be obtained with a larger number of participants.

Regarding scene recognition, although no difference was found between the conditions, interestingly, scores were high overall despite the low percentages of fixation time and count on the images. These results suggest that a large majority of participants processed the images of the AV content effectively, irrespective of subtitle position and of having dedicated large percentages of fixation time to viewing the subtitles. Moreover, these findings are consistent with those of Perego, et al. (2010), performed with adult participants, and may be due to the "phenomenon of picture superiority" (ibid., p.262). Moreover, the findings of experimental studies testing children's processing of media content suggest that children's cognitive processing of information presented in visual or AV mode is more effective than that of auditory information (Pezdek and Hartman, 1983; Walma Van Der Molen and Van Der Voort, 2000). As for content comprehension, while the integrated subtitles did not improve the participants' comprehension of the clips, it can be concluded that they do not seem to have had a detrimental effect.

Self-reported effort scores were low overall and participants did not perceive the integrated subtitles as more difficult to read or the clips with the integrated subtitles as more difficult to understand. Moreover, a large majority reported that neither subtitle conditions were too fast, and although the higher approval ratings for the standard subtitle position came close to statistical significance, the threshold of 5% was not met. Furthermore, approval ratings for both subtitle positions were high overall.

The interview data were found to be largely consistent with the questionnaire findings and also provided richer information on the participants' views. Several participants provided enlightening comments on the subtitle positions, with some preferring the standard subtitle position as they knew where each subtitle would appear, and others even identifying that the integrated subtitles allowed them to spend more time looking at the images, while the standard subtitles caused them to focus more on the bottom of the screen and miss more of the images.

7. Conclusion

This study seems to have been the first to combine eye tracking with performance measures and interviews to investigate children's reception of AV content with interlingual subtitles, and to test integrated subtitles with child viewers. In this way, it has sought to contribute to the existing research in this area. A limitation of this study is its small sample size, which limits the generalisability of the findings, and caution must be exercised when drawing inferences about child viewers in general. However, its main finding – that the children spent a significantly larger proportion of their viewing time looking at the images of the AV content and also fixated more times on the images when the subtitles were in the integrated position – is promising, and is consistent with the findings of existing studies performed with adults (Brooks and Armstrong, 2014; Brown, et al., 2015; Fox, 2018). These results, together with the findings that the integrated subtitles did not seem to have a detrimental effect on the children's viewing patterns, processing of the images or comprehension of the clips, and that, overall, the children approved of the integrated subtitles and perceived them as easy to read and understand indicate that further research on children's reception of integrated subtitles is warranted. Given the small sample size of the present study, future research performed with larger sample sizes would help to confirm whether integrated subtitles hold benefits for young viewers in terms of comprehension, cognitive load or subjective preferences. Moreover, future studies could use different devices or types of AV materials or investigate whether integrated subtitling provides advantages for young viewers with different reading abilities or who are deaf or hard of hearing. As noted by Romero-Fresco (2015), further research on gaze distribution would ensure a continuing focus on viewers' comprehension and enjoyment of both subtitles and images, and could provide research-based recommendations on aspects such as subtitle positioning.

References

Black, S., 2017. *Through a child's eyes: investigating the effects of subtitle placement and reading level on children's reception of subtitled audiovisual content.* PhD. Queen's University Belfast.

Bisson, M.J., Van Heuven, W.J.B., Conklin, K. and Tunney, R.J., 2014. Processing of native and foreign language subtitles in films: an eye tracking study. *Applied Psycholinguistics*, 35(2), pp.399–418.

Blythe, H.I., Liversedge, S.P., Joseph, H.S.S.L., White, S.J. and Rayner, K., 2009. Visual information capture during fixations in reading for children and adults. *Vision Research*, 49(12), pp.1583-1591.

Bravo, M., 2008. *Putting the reader in the picture: Screen translation and foreign-language learning.* PhD. Universitat Rovira i Virgili.

Brooks, M. and Armstrong, M., 2014. *TVX2014 Short Paper - Enhancing Subtitles.* [online] Salford: BBC. Retrieved from: http://www.bbc.co.uk/rd/blog/2014-10-tvx2014-short-paper-enhancing-subtitles

Brown, A., Jones, R., Crabb, M., Sandford, J., Brooks, M., Armstrong, M. and Jay, C., 2015. *White Paper WHP305. Dynamic Subtitles: The User Experience*. [pdf] Salford: BBC. Retrieved from: http://downloads.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP305.pdf

Cambra, C., Leal, A. and Silvestre, N., 2013. The interpretation and visual attention of hearing impaired children when watching a subtitled cartoon. *The Journal of Specialised Translation*, 20, pp.134-146.

Cambra, C., Penacchio, O., Silvestre, N. and Leal, A., 2014. Visual attention to subtitles when viewing a cartoon by deaf and hearing children: an eye-tracking pilot study. *Perspectives: Studies in Translatology*, 22(4), pp.607-617.

Chambers, C.T., Giesbrecht, K., Craig, K.D., Bennett, S.M. and Huntsman, E., 1999. A comparison of faces scales for the measurement of pediatric pain: children's and parents' ratings. *Pain,* 83(1), pp.25-35.

Coolican, H., 2014. *Research Methods and Statistics in Psychology*. 6th edn. Hove and New York: Psychology Press.

DeAngelus, M. and Pelz, J.B., 2009. Top-down control of eye movements: Yarbus revisited. *Visual Cognition*, 17(6), pp.790-811.

De Leeuw, E., 2011. Improving Data Quality when Surveying Children and Adolescents: Cognitive and Social Development and its Role in Questionnaire Construction and Pretesting. *Report prepared for the Annual Meeting of the Academy of Finland: Research Programs Public Health Challenges and Health and Welfare of Children and Young People.* [pdf] Naantali, Finland, 10-12 May, 2011. Retrieved from: http://www.aka.fi/globalassets/awanhat/documents/tiedostot/lapset/presentations-of-the-annual-seminar-10-12-may-2011/surveying-children-and-adolescents_de-leeuw.pdf

Department for Education (DfE), 2013. *English programmes of study: key stages 1 and 2. National Curriculum in England.* [pdf] England: DfE. Retrieved from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/335186/PRIMARY_national_curriculum_-_English_220714.pdf

D'Ydewalle, G. and De Bruycker, W., 2007. Eye Movements of Children and Adults While Reading Television Subtitles. *European Psychologist*, 12(3), pp.196-205.

D'Ydewalle, G. and Van Rensbergen, J., 1989. Developmental Studies of Text-Picture Interactions in the Perception of Animated Cartoons with Text. In: H. Mandl and J.R. Levin, eds. 1989. *Knowledge Acquisition from Text and Pictures.* North Holland: Elsevier Science Publishers B.V., pp.233-248.

Ford Williams, G., 2009. *Online Subtitling Editorial Guidelines V1.1.* [pdf] London: BBC. Retrieved from: http://www.bbc.co.uk/guidelines/futuremedia/accessibility/subtitling_guides/online_sub_editorial_guidelines_vs1_1.pdf

Fox, W., 2018. *Can integrated titles improve the viewing experience? Investigating the impact of subtitling on the reception and enjoyment of film using eye tracking and questionnaire data.* Berlin: Language Science Press.

Fresno, N., 2018. Watching accessible cartoons: The speed of closed captions for young audiences in the United States. *Perspectives: Studies in Translation Theory and Practice*, 26(3), pp.405-421.

Gambier, Y., 2012. The position of audiovisual translation studies. In: C. Millán and F. Bartrina, eds. 2012. *The Routledge Handbook of Translation Studies*. Abingdon: Routledge, pp.45-59.

Gerber-Morón, O., Szarkowska, A. and Woll, B., 2018. The impact of text segmentation on subtitle reading. *Journal of Eye Movement Research*, 11(4).

Gernsbacher, M.A., 2015. Video Captions Benefit Everyone. *Policy Insights from the Behavioral and Brain Sciences*, 2(1), pp.195-202.

Goldstein, R., Woods, R. and Peli, E., 2007. Where people look when watching movies: Do all viewers look at the same place? *Computers in Biology and Medicine*, 37(7), pp.957-964.

Jensema, C., 2003. *The Relation Between Eye Movement and Reading Captions and Print by School-Age Deaf Children.* Wheaton: Institute for Disability Research and Training, Inc.

Joseph, H.S.S.L., Liversedge, S.P., Blythe, H.I., White, S.J. and Rayner, K., 2009. Word length and landing position effects during reading in children and adults. *Vision Research*, 49(16), pp.2078-2086.

Kispal, A., 2008. *Effective Teaching of Inference Skills for Reading: Literature Review.* [pdf] Slough: National Foundation for Educational Research. Retrieved from: https://www.nfer.ac.uk/publications/EDR01/EDR01.pdf

Koolstra, C.M. and Beentjes, J.W.J., 1999. Children's Vocabulary Acquisition in a Foreign Language Through Watching Subtitled Television Programs at Home. *Educational Technology Research and Development*, 47(1), pp.51-60.

Koolstra, C. M., van der Voort, T.H.A. and d'Ydewalle, G., 1999. Lengthening the Presentation Time of Subtitles on Television: Effects on Children's Reading Time and Recognition. *Communications*, 24(4), pp.407-422.

Kothari, B. and Bandyopadhyay, T., 2014. Same language subtitling of Bollywood film songs on TV: Effects on literacy. *Information Technologies and International Development*, 10(4), pp.31-47.

Kress, J.E. and Fry, E.B., 2015. *The Reading Teacher's Book of Lists.* 6th edn. San Francisco: Jossey Bass.

Kruger, J.L., Doherty, S., Fox, W., and de Lissa, P., 2018. Multimodal measurement of cognitive load during subtitle processing: Same-language subtitles for foreign language viewers. In: I. Lacruz and R. Jääskeläinen, eds. 2018. *Innovation and Expansion in Translation Process Research*. London: John Benjamins Publishing Company. Ch.12.

Künzli, A. and Ehrensberger-Dow, M., 2011. Innovative subtitling: A reception study. In: C. Alvstad, A. Hild and E. Tiselius, eds. 2011. *Methods and Strategies of Process Research: Integrative Approaches in Translation Studies.* Amsterdam and Philadelphia: John Benjamins Publishing Company, pp.187-200.

La Leyenda de la Llorona, 2011. [DVD] Mexico: Ánima Estudios, Eficine 226.

Linebarger, D., 2001. Learning to Read From Television: The Effects of Using Captions and Narration. *Journal of Educational Psychology*, (93)2, pp.288-298.

Lorenzo, L., 2010. Criteria for elaborating subtitles for deaf and hard of hearing children in Spain: A guide of good practice. In: A. Matamala and P. Orero, eds. 2010. *Listening to Subtitles: Subtitles for the Deaf and Hard of Hearing.* Bern: Peter Lang AG, pp.139-147.

Masterson, J., Dixon, M., Stuart, M. and Lovejoy, S., 2002. *Children's Printed Word Database*. [online] Colchester: Department of Psychology, University of Essex. Retrieved from: http://www.essex.ac.uk/psychology/cpwd/

McClarty, R., 2013. *Film and Translation: The Art of Subtitling.* PhD. Queen's University Belfast.

Muñoz, C., 2017. The role of age and proficiency in subtitle reading. An eye tracking study. *System*, 67, pp.77-86.

Onwuegbuzie, A. J. and Collins, K. M. T., 2007. A Typology of Mixed Methods Sampling Designs in Social Science Research. *The Qualitative Report*, 12(2), pp.281-316.

Orrego-Carmona, D., 2015. *The Reception of (Non) Professional Subtitling.* PhD. Universitat Rovira i Virgili.

Orrego-Carmona, D., 2016. A reception study on non-professional subtitling. Do audiences notice any difference? *Across Languages and Cultures*, 17(2), pp.163-181.

Perego, E., del Missier, F., Porta, M. and Mosconi, M., 2010. The Cognitive Effectiveness of Subtitle Processing. *Media Psychology*, 13, pp.243-272.

Perego, E., Orrego-Carmona, D. and Bottiroli, S., 2016. An Empirical Take on the Dubbing vs. Subtitling Debate: An Eye Movement Study. *Lingue e Linguaggi*, 19, pp.255-274.

Pezdek, K. and Hartman, E.F., 1983. Children's Television Viewing: Attention and Comprehension of Auditory versus Visual Information. *Child Development*, 54(4), pp.1015-1023.

Rayner, K., 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), pp.372-422.

Romero-Fresco, P., 2018. Reception studies in live and pre-recorded subtitles for the deaf and hard of hearing. In: E. Di Giovanni and Y. Gambier, eds. 2018. *Reception Studies and Audiovisual Translation*. Amsterdam and Philadelphia: John Benjamins, pp. 199-224.

Romero-Fresco, P, ed. 2015. *The reception of subtitles for the deaf and hard of hearing in Europe*. Bern and Berlin: Peter Lang, pp.335–342.

Sawahata, Y., Khosla, R., Komine, K., Hiruma, N., Itou, T., Watanabe, S., Suzuki, Y., Hara, Y. and Issiki, N., 2008. Determining comprehension and quality of TV programs using eye-gaze tracking. *Pattern Recognition*, 41, pp.1610-1626.

Scott, J., 2000. Children as Respondents: The Challenge for Quantitative Methods. In: P. Christensen and A. James, eds. 2000. *Research with Children: Perspectives and Practices*. London and New York: Falmer Press. Ch.5.

Szarkowska, A. and Gerber-Morón, O., 2019. Two or three lines: a mixed-methods study on subtitle processing and preferences. *Perspectives: Studies in Translation Theory and Practice*, 27(1), pp.144-164.

Tamayo, A., 2015. *Estudio descriptivo y experimental de la subtitulación en TV para niños sordos: una propuesta alternativa.* PhD. Universitat Jaume I.

Tamayo, A. and Chaume, F., 2017. Subtitling for d/Deaf and Hard-of-Hearing Children: Current Practices and New Possibilities to Enhance Language Development. *Brain Sciences*, 7(7), DOI: 10.3390/brainsci7070075.

Tobii Technology, 2012. *User Manual – Tobii Studio.* [pdf] Retrieved from: http://www.tobii.com/en/eye-tracking-research/global/library/manuals/

Tyler, M. D., Jones, C., Grebennikov, L., Leigh, G., Noble, W. and Burnham, D., 2009. Effect of caption rate on the comprehension of educational television programmes by deaf school students. *Deafness & Education International*, 11(3), pp.152–162.

Walma Van Der Molen, J.H. and van Der Voort, T.H.A., 2000. The Impact of Television, Print, and Audio on Children's Recall of the News. *Human Communication Research,* 26, pp.3-26.

Ward, P., Wang, Y., Paul, P. and Loeterman, M., 2007. Near-verbatim captioning versus edited captioning for students who are deaf or hard of hearing: a preliminary investigation of effects on comprehension. *American Annals of the Deaf*, 152(1), pp.20–28.

Yarbus, A.L., 1967. *Eye Movements and Vision.* New York: Plenum Press.

Zárate, S., 2014. *Subtitling for deaf children: Granting accessibility to audiovisual programmes in an educational way.* PhD. University College London.

Zárate, S. and Eliahoo, J., 2014. Word recognition and content comprehension of subtitles for television by deaf children. *The Journal of Specialised Translation*, 21, pp.133-152.