

## **The role of age and proficiency in subtitle reading. An eye-tracking study**

**Carmen Muñoz**

**Universitat de Barcelona**

### **1. Introduction**

Research suggests that watching video material with the soundtrack in the second language (L2) and subtitles may improve L2 acquisition (e.g., Montero-Pérez, Van Den Noortgate, & Desmet, 2013). Studies have focused on different language dimensions, especially vocabulary and listening comprehension, but important issues concerning the adequacy of native language (L1) and second or foreign language<sup>1</sup> subtitling to different learner characteristics remain unexplored. In this respect, the use of eye tracking may provide promising insights, because it can shed light on how different learners read the screen text in the L1 and in the L2 (d'Ydewalle and De Bruycker, 2007; Winke, Sydorenko, & Gass, 2013 ).

The present study aims at exploring the effects of age and proficiency on the reading behaviour of foreign language learners by using eye-tracking methodology. These learner variables are of crucial importance for pedagogically oriented research concerned with the use of audiovisual material in the foreign language classroom. This study focuses on the two types of subtitling that foreign language learners are most likely to encounter in real-life situations (classroom, home)<sup>2</sup>: interlingual subtitling with the soundtrack in the L2 and the on-screen text in the L1 (L1 subtitles), and intralingual

subtitling with the soundtrack in the L2 and the subtitles in the same language (also known as same-language subtitles or captions).

## **2. Literature Review**

### *2.1. Subtitling and foreign language learning*

Ever since Price's (1983) pioneering study showed the benefits for L2 comprehension of L2 subtitles over non-subtitled audiovisual material, many studies have focused on this type of subtitling. The most frequently used explanation is the notion of bimodal reinforcement of sound and text: "...the double modal input may be processed more "deeply" because attention can alternate from the auditory to the visual format or be directed along parallel visual and auditory routes simultaneously" (Holobow, Lambert, & Sayegh, 1984, p.73). Similarly, Paivio's (1986) Dual Coding Principle explains that visual and verbal information are processed differently and along distinct channels, creating separate representations in the human mind. The activation of the verbal and imagery systems in processing leads to better recall, which is why L2 learning can be enhanced by combining visual images with verbal information. Based on this idea, Mayer's (2009) Multimedia Learning Principle states that when words are presented along with images, learning takes place through the formation of new mental representations or strengthening of existing ones; this is why people learn better from words and pictures than from words alone (p. 223).

#### *2.1.1. L2 and L1 subtitles*

The value of L2 subtitles has been demonstrated in studies focusing on productive and receptive skills (see Danan, 2004 for a review), on listening comprehension independently of written comprehension (e.g., Markham, 1999), on word recognition (e.g., Bird & Williams, 2002; Markham & Peter, 2003; Montero Pérez Peters, Clarebout, & Desmet, 2014; Rodgers, 2013), particularly the recognition of written word forms (though recognition of aural forms may be enhanced by non-subtitled video, see Sydorenko, 2010), and on speech segmentation (Charles & Trenkic, 2015). According to Winke, Gass, and Sydorenko (2010), L2 subtitles are beneficial for overall comprehension and vocabulary recognition because “they result in greater depth of processing by focusing attention, reinforce the acquisition of vocabulary through multiple modalities, and allow learners to determine meaning through the unpacking of language chunks” (p. 81). Similarly, L2 subtitles reduce learner-viewer anxiety (Danan, 2004; Winke et al., 2010), help learners ‘tune in’ to the language spoken and assist the development of listening skills over the longer term through frequent viewing (Lambert & Holobow, 1984; Vanderplank, 1988; 2010). In a meta-analysis, Montero Pérez and colleagues (2013) examined the effectiveness of L2 subtitles over non-subtitles for listening comprehension and vocabulary learning. Although the number of primary research studies in the meta-analysis was limited, it revealed a large effect of L2 subtitles on both listening comprehension (with a larger effect size for receptive than for productive tests) and vocabulary acquisition (with a large effect for both recognition and recall tests).

L2 subtitling has generally been claimed to be superior to L1 subtitling as well (see Danan, 2004). An insightful explanation of the finding that L2 subtitles may be more helpful than L1 subtitles is offered by Mitterer and McQueen (2009) in a study of word recognition in accented dialects of English by proficient Dutch users of English.

According to them, listeners use lexical knowledge to retune phonetic perception. When reading L2 subtitles, the phonological knowledge that is automatically retrieved is consistent (same language) with the English word forms, which facilitates speech perception. In contrast, when reading L1 subtitles (different language), the phonological knowledge that is also automatically retrieved creates lexical interference.

In turn, viewing programmes with L1 subtitles has been found to be more beneficial than without subtitles, one reason being that the effort of establishing paired equivalents through translation leads to deeper processing. Following Paivio's (1986) bilingual dual coding theory, Danan (2004, p.72) argues that in the case of L1 subtitled visual input "three independent systems are interconnected through triple associations between image, sound in one language, and text in another, which may lead to better processing and recall because of the additive effects of both image and translation". Furthermore, it has been demonstrated that reading of L1 subtitles does not prevent the processing of the soundtrack, regardless of the viewers' knowledge of the spoken language or their familiarity with subtitling (d'Ydewalle & Gielen, 1992; d'Ydewalle & Pavakanum, 1992).

A series of studies by Markham and colleagues (e.g., Markham & Peter, 2003) found an advantage of L1 subtitles over no subtitles in comprehension and vocabulary learning but, in contrast to other studies, they also found that their intermediate-level students who read L1 subtitles outperformed those who read L2 subtitles. Consequently, the authors suggest that L1 subtitles may be more suitable for low-level learners with a challenging video and that L2 subtitles may suit better high-level learners who may then progress to no subtitles. In the same line, Markham (1993) suggests that L2 subtitles are more helpful for intermediate and advanced learners when the video material is difficult.

### *2.1.2. Subtitles and learner characteristics*

Although the effect of learner proficiency in subtitling studies is under-researched, other studies have concurred with Markham and colleagues on highlighting the limitations of L2 subtitle use with lower-level learners. For example, Guillory (1998) found that if the material in the video is too advanced for the learners' proficiency level, L2 subtitles cannot sufficiently compensate for the fast rate of speech and the difficulty of the vocabulary. This author recommends the use of keywords rather than entire sentences for these learners, because L2 subtitles may impose a large cognitive load. Interviewing first, third, and fourth year college learners, Taylor (2005) observed that the beginner level learners reported difficulty in attending to sound, image, and L2 subtitles simultaneously, and found L2 subtitles distracting. Danan (2004) suggests that beginner level learners may not have reached the minimum language competency threshold to derive benefit from L2 subtitling and that the linguistic difficulty of the audiovisual material needs to be carefully matched to the learners' competency level. On this basis, Danan (2004) argues that L1 subtitling may be a good way of making incomprehensible visual input comprehensible. In contrast, Montero Pérez et al. (2013) conclude that L2 subtitling may be equally effective for all proficiency levels as long as the video materials match learners' actual level, a claim supported by other authors (e.g., Winke et al., 2010).

In his recent review of the research, Vanderplank (2016) notes several key variables regarding language proficiency which need to be taken into account in relation to the limits of L2 subtitles, the most important among them being reading speed. According to Vanderplank (2016), the problem for L2 viewing is not so much a matter of reading and listening comprehension, but the ability to follow (i.e., decode) what is

being said – an ability that obviously is limited in learners with low foreign language proficiency. L2 subtitles may support comprehension, but they require the ability to read and understand at reasonable speed. Both proficiency and reading speed would seem to limit the benefits of using captioned programmes with young learners, whose reading speed is still developing (Carver, 1990).<sup>3</sup> In this respect, Vanderplank observes that only at age 10 do children attain the reading speed needed to read subtitles in their L1.

In fact, in this area studies with children are scarce, and most of those published use L1 subtitles. Koolstra and Beentjes (1999) compared Dutch children in grades 4 and 6 watching a television programme with an English soundtrack and either Dutch subtitles or no subtitles at all (plus a control group to establish a baseline of English vocabulary knowledge). Vocabulary scores for those watching with subtitles were higher than for those watching without subtitles, and scores in this latter group were higher than those in the control group. Grade 6 children (following one year of English at school) outperformed children in grade 4 (with no previous instruction). However, high frequency of subtitled television viewing at home proved to have a more significant effect on performance. Several other studies with Dutch-speaking children have provided evidence of incidental learning of vocabulary from watching subtitled television programmes regularly over a long period. An example is the study by Kuppens (2010) who investigated the impact of long-term use of English language media in a large group of 11-year-old Flemish children. By means of oral tests (vocabulary, grammar, and direct and inverse translation of sentences), the study showed that children who “frequently watch subtitled English television programs and movies perform significantly better” (p. 65).

In a study with children of two different age groups, d'Ydewalle (2002) observed that the younger children (grade 4) tended to ignore the foreign spoken soundtrack when the L1 was made available in the subtitles, whereas the older children (grade 6) seemed to have sufficient resources to process both sources of information. Thus, the latter resembled secondary school students, who have also been observed to be sufficiently equipped to make use of the audio input and the subtitles (De Bot et al., 1986).

In sum, while studies on subtitle reading indicate benefits for L2 learning, there are still important gaps in relation to variables that are of crucial pedagogical relevance – such as how learners read the screen text, in the native language and the foreign language, as a function of their age and proficiency level. The use of the eye-tracking methodology may shed new light on these important issues.

## *2.2. Eye-tracking studies of subtitle reading*

The use of eye tracking in the investigation of reading has helped to enrich the knowledge of some key issues in L2 acquisition and processing (for a review, see Roberts & Siyanova-Chanturia, 2013). Basically, eye-tracking recordings are a real-time, online, direct measure of processing. Thus, during reading, rapid eye movements (*saccades*) are made from one point to another from left to right (for languages that are read from left to right), but also sometimes from right to left (*regressions*). In between these movements, the eyes become stable for as long as is needed to take in visual information. These stops are called *fixations*, but not all words are fixated in the same way, and some are not fixated at all. For example, short, frequent words like function

words are often skipped (Carpenter & Just, 1983); highly predictable words are skipped more frequently and are characterized by shorter fixation duration than words that are less contextually constrained (e.g., Rayner & Well, 1996). The measurement of these eye movements has provided very rich information about the cognitive processes involved in reading. Crucially, fixation duration corresponds to the duration of cognitive processing of the material located at fixation (Irwin, 2004; Rayner, 1998); that is, when reading is conceptually more complex, eye fixations become longer, indicating difficulties in cognitive processing (Duchowski, 2002) and higher cognitive load.

Research on subtitling reading using eye tracking is still very scarce, though the results reported so far are promising. Past research has focused mainly on viewers reading in their L1 or an unknown language and very few studies have investigated the eye movements of foreign language learners; to our knowledge, none has focused on the effects of learner age or proficiency. Issues that have been investigated with eye-tracking methodology include the differential effects of L2 subtitles across languages (for example, when the target language has the same or different script) or the amount of time (intermediate-level) foreign language learners spend reading L2 subtitles in different languages (Winke et al., 2010; Winke et al., 2013). Attention has also been paid to differences in reading patterns of L1 and L2 subtitles, suggesting that viewers skip subtitles in the native language more often than in the foreign language, which indicates a lower degree of processing when subtitles are redundant in the L1 (see Kruger, Hefer, & Matthew, 2014). Bisson, Van Heuven, Couklin, & Tunney (2014) used eye tracking to investigate the extent to which people process subtitles in different conditions in a group of English-L1 college students who had no knowledge of Dutch. The authors predicted that L1 subtitles would cause viewers to have more and longer fixations than L2 subtitles. However, in their study no significant differences were



found between the two conditions for total fixation duration, number of fixations, and skipped subtitles.

The reading behaviour of children with L1 subtitled material was investigated in studies by d'Ydewalle and his colleagues. D'Ydewalle and De Bruycker (2007) compared the reading behaviour of children in grades 5 and 6 and adults, and found that children's fixation durations were longer, their saccade amplitude was smaller, and their latency to jump to the subtitle was longer. Despite these variations, children did not show a radically different pattern from adults. However, in that study the target language (Swedish) was unknown to both age groups, so foreign language proficiency did not play a role in the results. D'Ydewalle and Van Rensbergen (1989) investigated eye movements of Dutch-speaking children in grades 2, 4, and 6 reading subtitles in their native language. They reported that children in grade 2 did not always read the subtitles, and that their reading depended on the degree to which the movie was verbally loaded.

As regards L2 proficiency, its influence has not been systematically explored in eye-tracking studies concerned with subtitling. In a recent study by Laskowska, Szarkowska, Pilipczuk, & Oliver (2015) adult participants viewed videos in English and in Norwegian with Polish (L1) subtitles. Their self-reported proficiency level in English was much higher than in Norwegian. It was found that the percentage of subtitles looked at when the soundtrack was in Norwegian was higher and that the mean fixation durations on (L1) subtitles were longer when the soundtrack was in English. The authors interpret this higher processing time as indicating that participants were comparing the content of the subtitle with the English audio. However, this study did not compare different proficiency levels within the same target language, but across languages and with subtitles in the participants' native language.

In sum, this review has identified important gaps in the limited body of research that has used eye-tracking methods with foreign language learners viewing L1 or L2 subtitled material, particularly in terms of learner age and proficiency. These learner variables may be of crucial relevance for pedagogically oriented research concerned with the use of audiovisual material in the foreign language classroom, particularly at a time when foreign language teaching/learning (in most cases English) has extended to primary schools all over the world. The present exploratory study addresses these research gaps and aims to shed light on the reading behaviour of learners of different ages and proficiency levels while viewing videos with L2 soundtrack and subtitles in the L1 or in the L2.

The following research questions underpinned the current study:

1. Does age influence the way individuals read L1 and L2 subtitles with an L2 soundtrack? Does this influence differ when the subtitles are in L1 and in L2?
2. Does L2 proficiency influence the way individuals read L1 and L2 subtitles with an L2 soundtrack? Does this influence differ when the subtitles are in L1 and in L2?

### **3. Methodology**

#### *3.1. Participants*

The sample comprised 40 Spanish-Catalan learners of English (27 females), distributed into three age groups: children, adolescents and adults. The children group comprised 19 learners in grade 5 or 6 of primary school, ranging in age from 10 to 12.7 (mean age 11.1), the adolescent group nine learners (age range 13.1 to 16, mean age

14.6), and the adult group 12 learners (age range 19 to 41, mean age 25.8). Three proficiency levels were established following the Common European Framework of Reference (CEFR): the beginner group (A1, A2) comprised 18 learners, the intermediate group (B1, B2) twelve, and the advanced group (C1, C2) ten. Information about participants' proficiency level was gathered for the primary school participants and the university students through teacher assessment in recent examinations. Because secondary school participants came from different schools it was considered more accurate to use the same instrument to measure their proficiency level. A cloze test that had been used in previous research in this learner population with satisfactory results was employed (see AUTHOR, 2006). Their performance was then assigned to a CEFR level by a teacher familiar with the age group and proficiency levels. Although there was a large overlap between age and proficiency grouping — most children (18) were in the beginner group, and most adults (10) in the advanced group — the intermediate group comprised nine adolescents, one child and two adults.

### *3.2. Materials*

The two clips used in this study were extracts from two different episodes of “The Simpsons”, a popular cartoon series with which all participants were familiar. Both clips had an English audio and were shown with either English or Spanish subtitles.<sup>4</sup> In order to maximize the relevance of this study to the typical classroom situation in which teachers have limited time to prepare materials, subtitles were not manipulated or shortened. The cartoon nature of the clips ensured that the language in the subtitles was not very complex. In the first clip, 83.64% of words (tokens) were among the most frequent 1000 words and 89.42% among the most frequent 3000 words;

in the other clip, the proportions were 84.53% and 91.85% respectively. The length of the two extracts differed (4:17 and 6:05 respectively with English subtitles, and 4:23 and 6:11 respectively with Spanish subtitles). The two clips in the English version had 114 and 134 subtitles, and the two clips in the Spanish version 120 and 140. The number of words in the subtitles also differed, ranging from 1049 to 750 words.

A questionnaire was filled out by the participants during the recording session. The first part elicited background information and was completed at the beginning of the session. The second part was completed after the participants had viewed the first of the two clips, and started by asking them about their perceptions of the degree of difficulty and effort during the viewing task<sup>5</sup>. Then, they were asked four simple comprehension questions about the content of the clip which required very short and easy answers. The third part was completed after the participants' viewing of the second clip and was identical to the previous one.

### *3.3. Procedure*

The data from the youngest group was collected on the students' school premises, whereas the adolescent and adult data were collected in an office at the university. For the youngest group consent was obtained from the parents through the school, and the older participants signed a consent form. As noted above, all participants filled out a questionnaire during the session. They were informed at the start that they would be asked comprehension questions about the clips to ensure that they kept their attention focused on the videos.

Each participant met individually with two researchers, one in charge of the eye-tracking device and the second in charge of the questionnaire. After filling out the

background section of the questionnaire, participants were given a brief explanation about eye tracking and the procedure to be followed. They were asked to watch the clips in the same way as they would do if they were at home, that is, as naturally as possible. They were informed that the soundtrack was in English in both clips but one had English subtitles and the other Spanish subtitles. The order of presentation of the L2 and L1 subtitles was counterbalanced.

For the eye-tracking recording, participants sat in a chair about 50 cm away from the eye-tracking screen. Before recording, a 5-point calibration routine was performed. Participants then watched one of the clips with either English or Spanish subtitles. When this finished, they filled out the questions about the clip from the second part of the questionnaire. Before watching the second clip, their responses were checked to ensure that they had engaged with the task and that comprehension was satisfactory. There was only one (adult) participant in whom this was not the case and she was asked to watch the first clip again. Before watching the second clip, the participants again followed the standard calibration procedure. Then, after watching the second clip, they answered the corresponding questions from the questionnaire. This time, all their responses provided evidence that they had engaged with the task. The session lasted about 30 minutes for each participant. No participant had to be excluded during the data collection process or due to poor data quality.

### *3.4. Analysis*

Clips were presented in a 1280 x 1024 resolution on a 17" TFT monitor of a Tobii T120 integrated eye-tracker (Tobii Publ AB, Stockholm, Sweden). Eye-tracking sampling rate was 120Hz, which equals a temporal resolution of 8 ms. Raw gaze data

for both eyes was later filtered into fixations using the I-VT filter, a fixation classification algorithm that takes into account the velocity at which the eyes move, identifying fixations when the velocity of eye movements is below the threshold of 30 degrees/second and the gaze remains in the same location for at least 60 ms. This non-intrusive eye tracker allows a large freedom of head movement (30x22x30cm).

Fixations were analysed using Tobii Studio software. To compute the fixation metrics in the relevant areas of the scene, dynamic areas of interest (AOIs)<sup>6</sup> were created for each subtitle with an area of approximately 775 x 125 pixels (slightly larger than the exact borders of the text in order to account for possible small horizontal or vertical inaccuracies in the recording of eye movements; see Bisson et al., 2014). The AOIs were synchronized with the dialogue of the clips, and, because there was almost constant dialogue, AOIs remained active most of the time. The following eight measures were calculated for the subtitle area for each clip: percentage of Spanish subtitles skipped (SKPL1), percentage of English subtitles skipped (SKPL2), total fixation count on Spanish text (FCL1), total fixation count on English text (FCL2), average fixation duration on Spanish text (AFDL1), average fixation duration on English text (AFDL2), total fixation duration on Spanish text (TFDL1) and total fixation duration on English text (TFDL2).

#### **4. Results**

Because of the importance of showing a fragment that was meaningful on its own, the length of the two extracts was not identical and the raw data were normalized. Thus, the fixation count for each participant was divided by the number of words in all subtitles, and a normalized value of the fixation count was obtained. Similarly, fixation

duration and total fixation duration of each participant were divided by the number of words per second in each subtitle language. The fixation counts for individual subtitles were used to calculate the number of subtitles skipped (subtitles that participants did not fixate upon) as a percentage of the total number of subtitles in each clip and condition. Finally, the mean value for the individual normalized data was computed for each measure and condition.

Shapiro-Wilk normality tests were run and two outliers were detected for certain measures. These values were changed to the next highest/lowest (non-outlier) values plus one unit increment higher/lower. The presence of those outliers and the small sample size recommended using non-parametric tests (the results did not substantially differ from the ones obtained with parametric tests).

#### *4.1. Age and eye-movement behaviour*

A Kruskal-Wallis test was run with the eight dependent variables. In all cases, age group was the independent variable, with three levels. Significant differences were revealed on all measures except for average fixation duration on L1 and L2 subtitles and total fixation duration on L1 subtitles (SKPL1:  $\chi^2(2) = 8.23, p = .016$ ; SKPL2:  $\chi^2(2) = 8.63, p = .013$ ; FCL1:  $\chi^2(2) = 10.91, p = .004$ ; FCL2:  $\chi^2(2) = 9.37, p = .009$ ; TFDL2:  $\chi^2(2) = 9.50, p = .009$ ). See descriptive statistics in Table 1 (Appendix).

To further explore the differences in viewing patterns between the three groups, Mann-Whitney U tests were run ( $\alpha = .016$ , Bonferroni corrected). Children displayed a lower percentage of skipped subtitles in L1 and L2 than adolescents (SKPL1:  $U = 36, p = .014, r = -.46$ ; SKPL2:  $U = 33, p = .009, r = -.49$ ) and than adults (SKPL1:  $U = 58.50, p = .023, r = -.41$ ; SKPL2:  $U = 61.50, p = .032, r = -.39$ ); the  $p$  values for the latter were

almost significant but the effect sizes were medium. Significant differences were also found on the measures of fixation count in L1 and L2 between children and adolescents and adults, with fairly large effect sizes. Children fixated significantly more times on L1 and L2 subtitles than adolescents (FCL1:  $U = 35$ ,  $p = .012$ ,  $r = -.47$ ; FCL2:  $U = 34$ ,  $p = .010$ ,  $r = -.48$ ), and adults (FCL1:  $U = 43$ ,  $p = .003$ ,  $r = -.52$ ; FCL2:  $U = 52.50$ ,  $p = .011$ ,  $r = -.45$ ). Significant differences were also revealed on the measures of total fixation duration between the group of children and the two older groups, as above with fairly large effect sizes. Children spent more time on L2 subtitles than adolescents (TFDL2:  $U = 38$ ,  $p = .019$ ,  $r = -.44$ ); the  $p$  value after Bonferroni correction was almost significant but the effect size was medium. Children also spent more time on L1 and L2 subtitles than adults (TFDL1:  $U = 63$ ,  $p = .039$ ,  $r = -.37$ ; TFDL2:  $U = 48$ ,  $p = .006$ ,  $r = -.48$ ). In the former case the  $p$  value approached significance and the effect size was medium. On the other hand, no significant differences were found in any measure between the group of adolescents and the group of adults.

To compare eye movements between L1 and L2 modes, a Wilcoxon Signed Ranks test was run for the whole group. Significant differences were found in only two measures: percentage of skipped subtitles and total fixation duration, with fairly large effect sizes. The percentage of subtitles skipped was significantly higher in the L1 condition ( $Z = -2.64$ ,  $p = .008$ ,  $r = -.42$ ) and, conversely, total fixation time was significantly higher on L2 subtitles than on L1 subtitles ( $Z = -3.29$ ,  $p = .001$ ,  $r = -.52$ ). Dividing the sample into age groups, differences were found only for the youngest and oldest groups. Children spent significantly more time on subtitles in L2 than in L1 ( $Z = -2.66$ ,  $p = .008$ ,  $r = -.61$ ), and adults displayed a higher percentage of skipped subtitles in L1 than in L2 ( $Z = -2.22$ ,  $p = .026$ ,  $r = -.64$ ). Note that although the  $p$  value of the last



comparison did not reach significance (after Bonferroni correction) the effect sizes of both differences were large.

#### 4.2. Proficiency and eye-movement behaviour

Similar procedures were followed with proficiency level as the independent variable with three levels: beginner, intermediate, and advanced. A Kruskal-Wallis test run revealed significant differences ( $p < .05$ ) in four measures (SKPL1:  $\chi^2(2) = 8.19, p = .017$ ; SKPL2:  $\chi^2(2) = 7.20, p = .027$ ; FCL1:  $\chi^2(2) = 9.95, p = .007$ ; TFDL2:  $\chi^2(2) = 6.27, p = .043$ ). Differences approached significance in two measures (FCL2:  $\chi^2(2) = 5.92, p = .052$ ; TFDL1:  $\chi^2(2) = 4.89, p = .087$ ). No significant differences were found in average fixations on L1 or L2 subtitles. Descriptive statistics for the three proficiency groups can be consulted in Table 2 (Appendix).

Next, to further explore the differences revealed by the Kruskal-Wallis test, two Mann-Whitney U tests were run, between beginner- and intermediate-level participants and between beginner- and advanced-level participants. The values of intermediate- and advanced-level participants were very similar and no further examination was conducted ( $\alpha = .025$ , Bonferroni corrected). Beginners displayed a significantly lower percentage of skipped L1 and L2 subtitles than intermediate-level participants (SKPL1:  $U = 47.50, p = .009, r = -.47$ ; SKPL2:  $U = 50.50, p = .013, r = -.45$ ). Beginners also skipped fewer subtitles than advanced-level participants. Differences did not reach significance after Bonferroni correction but the effect sizes were medium to large (SKPL1:  $U = 45.50, p = .031, r = -.40$ ; SKPL2:  $U = 50.00, p = .057, r = -.37$ ). Beginners fixated significantly more on L1 subtitles (FCL1:  $U = 41.00, p = .004, r = -.52$ ) than intermediate-level participants. Their number of fixations on L2 subtitles was also

higher (FCL2:  $U = 60.50$ ,  $p = .043$ ,  $r = -.37$ ) and although the difference was not significant ( $p > .025$ ) the effect size was medium. Similar results were found when beginners were compared with advanced-level participants (FCL1:  $U = 41.00$ ,  $p = .018$ ,  $r = -.44$ ; FCL2:  $U = 48.50$ ,  $p = .045$ ,  $r = -.38$ ). Differences were revealed in the measures of total fixation duration on L1 and L2 subtitles between beginners and intermediate-level participants with medium effect sizes although they did not reach significance after Bonferroni correction (TFDL1:  $U = 60.00$ ,  $p = .043$ ,  $r = -.37$ ; TFDL2:  $U = 61.00$ ,  $p = .048$ ,  $r = -.36$ ). A difference in the measure of total fixation duration on L2 subtitles was also revealed when comparing beginners and advanced-level participants with a medium to large effect size, though it was not significant after Bonferroni correction either (TFDL2:  $U = 46.00$ ,  $p = .035$ ,  $r = -.40$ ). No significant differences were found in the measures of average fixation duration between any of the groups.

Finally, the sample was divided into proficiency groups to compare eye movements between L1 and L2 modes. The Wilcoxon Signed Ranks test revealed significant differences in only two measures: percentage of skipped subtitles and total fixation duration (see above). It was found that beginner- and intermediate-level participants spent more time on subtitles in L2 than in L1 ( $Z = -2.20$ ,  $p = .028$ ,  $r = -.52$ ;  $Z = -2.27$ ,  $p = .023$ ,  $r = -.66$ ). It was also found that intermediate- and advanced-level participants skipped more subtitles in L1 than in L2 ( $Z = -2.04$ ,  $p = .041$ ,  $r = -.59$ ;  $Z = -1.84$ ,  $p = .07$ ,  $r = -.58$ ). Although the values only approached significance after Bonferroni correction for multiple comparisons ( $\alpha = .016$ ), the effect sizes are fairly large.

## **5. Discussion**

This study addressed two research questions concerning the differences in subtitle reading behaviour between learners of different ages and proficiency levels. The first research question investigated the possible influence of age on eye movement behaviour when watching audiovisual material with L1 and L2 subtitles. To answer this question, the eye-movement behaviour of participants in three age groups – children, adolescents and adults – was analysed using eight eye-tracking measures. No differences were found between the two older groups in any of the measures. In contrast, children differed from the older groups in most measures and the effect size ranged from medium to large. The primary school children skipped subtitles much less than adolescents and adults. Children also made more fixations on subtitles and spent a longer total time on them than adolescents and adults.

The finding that children make more eye fixations than the older participants may be related to several factors. First, on the basis of research findings in static reading it might be suggested that these children, aged 10-12, had a slower reading speed (Häikiö, Bertram, Hyönä, & Neimi, 2009) and a smaller perceptual span than older participants, which led them to make more eye fixations than the older and faster readers (Rayner, Slattery, & Bélanger, 2010). Second, the children's higher number of fixations may be indicative of more cognitive effort or processing difficulty<sup>7</sup>. Third, it may indicate that these learners rely more on subtitles than older learners, who do not need to read so much because they understand the auditory input better. Children's limited proficiency may also be an appropriate explanation for the finding that they skip fewer subtitles because they use them for support with the foreign language soundtrack. The fact that the children had a lower L2 proficiency level than the adults may also partly explain the divergence of the current result from the findings of d'Ydewalle and

De Bruycker (2007), who reported no such differences between children and adults in a study in which the foreign language was equally unknown to both groups.

As regards differences due to language mode, the analyses reveal that children spent more total time on subtitles in L2 than in L1, as a result of their higher number of fixations in the former than in the latter. This suggests that, for children, reading in the L2 may be more cognitively challenging<sup>8</sup>. On the other hand, adults were seen to skip more L1 subtitles than L2 subtitles, which corroborates the results in Kruger et al. (2014), indicating a lower degree of processing when subtitles are redundant in the L1. This finding may indicate that adults, who in this study had a higher proficiency level, do not need L1 subtitles for understanding, especially when the language of the text was not complex. Further, at least some of the adults in the study were advanced learners/users of English who may have trained themselves to rely less on the support provided by subtitles in their L1.

The second research question dealt with differences in subtitle reading behaviour between beginner, intermediate and advanced learners of English. Effect sizes were mostly medium to large. Beginners in the study were seen to make more fixations on L1 subtitles than intermediate and advanced learners and more fixations on L2 subtitles, although in this case the differences were not so large. Beginners' total time spent on fixations was also greater than that of intermediate and advanced learners but differences did not reach significance. Beginners were also found to skip L1 and L2 subtitles less than intermediate and advanced level learners. This finding provides an affirmative answer to the question posed by Winke and colleagues (2010) about whether upper-level learners ignore L2 subtitles more often than lower-level ones. The current study also established that the former ignore L1 subtitles more often than the latter.

Differences due to subtitle language mode were also found between proficiency groups, but they were more nuanced than when participants were distributed into age groups. Beginner- and intermediate-level learners spent more total time on subtitles in L2 than in L1, which suggests that the task was more challenging for them than for advanced learners.<sup>9</sup> This finding corroborates those of Taylor (2005) whose beginner participants found it difficult to attend to sound, image and L2 subtitles simultaneously and even found L2 subtitles distracting (see also Danan, 2004). It was also found that intermediate and advanced learners skipped more subtitles in L1 than in L2, corroborating the results in Kruger et al. (2014) and in Laskowska and colleagues (2015). However, the finding is at odds with the idea that L1 subtitles are read automatically (d'Ydewalle & Gielen, 1992). It is interesting to see that the current results differ from those in the study by Bisson and colleagues (2014), who reported no differences between the L1 and the L2 conditions. However, in their experiment participants were non-language learners and thus proficiency differences could not emerge (nor age differences, since they were all adults).

These proficiency-related differences only partially mirror age-related differences with respect to subtitle language mode. On the one hand, intermediate learners behave like beginners as regards their spending more time on L2 subtitles, which may be related to the still limited size of their vocabulary. On the other hand, intermediate learners behave like advanced learners with regard to their higher percentage of skipped subtitles in the L1 than in the L2, which appears to indicate a sufficient understanding of the soundtrack, possibly together with a somewhat negative attitude towards reading subtitles in the native language.

An unexpected result of this study was that no differences were found in the measure of average fixation duration between the age groups or the proficiency groups

in either mode. On the basis of evidence from L1 reading research, the young beginner learners might have been expected to spend more time on each fixation because of their slower reading rate (Rayner et al., 2010) and associated higher cognitive load (Duchowski, 2002). This was the result found in the study by d'Ydewalle and De Bruycker (2007) where children had longer fixations than adults, although there were no differences in proficiency because the language was equally unknown to both groups. In contrast, the study by Laskowska and colleagues (2015) with Polish adults who had different proficiency levels (in English and in Norwegian) found that the mean fixation duration was longer in the language in which participants were more proficient (English). This finding was interpreted as indicating that participants needed longer processing time to compare the content of the subtitle with the English audio. In contrast, and following from that explanation, it might be suggested that the young beginner learners in the present study did not attempt to compare the content of the subtitle with the English audio because of their perceived proficiency limitations. If this was the case, the learning benefits of subtitles may be more limited with beginner children. The impact of (perceived) proficiency on subtitle reading behaviour, and fixation duration in particular, is clearly an issue that needs further investigation.

In sum, this study found similar results in most age and proficiency comparisons, which was not surprising given the large overlap between the children and beginner groups at one extreme and the adults and advanced proficiency groups at the other, although the effect sizes tended to be larger when age groups were compared. Further, adolescents and adults exhibited similar eye movement behaviour, while intermediate-level participants did not differ from advanced-level participants in frequency of skipped subtitles although they were more similar to beginners in number of fixations. Slightly different results were also obtained in relation to language mode, and again age

comparisons revealed more differences than proficiency comparisons, suggesting perhaps that young age (associated with low proficiency) may have a stronger impact on eye movement behaviour than proficiency.

The findings of the current study have several pedagogical implications with regard to the choice of audiovisual material for learners of various ages and proficiency levels. Namely, arguments put forward in the literature concerning the use of L1 subtitles with children (see Vanderplank, 2016) have been confirmed through the use of the eye-tracking methodology, which showed that L1 subtitles may be more appropriate than L2 subtitles when reading speed is slow. Reading L2 text is a harder task for children and beginner learners when their vocabulary size is more limited. Alternatively, L2 subtitles might be manipulated (shorter length and longer time on screen) but this possibility is not always available to teachers. Adolescents and adults with higher proficiency levels cope better with L2 subtitles as a way to aid L2 learning.

## **6. Conclusions and further research**

The aim of this exploratory study was to investigate the effects of age and proficiency on subtitle viewing behaviour, given the potential of subtitles for second/foreign language learning. To our knowledge, this is the first study to explore differences in eye behaviour between three age groups (children, adolescents and adults) and three proficiency levels (beginner, intermediate, advanced) when reading subtitles in L1 and in L2 with L2 soundtrack.

The study has several limitations that further research should address. The first is the large overlap between age and proficiency groupings (which mirrors typical instruction situations). Further research could compare beginner- and advanced-level

adults. At the beginner/children extreme, one possibility would be to compare the viewing behaviour of fully bilingual children with that of foreign language learners of the same age. Another limitation is that the proficiency level of the participants could not be measured with the same instrument because of the large diversity in ages and proficiency levels. Also, the materials used in the study may not have been challenging enough for the advanced learners and may have biased their viewing behaviour away from a reliance on subtitles. Finding audio-visual material that is both appropriate across groups and attractive is not easy, but further research with different material is necessary in order to obtain more robust evidence. The two last limitations are common to most eye-tracking studies. One is that, because of the small sample size, findings can only be taken as indicative and must be confirmed by future research. In this study in particular, the small size of the adolescent group is an important limitation. The other limitation concerns the need for a more meticulous mechanism to distinguish mere attention allocation from reading behaviour. In this regard, the index proposed by Kruger and Steyn (2014) for the quantification of reading of dynamic texts seems promising because it provides a measure of the visual processing of subtitles that differentiates between the impact of seeing a video with subtitles and reading the subtitles from seeing the same video with subtitles but without reading them. It also seems a better alternative to a traditional measure of reading speed taken from static reading in which the reader does not have to adjust the pace of reading to the pace of presentation, as s/he must do when reading subtitles.

To conclude, this study contributes to the area of foreign language subtitling at a time when teaching activities need to accommodate, and capitalize on, the fact that classroom learners are “digital natives” (Prensky, 2001). As such, they will be more positively oriented towards audio-visual activities, which will enhance their learning.



Crucially, they will also benefit from the increased and optimized exposure to the target language provided by multimodal input in the classroom and will be encouraged to take advantage of language learning experiences of this kind outside the classroom as well.

---

<sup>1</sup> No distinction is made in this paper between the terms “foreign language” and “second language”.

<sup>2</sup> A third type of subtitling, reversed subtitling, has the soundtrack in the L1 and subtitles in the L2.

<sup>3</sup> According to Carver (1990), from grade 2 to college, reading rate increases on average 14 standard-length words per minute each year (one standard-length word defined as six characters in text, including punctuation and spaces).

<sup>4</sup> Due to the strong dubbing tradition in Spain, foreign television audiovisual programmes are commonly dubbed into Spanish or Catalan (on Catalan TV channels). However, when films in cinemas in Catalonia are shown in a subtitle format (very seldom), subtitles in Spanish are much more common than in Catalan, so the choice of subtitles in Spanish for the present study was not a marked choice for our participants. In fact, the percentage of films dubbed or subtitled in Catalan in cinemas in 2013 was only 3% of the total. [https://www.plataforma-llengua.cat/media/upload/pdf/informecat-2015\\_1430810660.pdf](https://www.plataforma-llengua.cat/media/upload/pdf/informecat-2015_1430810660.pdf)

<sup>5</sup> Due to space limitations, this paper does not report on these and other data provided by the participants through the questionnaire (e.g., familiarity with different viewing modes, incidental learning of vocabulary during the viewing). But see notes 7 to 9.

<sup>6</sup> AOIs define the regions of the audio-visual content displaying elements of interest such as subtitles or characters. Making the areas dynamic, it was possible to compute eye fixations to a particular subtitle or any other element regardless of their on-off behaviour and changes in size and position of any element over time, allowing for a more precise monitoring of attention allocation.

<sup>7</sup> Children’s perception of effort, as reported in the questionnaire, was higher than that of adults in both language modes. It was higher when subtitles were in English (see note 8).

<sup>8</sup> The fact that children were more challenged when subtitles were in L2 was confirmed by their own reported perceptions in the questionnaire. Children’s perception of effort when subtitles were in L2 was much higher (5.47 on a scale of 1 to 9) than in L1 (3.53). In contrast, adult learners’ perception of effort with L2 subtitles was much lower (2) and not so different from that of L1 subtitles (1.83).

<sup>9</sup> The difference in perception of effort with respect to L2 and to L1 subtitles had an inverse relationship with proficiency: the more proficient the smaller the difference. The pattern was similar to that in the previous note.

## References

- Bird, S. A. & Williams, J. N. (2002). The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling. *Applied Psycholinguistics*, 23, 509–533.

- Bisson, M.-J., Van Heuven, W. J. B., Couklin, K. & Tunney, R. J. (2014). Processing of native and foreign language subtitles in films. An eye tracking study. *Applied Psycholinguistics*, 35(2), 399–418.
- Carpenter, P. & Just, M. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye movements in reading, Perceptual and language processes* (pp. 275-307). San Diego, CA: Academic Press.
- Carver, R. P. (1990). *Reading rate, a review of research and theory*. Boston: Academic Press.
- Charles, T., & Trenkic, D. (2015). The effect of bi-modal input presentation on second language listening, The focus on speech segmentation. In Y. Gambier, A. Caimi & C. Mariotti (Eds.), *Subtitles and Language Learning*. Peter Lang.
- Danan, M. (2004). Captioning and Subtitling, Undervalued Language Learning Strategies. *Meta*, 49(1), 67–77.
- De Bot, K., Jagt, J. Janssen, H., Kessels, E. & Schils, E. (1986). Foreign television and language maintenance. *Second Language Research*, 2(1), 72–82.
- Duchowski, A. T. (2002). A breath-first survey of eye tracking applications. *Behavior Methods, Research, Instruments, and Computers*, 1, 1–15.
- d’Ydewalle, G. (2002). Foreign-language acquisition by watching subtitled television programs. *Journal of Foreign Language Education and Research*, 12, 59–77.
- d’Ydewalle, G., & De Bruycker, W. (2007). Eye movements of children and adults while reading television subtitles. *European Psychologist*, 12, 196–205.
- d’Ydewalle, G., & Gielen, I. (1992). Attention allocation with overlapping sound, image, and text. In K. Rayner (Ed.), *Eye Movements and Visual Cognition, Scene Perception and Reading* (pp. 414-427). New York: Springer-Verlag.

- d'Ydewalle, G., & Pavakanum, U. (1992). Watching foreign television programs and language learning. In F. Engels, D. Bouwhuis, T. Bösser & G. d'Ydewalle (Eds.), *Cognitive Modelling and interactive environments in language learning* (pp. 193–198). Berlin: Springer-Verlag.
- d'Ydewalle, G., & Van Rensbergen, J. (1989). Developmental studies of text-picture interactions in the perception of animated cartoons with text. In H. Mandl & J. R. Levin (Eds.), *Knowledge acquisition from text and pictures* (pp. 233–248). Amsterdam: Elsevier Science Publishers.
- Guillory, H. G. (1998). The effects of keyword captions to authentic French video on learner comprehension. *Calico Journal*, 15(1/3), 89–108.
- Häikiö T, Bertram R, Hyönä J., & Neimi P. (2009). Development of the letter identity span in reading: Evidence from the eye movement moving window paradigm. *Journal of Experimental Child Psychology*, 102, 167–181. doi: 10.1016/j.jecp.2008.04.002.
- Holobow, N., Lambert, W. E., & Sayegh, L. (1984). Pairing script and dialogue, Combinations that show promise for second or foreign language acquisition. *Language Learning*, 34(4), 59–74.
- Irwin, D. E. 2004. Fixation location and fixation duration as indices of cognitive processing. In J.M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action, Eye movements and the visual world* (pp. 105–133). New York, NY: Psychology Press.
- Koolstra, C. M., & Beentjes, J. W. J. (1999). Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home. *Educational Technology Research & Development*, 47(1), 51–60.

- Kruger, J.-L., Hefer, E., & Matthew G. (2014). Attention distribution and cognitive load in a subtitled academic lecture, L1 vs. L2. *Journal of Eye Movement Research*, 7(5), 1–15.
- Kruger, J.-L., & Steyn, F. (2014) Subtitles and Eye Tracking, Reading and Performance. *Reading Research Quarterly*, 49(1), 105–120.
- Kuppens, A.H. (2010). Incidental foreign language acquisition from media exposure. *Learning, Media and Technology*, 35(1), 65–85.
- Lambert, W. E., & Holobow, N. H. (1984). Combinations of printed script and spoken dialogues that show promise for beginning students of a foreign language. *Canadian Journal of Behavioural Science*, 16, 1–11.
- Laskowska, M., Szarkowska, A., Pilipczuk, O., & Oliver, A. (2015). The influence of the original language of the film on subtitle reading. Presentation at the III Polish Eyetracking Conference, 5-6 March 2015.
- Markham, P. (1993). Captioned television videotapes, Effects of visual support on second language comprehension. *Journal of Educational Technology Systems*, 29(3), 183–191.
- Markham, P. (1999). Captioned videotapes and second-language listening word recognition. *Foreign Language Annals*, 32(2), 321–328.
- Markham, P., & Peter, L. (2003). The influence of English language and Spanish language captions on foreign language listening/reading comprehension. *Journal of Educational Technology Systems*, 31(1), 331–341.
- Mayer, R. E. (2009). *Multimedia Learning*. Second Edition. New York: Cambridge University Press.

- Mitterer, H., & McQueen J. M. (2009). Foreign subtitles help but native-language subtitles harm foreign speech perception. *Plos One* 4(11) doi, 10.1371/journal.pone.0007785.
- Montero Pérez, M., Peters, E., Clarebout, G., & Desmet, P. (2014). Effects of captioning on video comprehension and incidental vocabulary learning. *Language Learning and Technology*, 18(1), 118–141.
- Montero Pérez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning, A meta-analysis. *System*, 41, 720–739.
- Muñoz, C. (Ed.) 2006. Age and the rate of foreign language learning. Clevedon: Multilingual Matters.
- Paivio, A. (1986). *Mental representations, A dual coding approach*. Oxford: Oxford University Press.
- Prensky, M. (2001). Digital Natives, Digital Immigrants. *On the Horizon*, 9(5), 1–6.
- Price, K. (1983). Closed-captioned TV, An untapped resource. *MATESOL Newsletter*, 121, 1–8.
- Rayner, K. (1998). Eye movements in reading and information processing, 20 years of research. *Psychological Bulletin*, 124, 372–422.
- Rayner, K., Slattery, T. J., and Bélanger, N. N. (2010). Eye movements, the perceptual span, and reading speed. *Psychonomic Bulletin & Review*, 17(6), 834–839. doi: [10.3758/PBR.17.6.834](https://doi.org/10.3758/PBR.17.6.834)
- Rayner, K., & Well, A. (1996). Effects of contextual constraints on eye movements in reading, A further examination. *Psychonomic Bulletin and Review*, 3, 504–509.
- Roberts, L., & Siyanova-Chanturia, A. (2013). Using eye-tracking to investigate topics in L2 acquisition and L2 processing. *Studies in Second Language Acquisition*, 35, 213–235.

- Rodgers, M. (2013). *English language learning through viewing television, An investigation of comprehension, incidental vocabulary acquisition, lexical coverage, attitudes, and captions*. Doctoral dissertation. Victoria University of Wellington.
- Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning Technology, 14*(2), 50–73.
- Taylor, G. (2005). Perceived processing strategies of students watching captioned video. *Foreign Language Annals, 38*(3), 422–427.
- Vanderplank, R. (1988). The value of teletext subtitles in language learning. *ELT Journal, 42*(4), 272–281.
- Vanderplank, R. (2010). “Deja Vu? A Decade of Research on Language Laboratories, Television and Video in Language Learning”. *Language Teaching, 43*(1), 1–37.
- Vanderplank, R. (2016). ‘Effects of’ and ‘effects with’ captions. How exactly does watching a TV programme with same-language subtitles make a difference to language learners? *Language Teaching, 49*(2), 235–250.
- Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language learning activities. *Language Learning & Technology, 14*(1), 65–86.
- Winke, P., Sydorenko, T., & Gass, S. (2013). Factors influencing the use of captions by foreign language learners, An eye-tracking study. *The Modern Language Journal, 97*(1), 254–275.

## Appendix

**Table 1**

Descriptive statistics for age groups.

	Children		Adolescents		Adults	
	<i>N</i> = 19		<i>N</i> = 9		<i>N</i> = 12	
	Median	Min-Max	Median	Min-Max	Median	Min-Max
SKPL1	2.50	.0-28.33	15.00	.0-73.33	15.42	.0-86.67
SKPL2	.75	.0-16.42	13.43	.0-58.96	8.58	.0-91.79
FCL1	.91	.40-1.32	.57	.11-1.03	.42	.04-1.11
FCL2	.95	.29-1.34	.45	.09-1.09	.48	.02-1.16
AFDL1	.08	.05-.10	.09	.06-.12	.08	.05-.12
AFDL2	.08	.05-.11	.09	.05-.10	.07	.05-.10
TFDL1	59.78	6.53-98.75	35.08	7.12-89.45	25.65	2.90-78.70
TFDL2	70.60	7.54-127.30	43.14	4.84-96.30	31.67	1.63-95.30

**Table 2**

Descriptive statistics for proficiency groups.

	Beginner		Intermediate		Advanced	
	<i>N</i> = 18		<i>N</i> = 12		<i>N</i> = 10	
	Median	Min-Max	Median	Min-Max	Median	Min-Max
SKPL1	2.50	.0-28.33	16.25	.0-73.33	15.42	.0-86.67

---

SKPL2	.75	.0-35.82	8.95	.0-58.96	8.58	.0-91.79
FCL1	.92	.40-1.32	.47	.11-1.03	.53	.04-1.11
FCL2	.89	.29-1.34	.58	.09-1.09	.46	.02-1.16
AFDL1	.08	.05-.10	.08	.05-.12	.08	.05-.12
AFDL2	.08	.05-.11	.07	.05-.10	.08	.05-.10
TFDL1	62.77	6.53-98.75	31.79	7.12-89.45	34.80	2.90-78.70
TFDL2	70.19	7.54-127.30	42.64	4.84-96.30	38.93	1.63-95.30

---