Journal of Eye Movement Research
11(3):2

Line breaks in subtitling: an eye tracking study on viewer preferences

Olivia Gerber-Morón
Universitat Autònoma de Barcelona, Spain

Agnieszka Szarkowska
University College London, UK
University of Warsaw, Poland

There is a discrepancy between professional subtitling guidelines and how they are implemented in real life. One example of such discrepancy are line breaks: the way the text is divided between the two lines in a subtitle. Although we know from the guidelines how subtitles *should* look like and from watching subtitled materials how they *really* look like, little is known about what line breaks viewers would prefer. We examined individual differences in syntactic processing and viewers' preferences regarding line breaks in various linguistic units, including noun, verb and adjective phrases. We studied people's eye movements while they were reading pictures with subtitles. We also investigated whether these preferences are affected by hearing status and previous experience with subtitling. Viewers were shown 30 pairs of screenshots with syntactically segmented and non-syntactically segmented subtitles and they were asked to choose which subtitle in each pair was better. We tested 21 English, 26 Spanish and 21 Polish hearing people, and 19 hard of hearing and deaf people from the UK. Our results show that viewers prefer syntactically segmented line breaks. Eye tracking results indicate that linguistic units are processed differently depending on the linguistic category and the viewers' profile.

Keywords: Eye movements, eye tracking, reading, subtitling, line breaks, individual differences, segmentation, audiovisual translation, syntactic processing

Introduction

It is a truth universally acknowledged that subtitles should be easy to read and not stand in viewers' enjoyment of a film. One way of enhancing subtitle readability is segmentation, i.e. the way the text is divided between the two lines in a subtitle. Both subtitling scholars and professionals believe that subtitle segmentation should follow syntactic rules (Baker, Lambourne, & Rowston, 1984; BBC, 2017; Díaz Cintas & Remael, 2007; Gambier, 2006; Ivarsson & Carroll, 1998; Karamitroglou, 1998; Ofcom, 2017; Perego, 2008b). This means that linguistic units should be kept together in one line. For instance, rather than having a subtitle segmented in this way (BBC, 2017):

We are aiming to get a
better television service.

Received January 30, 2018; Published May 17, 2018.
Citation: Gerber-Morón, O. & Szarkowska, A. (2018). Line breaks in subtitling: an eye tracking study on viewer preferences. *Journal of Eye Movement Research*, *11*(3):2
Digital Object Identifier: 10.16910/jemr.11.3.2
ISSN: 1995-8692

a well-segmented subtitle would have the indefinite article '*a*' in the second line together with the rest of the noun phrase it belongs to:

We are aiming to get
a better television service.

As subtitles compete for screen space and viewers' attention with images, good subtitle segmentation is crucial to optimise readability and to enhance viewers' enjoyment of the film (Díaz Cintas & Remael, 2007). In this study, we look into viewers' preferences on subtitle segmentation and its impact on readability.

Syntactically-cued text and reading

When reading, people make sense of words by grouping them into phrases – a process known as parsing (Warren, 2012). Parsing is done incrementally, word by word: readers do not wait until the end of the sentence to interpret it, but try to make sense of it while they are reading (Frazier & Rayner, 1982; Rayner, Pollatsek, Ashby, & Clifton, 2012). To understand a sentence, readers must "first identify its syntactic relations" (Rayner et al., 2012, p. 223). If text is not syntactically cued, the reader's

comprehension may be disrupted. Syntactic ambiguities leading the reader to an incorrect interpretation, known as "garden path" sentences, need to be reanalysed and disambiguated (Frazier, 1979; Rayner et al., 2012). These ambiguities and disruptions affect eye movements, as readers make longer fixations and regress to earlier parts of the sentence to disambiguate unclear text (Frazier & Rayner, 1982).

Previous studies on reading printed text showed that syntactically-cued text facilitates reading (Levasseur, 2004; Murnane, 1987; Weiss, 1983), resulting in fewer dysfluencies at line breaks than uncued texts (Levasseur, 2004). Dividing phrases based on syntactic units has also been found to improve children's reading comprehension (Murnane, 1987; Weiss, 1983). From previous eye tracking literature, we know that some grammatical structures are more difficult to process than others, resulting in regressive eye movements and longer reading times (Ehrlich & Rayner, 1981; Rayner, Ashby, Pollatsek, & Reichle, 2004; Rayner & Well, 1996). In this study, we expect to find eye movement disfluencies (revisits, longer dwell time) in non-syntactically segmented text.

Linguistic units in subtitle segmentation

Subtitling guidelines recommend that subtitle text should be presented in sense blocks and divided based on linguistic units (Baker et al., 1984; Carroll & Ivarsson, 1998; Luyken, Herbst, Langham-Brown, Reid, & Spinhof, 1991; Perego, 2008a), at the highest syntactic nodes possible (Karamitroglou, 1998). At the phrase level, it is believed (Perego, 2008b) that the following phrases should be displayed on the same subtitle line: noun phrases (nouns preceded by an article); prepositional phrases (simple and/or complex preposition heading a noun or noun phrase); and verb phrases (auxiliaries and main verbs or phrasal verbs). At the clause and sentence level, constructions that should be kept on the same subtitle line include (Perego, 2008b): coordination constructions (sentential conjunctions such as 'and' and negative constructions with 'not'); subordination constructions (clauses introduced by the conjunction 'that'); *if*-structures and comparative constructions (clauses preceded by the conjunction 'than').

Similar rules regarding line breaks are put forward in many subtitling guidelines endorsed by television broadcasters and media regulators (ABC, 2010; BBC, 2017; DCMP, 2017; Media Access Australia, 2012; Netflix, 2016; Ofcom, 2017). According to them, the parts of speech that should not be split across a two-line subtitle are: article and noun; noun and adjective; first and last name; preposition and following phrase; conjunction and following phrase/clause; prepositional verb and preposition; pronoun and verb; and parts of a complex verb. However, when there is a conflict, synchronisation with the soundtrack should take precedence over line breaks (BBC, 2017).

Geometry in subtitle segmentation

Apart from sense blocks and syntactic phrases, another important consideration in how to form a two-line subtitle is its geometry (Baker et al., 1984; Díaz Cintas & Remael, 2007; Ivarsson & Carroll, 1998; Karamitroglou, 1998). When watching subtitled videos, viewers may not be aware of syntactic rules used to split linguistic units between the lines. What they may notice instead is subtitle shape: either closer to a pyramid or trapezoid with one line shorter than the other, or a rectangle with two lines of roughly equal length.

It is generally believed that lines within a subtitle should be proportionally equal in length because "untidy formats are disliked by viewers" (Baker et al., 1984, p. 13) and people are used to reading printed material in a rectangular format (Karamitroglou, 1998). When two lines of unequal length are used, "the upper line should preferably be shorter to keep as much of the image as free" (Carroll & Ivarsson, 1998, p. 2). If geometry is in conflict with syntax, then preference is given to the latter (Karamitroglou, 1998).

In view of the above, it is plausible that viewers make their preferences based on the shape rather than syntax (Baker et al., 1984; TED, 2015). Tests with viewers are therefore needed to understand subtitle segmentation preferences and to establish the effects of line breaks on subtitling processing.

Empirical studies on subtitle segmentation

Previous research on subtitle segmentation, including studies with eye tracking, has been limited and inconclusive. In a study on the cognitive effectiveness of subtitle processing (Perego, Del Missier, Porta, & Mosconi, 2010), no differences were found in processing subtitles with and without syntactic-based segmentation, except for longer fixations in non-syntactically segmented text. Similarly, Gerber-Morón & Szarkowska (forthcoming) did not find differences in comprehension between syntactically and non-syntactically segmented subtitles, but reported

higher cognitive load in the latter. In contrast, a study on text chunking in live subtitles (Rajendran, Duchowski, Orero, Martínez, & Romero-Fresco, 2013) showed that subtitles segmented following linguistic phrases facilitate subtitle processing. They found a significant difference in the number of eye movements between the subtitles and the image compared to non-syntactically segmented subtitles displayed word by word.

Different types of viewers

People may watch subtitled films differently depending on whether or not they are familiar with subtitling. Yet, despite an increasingly growing number of eye tracking studies on subtitling (Bisson, Van Heuven, Conklin, & Tunney, 2014; Krejtz, Szarkowska, & Krejtz, 2013; Kruger & Steyn, 2014; Kruger, Szarkowska, & Krejtz, 2015), little is known about the role of viewers' previous experience with subtitling on the way they process subtitled videos. Perego et al. (2016) conducted a cross-national study on subtitle reception and found that Italians, who are not habitual subtitle users, spent most of the watching time on reading subtitles and took more effort processing subtitles. In a study on eye movements of adults and children while reading television subtitles (d'Ydewalle & De Bruycker, 2007), longer fixations in the text were observed in children, who were less experienced in subtitling than adults. Similar fixation durations were obtained in another study on the processing of native and foreign language subtitles in native English speakers (Bisson et al., 2014), which was attributed to the lack of familiarity with subtitles.

Apart from previous experience with subtitling, another factor that impacts on the processing of subtitled videos is hearing status (de Linde, 1996). Burnham et al. (2008) note that "hearing status and literacy tend to covary" (p. 392). Early deafness has been found to be a predictor of poor reading (Albertini & Mayer, 2011; Antia, Jones, Reed, & Kreimeyer, 2009; Karchmer & Mitchell, 2003; Marschark, 1993; Marschark, Lang, & Albertini, 2002; Qi & Mitchell, 2012; Schirmer & McGough, 2005). In consequence, deaf viewers may experience difficulties when reading subtitles and their comprehension of subtitled content may be lower than that of hearing viewers (Cambra, Silvestre, & Leal, 2009; Monreal & Hernandez, 2005; Szarkowska, Krejtz, Klyszejko, & Wieczorek, 2011). One of the difficulties experienced by deaf people when reading is related to definite and indefinite articles (Channon & Sayers, 2007; Wolbers, Dostal, & Bowers, 2012). Deaf people spend more time reading function words in subtitles (such as determiners, prepositions, conjunctions or auxiliary verbs) than hard of hearing and hearing viewers (Krejtz, Szarkowska, & Łogińska, 2016). This has been attributed to the fact that many function words do not exist in sign languages, that such words tend to be short and unstressed, and therefore more difficult to identify, and that they have "low fixed semantic content outside of specific context in which they occur" (Channon & Sayers, 2007, p. 92). Given that function words are an important part of the linguistic units split between the two subtitle lines, in this study we investigate whether hearing status and previous experience with subtitling affects the preferences for or against syntactically-cued text.

Overview of the current study

This study adopts the viewers' perspective on subtitle segmentation by analysing people's preferences and reactions to different types of line breaks. To investigate these issues, the approach we developed was three-fold. First, we examined the preferences of different groups of subtitle viewers with the goal of identifying any potential differences depending on their experience with subtitling, their hearing status and the nature of the linguistic units. Second, we analysed viewers' eye movements while they were reading syntactically segmented and non-syntactically segmented subtitles. Drawing on the assumption that processing takes longer in the case of more effortful texts (Paas, Tuovinen, Tabbers, & Van Gerven, 2003), we predicted that syntactically segmented text would be preferred by viewers, whereas non-syntactically segmented text would take more time to read and result in higher mean fixation durations, particularly in the case of viewers less experienced with subtitling or deaf, given their known difficulties with processing syntactic structures (Brasel & Quigley, 1975; Brown, 1973; Conrad, 1979; Odom & Blanton, 1970; Quigley & Paul, 1984; Savage, Evans, & Savage, 1981). Finally, we invited participants to a short semi-structured interview to elicit their views on subtitle segmentation.

This study consists of two experiments: in Experiment 1 we tested hearing viewers from the UK, Poland, and Spain, while in Experiment 2 we tested British deaf, hard of hearing and hearing people. In each experiment, participants were asked to choose subtitles which they thought were better from 30 pairs of screenshots (see the Methods section). In each pair, one subtitle was segmented following the established subtitling rules, as

described in the Introduction, and the other violated them, splitting linguistic units between the two lines. After the experiment, participants were also asked whether they made their choices based on linguistic considerations or rather on subtitle shape.

Using a mixed-methods approach, where we combined preferences, eye tracking and interviews, has enabled us to gain unique insights into the reception of subtitle segmentation among different groups of viewers. To the best of our knowledge, no previous research has been conducted into viewers' preferences on subtitle segmentation, using such a wide selection of linguistic units. The results of this study are particularly relevant in the context of current subtitling practices and subtitle readability.

Methods

The study took place at University College London. Two experiments were conducted, using the same methodology and materials. The study received full ethical approval from the UCL Research Ethics Committee.

Participants

Experiment 1 involved 68 participants (21 English, 21 Polish, and 26 Spanish native speakers) ranging from 19 to 42 years of age (M=26.51, SD=6.02). Spanish speakers were included given their exposure to dubbing. Polish speakers were more accustomed to watching subtitles in comparison with Spanish speakers. English speakers were used as a control group. However, even though the participants came from different audiovisual translation traditions, most of them declared that subtitling is their preferred type of watching foreign films. They said they either use subtitles in their mother tongue or in English, which is not surprising given that the majority of the productions they watch are in English. This can be on the one hand be explained by changing viewers habits (Matamala, Perego, & Bottiroli, 2017) and on the other by the fact that our participants were living in the UK. The fact that they are frequent subtitle users also makes them a good group to ask about certain solutions used in subtitles, such as line breaks.

As the subtitles in this study were in English, we asked Polish and Spanish participants to evaluate their proficiency in reading English using the Common European Framework of Reference for Languages (from A1 to C2). All the participants declared a reading level equal or higher than B1. Of the total sample of Polish participants, 3 had a C1 level and 18 had a C2 level. In the sample of Spanish participants, 1 had a B1 level, 4 had a B2 level, 5 had a C1 and 16 had a C2 level. No statistically significant differences were found between the proficiency of Polish and Spanish participants, $\chi^2(3)=5.144$, p=.162.

Experiment 2 involved either hearing, hard of hearing, or deaf participants from the UK. We recruited 40 participants (21 hearing, 10 hard of hearing and 9 deaf) ranging from 20 to 74 years of age (M=35.59, SD=13.7). Before taking part in the experiment, hard of hearing and deaf participants completed a demographic questionnaire with information on their hearing impairment, age of hearing loss onset, communication preferences, etc. and were asked if they described themselves as either deaf or hard of hearing. Of the total sample of deaf and hard of hearing participants, 10 were profoundly deaf, 6 were severely deaf and 3 had a moderate hearing loss. In relation to the age of onset, 7 were born deaf or hard of hearing, 4 lost hearing under the age of 8, 2 lost hearing between the ages of 9-17, and 6 lost hearing between the ages of 18-40. Except for two participants who used a BSL interpreter, other hard of hearing and deaf participants chose spoken and written English to communicate during the experiment.

Participants were recruited using the UCL Psychology pool of volunteers, social media (Facebook page of the SURE project, Twitter), and personal networking. Hard of hearing and deaf participants were recruited with the help of the National Association of Deafened People and the UCL Deafness, Cognition and Language Centre participant pool. Hearing participants were paid £10 for participating in the experiment, following UCL hourly rates for experimental participants. Hard of hearing and deaf participants received £25 in recognition of the greater difficulty in recruiting special populations.

Design

In each experiment, we employed a mixed factorial design. The independent between-subject variables were language in Experiment 1 (English, Polish, Spanish) or hearing loss in Experiment 2 (hearing, hard of hearing and deaf), and the type of segmentation (syntactically segmented subtitles vs. non-syntactically segmented subtitles, henceforth referred to as SS and NSS, respectively). The main dependent variables were preferences on line breaks (SS and NSS) and eye tracking measures (dwell time, mean fixation duration and revisits).

Materials

The subtitles used in this study were in English. One reason for this choice was that it would be difficult to test line breaks and subtitle segmentation across different languages. For instance, as opposed to English and Spanish, the Polish language does not have articles, so it would be impossible to compare this linguistic unit across the languages of study participants. Another reason for using English subtitles was that it is particularly in intralingual English-to-English subtitles on television in the UK (where our study materials came from and there this study was based) that non-syntactic based segmentation is common despite the current subtitling guidelines (BBC, 2017; Ofcom, 2017).

The stimuli were 30 pairs of screenshots with subtitles in English from the BBC's *Sherlock, Series 4* (2017, dir. Mark Gatiss and Steven Moffat). Each pair contained exactly the same text, but differently segmented lines (see Figure 1).

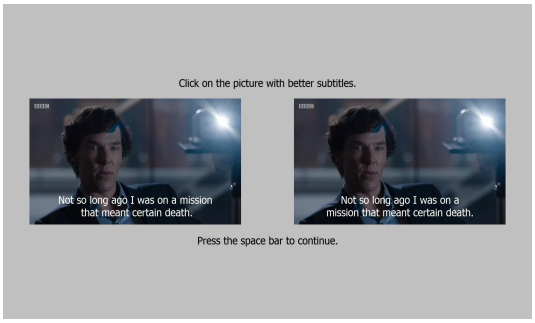

Figure 1. Stimulus example with syntactically segmented (left) and non-syntactically segmented text (right).

In one version, the two lines were segmented in accordance to subtitling standards, using syntactic rules to keep linguistic units on a single line (SS version). In the other version, syntactic rules were not followed and linguistic units were split between the first and the second line of the subtitle (NSS version). The following ten categories of the most common linguistic units (Biber, Johansson, Leech, Conrad, & Finegan, 1999) were manipulated in the study:

1. Indefinite article + noun (*IndArt*)
2. Definite article + noun (*DefArt*)
3. To + infinitive (*ToInf*)
4. Compound (*Comp*)
5. Auxiliary + lexical verb (*AuxVerb*)
6. Sentence + sentence (*SentSent*)
7. Preposition (*Prep*)
8. Possessive (*Poss*)
9. Adjective + noun (*AdjN*)
10. Conjunction (*Conj*)

For each of these categories, three instances, i.e. three different sentence stimuli, were shown (see Table 1 for examples). The presentation of screenshots (right/left) was counterbalanced, with 15 sentences in the SS condition displayed on the left, and 15 on the right. The order of presentation of the pairs (and therefore of different linguistic units) was randomised using SMI Experiment Centre.

Table 1. Examples of linguistic units manipulated in the syntactically segmented and non-syntactically segmented versions.

Category (*Abbreviation*)	Syntactic segmentation (SS)	Non-syntactic segmentation (NSS)
Indefinite article (*IndArt*)	No chance for you to be a hero this time, Mr Holmes.	No chance for you to be a hero this time, Mr Holmes.
Definite article (*DefArt*)	Because I'll know the truth when I hear it.	Because I'll know the truth when I hear it.
To + infinitive (*ToInf*)	Rest assured we have the tech to doctor a bit of security footage.	Rest assured we have the tech to doctor a bit of security footage.
Compound (*Comp*)	He's looking for the memory stick he managed to hide.	He's looking for the memory stick he managed to hide.
Auxiliary (*AuxVerb*)	Perhaps he was trying to frighten you.	Perhaps he was trying to frighten you.
Sentence + sentence (*SentSent*)	John, you amaze me. You know what happened?	John, you amaze me. You know what happened?
Preposition (*Prep*)	There were two types of vinyl in the burnt-out remains of the car.	There were two types of vinyl in the burnt-out remains of the car.
Possessive (*Poss*)	Charlie was our whole world, Mr Holmes.	Charlie was our whole world, Mr Holmes.
Adjective + noun (*AdjN*)	The memory stick is the easiest way to track you down.	The memory stick is the easiest way to track you down.
Conjunction (*Conj*)	I know you'll try to find me but there is no point.	I know you'll try to find me but there is no point.

Apparatus

SMI Red 250 mobile eye tracker was used with a two-screen set-up, one for experimenter and the other for the participant. Participants' eye movements were recorded with the sampling rate of 250Hz. The minimum duration of a fixation was set at 80 ms. We used the SMI velocity-based saccade detection algorithm. Participants with tracking ratio below 80% were excluded from eye tracking analyses. The experiment was designed and conducted using the SMI Experiment Suite. SMI BeGaze and SPSS v. 24 were used to analyse the data.

Dependent variables

The dependent variables were the preference score and three eye tracking measures (see Table 2). The preference score was calculated based on the preference expressed by a participant regarding each linguistic unit: as a percentage of people preferring SS or NSS subtitles in each linguistic unit. As there were three examples per unit, their scores were averaged per participant per unit. Participants expressed their preference by clicking on the picture with subtitles they thought were better (see Figure 2.).

Figure 2. Visualisation of mouse clicks on syntactically segmented (left) and non-syntactically segmented (right) subtitles (SentSent condition).

After completing the test with 30 pairs of subtitles, participants were asked a multiple-choice follow-up question displayed on the screen: *What was most important for you when deciding which subtitles were better*? The following options were provided: *I chose those that looked like a pyramid/trapeze (shape), I chose those that looked like a rectangle (shape), I chose those that had semantic and syntactic phrases together, I don't know*. In the post-test interview, we asked the participants if they prefer to have the first line in the subtitle shorter, longer or the same length as the second line, which prompted them to elaborate on their choices and allowed us to elicit their views on line breaks in subtitling.

Eye tracking analyses were conducted on data from areas of interest (AOIs) drawn for each subtitle in each screenshot. The three eye tracking measures used in this study are described in Table 2.

Table 2. Description of the eye tracking measures.

Eye tracking measure	**Description**
Dwell time	The sum of durations of all fixations and saccades in an AOI starting with the first fixation (reported in milliseconds). Higher dwell time may be indicative of higher cognitive effort and processing difficulties (Holmqvist et al., 2011)
Mean fixation duration (MFD)	The duration of a fixation in a subtitle AOI, averaged per clip and per participant (reported in milliseconds). Longer fixation duration is related to higher processing effort and higher difficulty of the text being read (Rayner, 1998).
Revisits	The number of glances a participant made to the subtitle AOI after visiting the subtitle for the first time (reported as a count) (Doherty & Kruger, 2018).

Procedure

Participants were tested individually in a lab. They were informed the study was on the quality of subtitles. The details of the experiment were not revealed until the end of the test during the debrief.

Before starting the test, participants read the information sheet, signed an informed consent form and underwent a 9-point calibration procedure. Participants saw 30 pairs of screenshots in randomised order. From each pair, participants had to select (i.e. click on) the screenshot with the subtitle segmentation they preferred (SS or NSS). Participants then answered the question on segmentation style preference. At the end, they undertook a short interview in which they expressed their views on subtitle segmentation based on the test and their personal experience with subtitles. The experiment concluded with the debrief of the study. The experiment lasted approximately 15 minutes, depending on the time it took the participants to answer the questions and participate in the interview.

Results

All raw data, results and experimental protocols from this experiment are openly availably in RepOD repository (Szarkowska & Gerber-Morón, 2018).

Experiment 1

Preferences

We conducted a 2 x 3 mixed ANOVA with segmentation (SS vs. NSS subtitles) as a within-subjects factor and language (English, Polish, Spanish) as a between-subjects factor with a percentage of preference for a particular linguistic unit as a dependent variable. In all linguistic parameters tested, we found a large main effect of segmentation (see Table 3). The SS subtitles were preferred over the NSS ones. Figure 3 shows preferences by linguistic units and Table 3 by participant groups. There were no differences between groups in any of the linguistic conditions and no interactions. This means that regardless of their mother tongue, all participants had similar preferences.

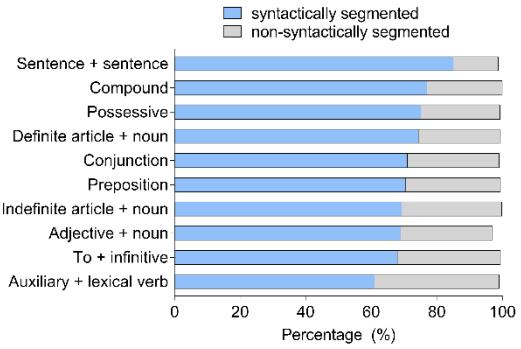

Figure 3. Preferences for SS and NSS subtitles by linguistic units in Experiment 1.

Table 3. Percentage of participants who preferred the syntactically segmented condition.

	Language						
Linguistic unit	English	Polish	Spanish	*df*	*F*	*p*	η_p^2
Indefinite article	69	76	63	1,66	28.426	.000*	.301
Definite article	74	77	71	1,66	45.264	.000*	.407
To infinitive	69	68	67	1,66	20.465	.000*	.237
Compound	82	87	69	1,66	56.267	.000*	.460
Auxiliary + verb	57	69	58	1,66	8.256	.005*	.111
Sentence + sentence	85	95	77	1,66	114.569	.000*	.634
Preposition	73	74	65	1,66	31.147	.000*	.321
Possessive	78	74	72	1,66	48.890	.000*	.426
Adjective + noun	73	64	68	1,66	21.291	.000*	.244
Conjunction	77	71	65	1,66	40.303	.000*	.379

As shown by Figure 4, the overwhelming majority of participants made their choices based on semantic and syntactic units rather than subtitle shape. Most Polish participants declared to prioritize semantic and syntactic units, whereas for English and Spanish participants pyramid shape was also considered as a choice.

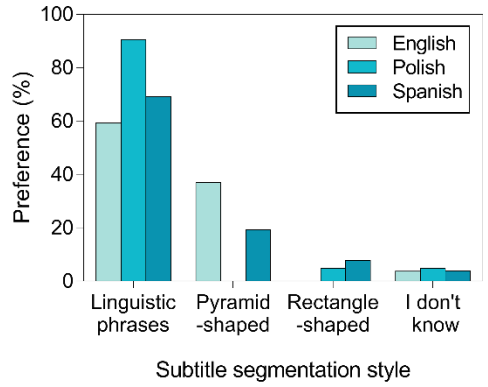

Figure 4. Segmentation preferences by group and style.

Eye tracking measures

Due to data quality issues, eye tracking analyses in Experiment 1 were conducted on 16 English, 16 Polish and 18 Spanish participants.

Dwell time

There was a main effect of segmentation on dwell time in all linguistic units apart from *ToInf*, *SentSent* and *Prep* (see Table 4). Dwell time was higher in most SS noun phrases (*IndArt*, *DefArt*, *Comp*, *Poss*) as well as in SS *Conj*, and lower in NSS *AuxVerb* and *AdjN*. There was no main effect of language on dwell time in any of the linguistic units. We found an interaction, approaching statistical significance, between segmentation and language in *Poss*, $F(2,47)=3.092$, $p=.055$, $\eta_p^2=.116$. We decomposed this interaction with simple effects with Bonferroni correction and found that for English participants there was a main effect of segmentation on dwell time in *Poss*, $F(1,15)=13.217$, $p=.002$, $\eta_p^2=.468$. Their dwell time was higher in the SS condition than in the NSS condition. There was no main effect for either Polish or Spanish participants.

Table 4. Dwell Time on subtitles by linguistic unit and segmentation (ms).

Linguistic unit split	Language			*df*	*F*	*P*	η_p^2
	English	Polish	Spanish				
Indefinite article				1,47	23.604	.000*	.334
SS	2000	1976	2185				
NSS	1536	1648	1719				
Definite article				1,47	23.913	.000*	.337
SS	1829	1821	1946				
NSS	1432	1456	1426				
To + infinitive				1,47	3.131	.083	.062
SS	1687	1603	1580				
NSS	1934	1868	1694				
Compound				1,47	5.998	.018*	.113
SS	1463	1618	1486				
NSS	1184	1473	1288				
Auxiliary + verb				1,47	9.789	.003*	.172
SS	1430	1686	1441				
NSS	1867	2132	1733				
Sentence + sentence				1,47	1.260	.267	.026
SS	1111	1167	1249				
NSS	977	1262	1010				
Preposition				1,47	1.302	.260	.027
SS	1819	1968	1866				
NSS	2079	1995	2049				
Possessive				1,47	14.284	.000*	.233
SS	1958	1649	1477				
NSS	1328	1501	1280				
Adjective + noun				1,47	12.845	.001*	.215
SS	1500	1737	1533				
NSS	1750	2365	1917				
Conjunction				1,47	7.834	.007*	.143
SS	1381	1695	1553				
NSS	1221	1377	1298				

Mean fixation duration (MFD)

There was a main effect of segmentation on MFD only in one linguistic unit: *AdjN* (Table 5), where the SS condition resulted in higher MFD than the NSS one. We also found an interaction between segmentation and language in *DefArt*, $F(2,41)=3.199$, $p=.051$, $\eta_p^2=.135$. We decomposed this interaction with simple effects with Bonferroni correction and found that for Polish participants there was a main effect of segmentation on MFD in *DefArt*, $F(1,12)=8.215$, $p=.014$, $\eta_p^2=.140$, their mean fixation duration was longer for the NSS condition. There was no main effect for English or Spanish participants.

There was a main effect of language on MFD in a number of linguistic units (see Table 6). Post-hoc Bonferroni tests showed that Polish had significantly shorter MFD than Spanish participants in *IndArt*, $p=.042$, 95% CI [-74.52, -1.06]; *DefArt*, $p=.020$, 95% CI [-60.83, -4.21]; *ToInf*, $p=.009$, 95% CI [-68.47, -7.97]; *Comp*, $p=.029$, 95% CI [-61.92, -2.62]; and *Prep*, $p=.034$, 95% CI [-1.95, -66.18]. English participants did not differ from Polish or Spanish participants.

Table 5. Mean fixation duration by linguistic unit and segmentation.

	Language						
Linguistic unit split	English	Polish	Spanish	*df*	*F*	*p*	η_p^2
Indefinite article				1,41	.429	.516	.010
SS	217	210	236				
NSS	215	192	242				
Definite article				1,41	.331	.568	.008
SS	219	180	225				
NSS	200	208	228				
To + infinitive				1,41	.221	.641	.005
SS	219	204	241				
NSS	223	195	236				
Compound				1,41	.019	.890	.000
SS	195	190	232				
NSS	202	197	219				
Auxiliary + verb				1,41	.922	.343	.022
SS	235	241	238				
NSS	218	220	242				
Sentence + sentence				1,41	2.110	.154	.049
SS	196	187	210				
NSS	172	179	202				
Preposition				1,41	.334	.566	.008
SS	211	210	233				
NSS	214	191	236				
Possessive				1,41	1.552	.220	.036
SS	216	202	225				
NSS	205	191	227				
Adjective + noun				1,41	6.103	.018*	.130
SS	220	207	230				
NSS	183	194	215				
Conjunction				1,41	.160	.691	.004
SS	213	203	225				
NSS	209	207	215				

Table 6. ANOVA results for between-subject effects in mean fixation duration in Experiment 1.

Measure	*df*	*F*	*p*	η_p^2
Indefinite article	2,41	3.416	.042*	.143
Definite article	2,41	4.154	.023*	.169
To + infinitive	2,41	4.975	.012*	.195
Compound	2,41	4.519	.017*	.181
Auxiliary + verb	2,41	.394	.677	.019
Sentence + sentence	2,41	2.561	.090	.111
Preposition	2,41	3.715	.033*	.153
Possessive	2,41	2.163	.128	.095
Adjective + noun	2,41	1.583	.218	.072
Conjunction	2,41	.548	.582	.026

Revisits

To see whether NSS subtitles induced more re-reading, which would show their lower readability, we analysed the number of revisits to the subtitles. We found a main effect of segmentation on revisits in all linguistic units apart from *SentSent*, *Prep* and *Conj* (see Table 7). Contrary to expectations, the number of revisits was higher in the SS condition for noun phrases (*IndArt*, *DefArt*, *Comp*, *Poss*). As for verb phrases (*ToInf*, *AuxVerb*) and *AdjN*, revisits were higher in the NSS condition.

We found interactions between segmentation and language in *Poss*, $F(2,53)=3.418$, $p=.040$, $\eta_p^2=.114$, and *AdjN,* $F(2,53)=7.696$, $p=.001$, $\eta_p^2=.225$. We decomposed these interactions with simple effects with Bonferroni correction and found that for English participants there was a main effect of segmentation on revisits in *Poss*, $F(1,17)=20.823$, $p=.000$, $\eta_p^2=.551$, and *AdjN*, $F(1,17)=5,017$, $p=.039$, $\eta_p^2=.228$. *Poss* was higher in the SS condition and *AdjN* was higher in the NSS condition. For Polish participants, there was no main effect of segmentation in *Poss*, but there was a main effect in *AdjN*, $F(1,15)=26.340$, $p=.000$, $\eta_p^2=.637$, being higher in the NSS condition. For Spanish participants, we found a main effect in *Poss*, $F(1,21)=5.469$, $p=.029$, $\eta_p^2=.207$, but only a tendency in *AdjN*, $F(1,21)=3.980$, $p=.059$, $\eta_p^2=.159$. They had more revisits for *Poss* in the SS condition, whereas there were more revisits for *AdjN* in the NSS condition.

There was no main effect of language on revisits in any of the linguistic units, apart from *AuxVerb,* $F(2,53)=6.437$, $p=.003$, $\eta_p^2=.195$. Post-hoc Bonferroni tests showed that Polish participants made significantly more revisits than Spanish participants, $p=.003$, 95% CI [.37, 2.10], being higher in the NSS for both groups.

Table 7. Revisits by linguistic unit and segmentation.

	Language						
Linguistic unit split	English	Polish	Spanish	*df*	*F*	*p*	η_p^2
Indefinite article				1,53	7.993	.007*	.131
SS	2.37	2.18	2.28				
NSS	1.72	2.14	1.66				
Definite article				1,53	18.767	.000*	.261
SS	2.13	2.54	1.86				
NSS	1.79	1.79	1.28				
To + infinitive				1,53	7.656	.008*	.126
SS	2.03	1.77	1.83				
NSS	2.50	2.35	1.97				
Compound				1,53	9.375	.003*	.150
SS	1.80	1.97	1.33				
NSS	1.32	1.28	1.31				
Auxiliary + verb				1,53	20.877	.000*	.283
SS	1.47	2.12	1.11				

NSS	2.58	2.96	1.50				
Sentence + sentence				1,53	.408	.526	.008
SS	.916	1.43	1.15				
NSS	1.13	1.28	.86				
Preposition				1,53	.732	.396	.014
SS	1.96	2.50	2.07				
NSS	2.18	2.45	2.25				
Possessive				1,53	24.937	.000*	.320
SS	2.46	2.02	1.74				
NSS	1.36	1.66	1.30				
Adjective + noun				1,53	36.361	.000*	.407
SS	1.61	1.90	1.77				
NSS	2.22	3.81	2.20				
Conjunction				1,53	1.924	.171	.035
SS	1.55	2.00	1.50				
NSS	1.21	1.87	1.43				

Discussion

All participants preferred SS than NSS subtitles. The strongest effect was found in the SS *SentSent* condition, with 86% participants expressing preference for the syntactically cued subtitles compared to 14% for non-syntactically cued ones. Most participants stated they prefer subtitles to be segmented according to semantic and syntactic phrase structures, and not shape.

Two interesting patterns emerged from eye tracking results on the time spent reading the noun and verb phrases in the subtitles. SS subtitles consistently induced longer dwell time for noun phrases (*IndArt, DefArt, Comp, Poss*), whereas NSS subtitles induced longer dwell time for verb phrases (*AuxVerb* and *ToInf*). We observed an interaction effect in English participants: for *Poss*, they had longer dwell time in the SS condition than Spanish and Polish participants.

Results in revisits followed the same pattern: participants made more revisits in the SS subtitles in noun phrases (*IndArt*, *DefArt*, *Comp*, *Poss*) and more revisits in NSS subtitles in verb phrases (*ToInf*, *AuxVerb*). The interactions indicated that there were more revisits for *Adj* in the SS condition across the three groups and for *Poss* in the SS condition for English and Spanish participants. These results seem to indicate that noun phrases are more difficult to process in SS condition, and verb phrases in the NSS condition.

In line with our predictions, Spanish participants, who come from dubbing tradition, showed longer mean fixation duration than English and Polish participants in both SS and NSS subtitles. There was an interaction showing that Polish had more difficulties processing *DefArt* in the NSS condition, with longer mean fixation duration.

Results

Experiment 2

Preferences

Similarly, to Experiment 1, we conducted a 2 x 3 mixed ANOVA with segmentation (SS vs. NSS subtitles) as a within-subject factor and hearing loss (hearing, hard of hearing, and deaf) as a between-subjects factor with a percentage of preference for a linguistic unit as a dependent variable.

This time we found a main effect of segmentation in all linguistic parameters apart from *AuxVerb* and *AdjN*: the SS subtitles were preferred over the NSS ones. Figure 5 presents general preferences for all linguistic units and Table 8 shows how they differed by hearing loss.

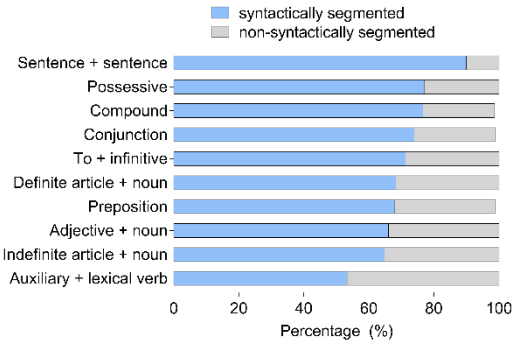

Figure 5. Preferences for SS and NSS subtitles by linguistic units in Experiment 2.

Table 8. Percentage of Experiment 2 participants who preferred the syntactically segmented condition.

Linguistic unit	Degree of hearing loss			*df*	*F*	*p*	η_p^2
	Hearing	Hard of hearing	Deaf				
Indefinite article	69	56	62	1,37	6.652	.014*	.152
Definite article	74	76	44	1,37	7.490	.009*	.168
To + infinitive	69	73	74	1,37	18.423	.000*	.332
Compound	82	73	66	1,37	22.994	.000*	.383
Auxiliary + verb	55	46	55	1,37	.255	.617	.007
Sentence + sentence	85	95	94	1,37	147.509	.000*	.799
Preposition	73	70	55	1,37	12.453	.001*	.252
Possessive	78	83	66	1,37	23.792	.000*	.391
Adjective + noun	73	65	50	1,37	2.687	.110	.068
Conjunction	77	83	55	1,37	24.441	.000*	.398

We found an almost significant interaction between segmentation and hearing loss in *DefArt*, $F(2,37)=3.086$, $p=.058$, $\eta_p^2=.143$. We decomposed it with simple effects with Bonferroni correction and found that for hearing participants there was a main effect of preference on segmentation in *DefArt*, $F(1,20)=19,375$, $p=.000$, $\eta_p^2=.492$, as well as for hard of hearing participants, $F(1,9)=7.111$, $p=.026$, $\eta_p^2=.441$, but there was no effect for deaf participants. This means that deaf participants expressed a slight preference towards NSS, but it was not significant.

There was a main effect of hearing loss in AdjN, $F(2,37)=3.469$, $p=.042$, $\eta_p^2=.158$ and a tendency approaching significance in Comp, $F(2,37)=3.063$, $p=.059$, $\eta_p^2=.142$. Post-hoc Bonferroni tests showed that hearing participants tended to express higher preference for SS *AdjN* than hard of hearing participants, $p=.051$, 95% CI [-.0009, .0834], as well as for SS *Comp*, $p=.057$, 95% CI [-.1001, .0001]. No statistically significant difference was reached in the group of deaf participants.

When asked about their choices, most hearing and hard of hearing participants declared to prioritize semantic and syntactic units, whereas for deaf participants it was the subtitle shape that was more important, as shown on Figure 6.

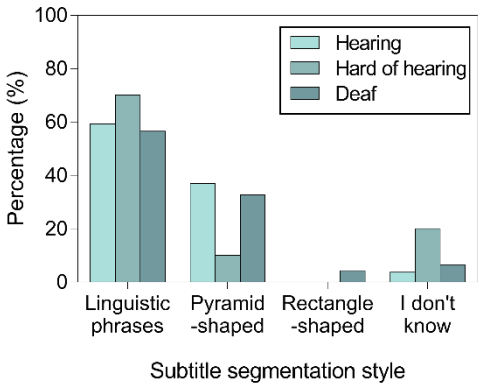

Figure 6. Segmentation preferences by group.

Eye tracking measures

Due to data quality issues, eye tracking analyses in Experiment 2 were conducted on 16 English, 8 hard of hearing and 5 deaf participants.

Dwell time

We found a significant main effect of segmentation on dwell time in *IndArt*, *AuxVerb* and *Poss* (see Table 9). Dwell time was higher for *IndArt* in the SS condition and for *AuxVerb* in the NSS condition.

We found interactions between segmentation and hearing loss in dwell time for *AdjN*, $F(2,26)=7.898$, $p=.002$, $\eta_p^2=.378$, and *Conj*, $F(2,26)=4.334$, $p=.024$, $\eta_p^2=.250$. We decomposed these interactions with simple effects with Bonferroni correction and found that for hard of hearing participants there was a main effect of segmentation on dwell time in *AdjN*, $F(1,7)=31.727$, $p=.001$, $\eta_p^2=.819$, and *Conj*, $F(1,7)=8,306$, $p=.024$, $\eta_p^2=.543$. Dwell time was higher for *AdjN* in the NSS condition and for *Conj* in the SS condition. Main effect of segmentation of

Poss for hard of hearing was higher in the SS condition. As for deaf participants, the main effect of segmentation on dwell time for *Poss* was higher in the NSS condition. There was no effect for hearing or deaf participants in *AdjN* and *Conj*.

Between-subject analysis showed a significant main effect of hearing loss in *DefArt* ($F(2,26)=3.846$, $p=.034$, $\eta_p^2= .228$) and a tendency approaching significance in *SentSent* ($F(2,26)=3.241$, $p=.055$, $\eta_p^2 =.200$). Post-hoc tests with Bonferroni correction showed that deaf participants had significantly lower dwell time than hard of hearing in *DefArt*, $p=.032$, 95% CI [-1801.76, -64.33]. Hard of hearing participants tended to have higher dwell time than hearing participants in *SentSent*, $p=.053$, 95% CI [-962.76, -4.14].

Table 9. Dwell Time by linguistic unit and segmentation (ms).

Linguistic unit split	Degree of hearing loss			*df*	*F*	*p*	η_p^2
	Hearing	Hard of hearing	Deaf				
Indefinite article				1,26	5.389	.028*	.172
SS	2000	2434	1803				
NSS	1536	2315	1442				
Definite article				1,26	2.405	.133	.085
SS	1829	2271	1053				
NSS	1432	1873	1225				
To + infinitive				1,26	.796	.381	.030
SS	1687	1908	1578				
NSS	1934	2088	1646				
Compound				1,26	1.481	.235	.054
SS	1463	1767	1502				
NSS	1184	1697	1464				
Auxiliary + verb				1,26	19.105	.000*	.424
SS	1430	1248	991				
NSS	1867	2402	1479				
Sentence + sentence				1,26	.093	.762	.004
SS	1111	1679	985				
NSS	977	1367	1331				
Preposition				1,26	3.828	.061	.128
SS	1819	2065	2238				
NSS	2079	2349	2371				
Possessive				1,26	8.795	.006*	.253
SS	1958	1806	1118				
NSS	1328	1228	1176				
Adjective + noun				1,26	2.929	.099	.101
SS	1500	2382	2328				
NSS	1750	3324	1823				
Conjunction				1,26	3.423	.076	.116
SS	1381	2246	1023				
NSS	1221	1425	1240				

Mean fixation duration (MFD)

Segmentation had no effect on MFD (Table 10) and there were no interactions between segmentation and degree of hearing loss.

There was a main effect of hearing loss on mean fixation duration in *SentSent*, $F(2,20)=3.603$, $p=.046$, $\eta_p^2=.265$.

Post-hoc Bonferroni tests showed that hard of hearing participants had significantly longer mean fixation durations than hearing participants in *SentSent*, $p=.044$, 95% CI [-59.84, -64]. Mean fixation duration for *SentSent* was higher in the SS condition for both groups.

Table 10. Mean Fixation Duration by linguistic unit and segmentation.

Linguistic unit split	Degree of hearing loss			*df*	*F*	*p*	η_p^2
	Hearing	Hard of hearing	Deaf				
Indefinite article				1,20	.370	.550	.018
SS	217	209	227				
NSS	215	224	193				
Definite article				1,20	2.977	.100	.130
SS	219	222	219				
NSS	200	207	190				
To + infinitive				1,20	.097	.758	.005
SS	219	222	212				
NSS	223	213	230				
Compound				1,20	1.118	.303	.053
SS	195	205	273				
NSS	202	207	222				
Auxiliary + verb				1,20	3.517	.075	.150
SS	235	260	267				
NSS	218	220	235				
Sentence + sentence				1,20	1.601	.220	.074
SS	196	229	186				
NSS	172	200	192				
Preposition				1,20	.295	.593	.015
SS	211	220	218				
NSS	214	202	215				
Possessive				1,20	2.496	.130	.111
SS	216	228	217				
NSS	205	219	199				
Adjective + noun				1,20	3.040	.097	.132
SS	220	222	254				
NSS	183	223	218				
Conjunction				1,20	2.927	.103	.128
SS	213	215	236				
NSS	209	216	171				

Revisits

We found a significant main effect of segmentation on revisits in *IndArt*, *AuxVerb* and *Poss*. The number of revisits was higher for *IndArt* and *Poss* in the SS condition and for *AuxVerb* in the NSS condition.

We also found interactions between segmentation and hearing loss in revisits in *ToInf*, $F(2,29)= 41.48$, $p=.026$, $\eta_p^2=.222$.

We decomposed these interactions with simple effects with Bonferroni correction and found that deaf participants tended to have more revisits for *ToInf* in the SS condition $F(1,4)=6.968$, $p=.058$, $\eta_p^2=.635$. There was no effect for English or hard of hearing participants.

Table 11. Revisits by linguistic unit and segmentation.

Linguistic unit split	Degree of hearing loss			*df*	*F*	*P*	η_p^2
	Hearing	Hard of hearing	Deaf				
Indefinite article				1,29	4.771	.037*	.141
SS	2.37	2.70	3.33				
NSS	1.72	2.48	2.60				
Definite article				1,29	.814	.374	.027
SS	2.13	2.12	1.40				
NSS	1.79	1.57	1.80				
To + infinitive				1,29	.000	.994	.000
SS	2.03	1.83	2.93				
NSS	2.50	2.55	1.73				
Compound				1,29	1.578	.219	.052
SS	1.80	1.92	2.13				
NSS	1.32	1.46	2.33				
Auxiliary + verb				1,29	19.002	.000*	.396
SS	1.47	1.22	1.60				
NSS	2.58	3.33	2.10				
Sentence + sentence				1,29	.181	.673	.006
SS	.916	1.66	1.50				
NSS	1.13	1.61	1.60				
Preposition				1,29	3.026	.093	.094
SS	1.96	2.05	2.46				
NSS	2.18	2.51	2.93				
Possessive				1,29	12.984	.001*	.309
SS	2.46	2.22	1.46				
NSS	1.36	1.33	1.20				
Adjective + noun				1,29	3.495	.072	.108
SS	1.61	2.27	3.60				
NSS	2.22	3.55	3.30				
Conjunction				1,29	.502	.484	.017
SS	1.55	1.55	1.10				
NSS	1.21	1.51	1.06				

Discussion

Similarly to Experiment 1, most participants expressed a marked preference towards SS subtitles. Again, the strongest effect was in *SentSent* cases with 90% for the SS condition compared to 10% for NSS. Deaf participants showed lower preferences than the other groups for SS subtitles in function words, such as *DefArt*, *Conj*, *Poss* and *Prep*.

Hearing and hard of hearing participants stated clearly they chose subtitles based on semantic and syntactic phrases, whereas deaf participants based their decisions on shape, with the preference towards the pyramid-shaped subtitles.

Deaf participants seemed to have more difficulties processing definite and indefinite articles than the other groups, as shown by eye tracking results: they tended to have more revisits for the SS *ToInf* compared to hearing and hard of hearing participants.

Interviews

In the post-task interviews, more than half of the participants of all the groups stated that they preferred line breaks that follow syntactic and semantic rules. However, a number of participants opted for non-syntactic line breaks, stating they give them a sense of continuity in reading, especially for some linguistic categories such as *ToInf* or *IndArt*. Many participants commented that segmentation should keep syntax and shape in balance; subtitles should be chunked according to natural thoughts, so that they can be read as quickly as possible. Other participants specified that segmentation might be an important aspect for slow readers. One interesting observation by a hard of hearing participant was that "line breaks have their value, yet when you are reading fast most of the time it becomes less relevant."

General discussion

In this study we investigated the preferences and reactions of viewers to syntactically segmented (SS) and non-syntactically segmented (NSS) text in subtitles. Our study combined an offline, metalinguistic measure of preference with online eye tracking-based reading time measures. To determine whether these measures depend on previous experience with subtitling or on hearing loss, we tested participants from countries with different audiovisual translation traditions: hearing people from the UK, Poland and Spain as well as British deaf, hard of hearing, and hearing viewers. We expected participants to prefer SS subtitles as this type of segmentation follows the "natural sentence structure" (Luyken et al., 1991, p. 47). We also hypothesized that NSS text would be more difficult to read, resulting in longer reading times. Our predictions were confirmed in relation to preferences, but only partially confirmed when it comes to eye tracking measures.

The most important finding of this study is that viewers expressed a very clear preference for syntactically segmented text in subtitles. They also declared in post-test interviews that when making their decisions, they relied more on syntactic and semantic considerations rather than on subtitle shape. These results confirm previous conjectures expressed in subtitling guidelines (Ivarsson & Carroll, 1998; Karamitroglou, 1998) and provide empirical evidence in their support.

SS text was preferred over NSS in nearly all linguistic units by all types of viewers except for the deaf in the case of the definite article. The largest preference for SS was found in the *SentSent* condition, whereas the lowest in the case of *AuxVerb*. The *SentSent* condition was the only one in our study which included punctuation. The two sentences in a subtitle were clearly separated by a full stop, thus providing participants with guidance on where one unit of meaning finished and another began. Viewers preferred punctuation marks to be placed at the end of the first line and not separating the subject from the predicate in the second sentence, thus supporting the view that each subtitle line should contain one clause or sentence (Karamitroglou, 1998). In contrast, in the *AuxVerb* condition, which tested the splitting of the auxiliary from the main verb in a two-constituent verb phrase, the viewers preferred SS text, but their preference was not as strong as in the case of the *SentSent* condition. It is plausible that in order to fully integrate the meaning of text in the subtitle, viewers needed to process not only the verb phrase itself (auxiliary + main verb), but also the verb complement.

Contrary to our predictions, some linguistic units took longer to read in the SS rather than NSS condition, as reflected by longer dwell time and more revisits. To interpret the differences between linguistic units, we classified some of them as noun or verb phrases. The *IndArt*, *DefArt*, *Comp* and *Poss* conditions were grouped under the umbrella term 'noun phrases', whereas *AuxVerb* as 'verb phrases'. In general, people spent more time reading the SS text in noun phrases, and less time reading the NSS text in the *AuxVerb*. This finding goes against the results reported by Perego et al. (2010), who tested 'ill-segmented' and 'well-segmented' noun phrases in Italian subtitles on a group of hearing people, and found no differences in the number of fixations or proportion of fixation time between the SS and NSS conditions. Interestingly, the authors also found a slightly longer mean fixation duration on NSS subtitles (228 ms in NSS compared to 216 ms in SS) – a result which was not confirmed by our data. In fact, in our study the mean fixation duration in the noun phrase *AdjN* in Experiment 1 was longer in the SS than in the NSS condition. That readers looked longer at this noun phrase category in the SS condition may be attributed to its final position at the end of the first subtitle line.

Compare, for instance:

(SS) He's looking for the memory stick
he managed to hide.

and

(NSS) He's looking for the memory
stick he managed to hide.

where in the SS condition, the complete noun phrase *Comp* is situated at the end of the first subtitle line. (Rayner, Kambe, & Duffy, 2000) found that readers looked longer at noun phrases when they were in the clause-final position. Syntactically segmented text in subtitles is characterized by the presence of complete phrases at the end of lines (Karamitroglou, 1998). According to Rayner et al. (2000), readers "fixate longer on a word when it ends a clause than when the same word does not end a clause," which could explain the longer fixation time. This result may be taken as an indication that people integrate the information from the clause at its end, including any unfinished processing before they move on, which has been referred to in literature as "clause wrap-up effect" (Just & Carpenter, 1980; Rayner et al., 2000).

This study also brought to light some important difference between how various types of viewers process line breaks in subtitling. Spanish viewers, who are generally less accustomed to subtitling and more to dubbing, had longest mean fixation duration in a number of linguistic units, indicating more effortful cognitive processing (Holmqvist et al., 2011) compared to Polish participants, who were more accustomed to subtitling. This result is not necessarily related to the nature of text segmentation, but rather to participant characteristics.

We also discovered interesting patterns of results depending on hearing loss. Deaf participants were not as concerned about syntactic segmentation as other groups, which was demonstrated by a lack of effect of segmentation on preferences in some linguistic units. This finding confirms our initial prediction about deaf people experiencing more difficulties in processing syntactic structures. The fact that there was no effect of segmentation in *DefArt* for deaf participants, combined with their longer dwell time spent on reading sentences in the *DefArt* condition, should perhaps be unsurprising, considering that deaf people with profound and severe prelingual hearing loss tend to experience difficulties with function words, including articles (Channon & Sayers, 2007; Krejtz et al., 2016; Wolbers et al., 2012). This effect can be attributed to the absence of many function words in sign languages, their context-dependence and low fixed semantic content (Channon & Sayers, 2007; Trezek, Wang, & Paul, 2010).

One important limitation of this study is that we tested static text of subtitles rather than dynamically changing subtitles displayed naturally as part of a film. The reason for this was that this approach enabled us to control linguistic units and to present participants with two clear conditions to compare. However, this self-paced reading allowed participants to take as much time as they needed to complete the task, whereas in real-life subtitling, viewers have no control over the presentation speed and have thus less time to process subtitles. The understanding of subtitled text is also context-sensitive, and as our study only contained screenshots, it did not allow participants to rely more on the context to interpret the sentences, as they would normally do when watching subtitled videos. Another limitation is the lack of sound, which could have given more context to hearing and hard of hearing participants. Yet, despite these limitations in ecological validity, we believe that this study contributes to our understanding of processing different linguistic units in subtitles.

Future research could look into subtitle segmentation in subtitled videos (see also Gerber-Morón & Szarkowska (forthcoming)), using other languages with other syntactic structures than English, which was the only language tested in this study. Further research is also required to fully understand the impact of word frequency and word length on the reading of subtitles (Moran, 2009; Rayner, 2015). Subtitle segmentation implications could also be explored across subtitles, when a sentence runs over two or more subtitles.

Our findings may have direct implications on current subtitling practices: if possible, text in the subtitles should be segmented to keep syntactic phrases together. This is particularly important in the case of two clauses or sentences separated by a punctuation mark. It is perhaps less important in the case of verb phrases like auxiliary and main verb. Following syntactic rules for segmenting subtitles can facilitate the reading process to viewers less experienced with subtitling, and can benefit deaf viewers from improving their syntax.

Ethics and Conflict of Interest

The authors declare that the contents of the article are in agreement with the ethics described in http://biblio.unibe.ch/portale/elibrary/BOP/jemr/ethics.html and that there is no conflict of interest regarding the publication of this paper.

Acknowledgements

The research reported here has been supported by a grant from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 702606, "La Caixa" Foundation (E-08-2014-1306365) and Transmedia Catalonia Research Group (SGR2005/230).

Many thanks for Pilar Orero and Judit Castellà for their comments on an earlier version of the manuscript.

References

ABC. (2010). *Industry Guidelines on Captioning Television Programs*. Retrieved from www.abc.net.au/mediawatch/transcripts/1105_freetvguidelines.pdf

Albertini, J., & Mayer, C. (2011). Using miscue analysis to assess comprehension in deaf college readers. *Journal of Deaf Studies and Deaf Education*, *16*, 35–46. https://doi.org/doi:10.1093/deafed/enq017

Antia, S. D., Jones, P. B., Reed, S., & Kreimeyer, K. H. (2009). Academic status and progress of deaf and hard-of-hearing students in general education classrooms. *Journal of Deaf Studies and Deaf Education*, *14*, 293–311. https://doi.org/doi:10.1093/deafed/enp009

Baker, R. G., Lambourne, A., & Rowston, G. (1984). *Handbook for Television Subtitlers*. Winchester: Independent Broadcasting Authority.

BBC. (2017). *BBC subtitle guidelines*. London: The British Broadcasting Corporation. Retrieved from http://bbc.github.io/subtitle-guidelines/

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Longman.

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

Brasel, K. E., & Quigley, S. (1975). *The influence of early language and communication environments on the development of language in deaf children*. Urbana, IL: University of Illinois.

Brown, R. (1973). *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.

Burnham, D., Leigh, G., Noble, W., Jones, C., Tyler, M., Grebennikov, L., & Varley, A. (2008). Parameters in television captioning for deaf and hard-of-hearing adults: effects of caption rate versus text reduction on comprehension. *J Deaf Stud Deaf Educ*, *13*(3), 391–404. https://doi.org/10.1093/deafed/enn003

Cambra, C., Silvestre, N., & Leal, A. (2009). Comprehension of Television Messages by Deaf Students at Various Stages of Education. *American Annals of the Deaf*, *153*(5), 425–434. https://doi.org/10.1353/aad.0.0065

Carroll, M., & Ivarsson, J. (1998). *Code of Good Subtitling Practice*. Berlin. Retrieved from https://www.esist.org/wp-content/uploads/2016/06/Code-of-Good-Subtitling-Practice.PDF.pdf

Channon, R., & Sayers, E. E. (2007). Toward a Description of Deaf College Students' Written English: Overuse, Avoidance, and Mastery of Function Words. *American Annals of the Deaf*, *152*(2), 91–103. https://doi.org/10.1353/aad.2007.0018

Conrad, R. (1979). *The Deaf School Child*. London: Harper and Row.

d'Ydewalle, G., & De Bruycker, W. (2007). Eye Movements of Children and Adults While Reading Television Subtitles. *European Psychologist*, *12*(3), 196–205. https://doi.org/10.1027/1016-9040.12.3.196

DCMP. (2017). *DCMP Captioning Key*. Retrieved from http://www.captioningkey.org/

de Linde, Z. (1996). Le sous-titrage intralinguistique pour les sourds et les mal entendants. In Y. Gambier (Ed.), *Les transferts linguistiques dans les medias audiovisuels* (pp. 173–183). Villeneuve d'Ascq: Presses Universitaires du Septentrion.

Díaz Cintas, J., & Remael, A. (2007). *Audiovisual translation: subtitling*. Manchester: St. Jerome.

Doherty, S., & Kruger, J.-L. (2018). The development of eye tracking in empirical research on subtitling and captioning. In J. Sita, T. Dwyer, S. Redmond, & C. Perkins (Eds.), *Seeing into Screens*. London: Bloomsbury.

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, *20*, 641–655.

Frazier, L. (1979). *On comprehending sentences: syntactic parsing strategies*. University of Connecticut, Storrs.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*(2), 178–210. https://doi.org/10.1016/0010-0285(82)90008-1

Gambier, Y. (2006). Subtitling. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics* (2nd ed., pp. 258–263). Oxford, UK: Elsevier. https://doi.org/10.1016/B0-08-044854-2/00472-7

Gerber-Morón, O., & Szarkowska, A. (forthcoming). The impact of text segmentation on subtitle reading: an eye tracking study on cognitive load and reading performance. *Journal of Eye Movement Research*.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: a comprehensive guide to methods and measures*. Oxford: Oxford University Press.

Ivarsson, J., & Carroll, M. (1998). *Subtitling*. Simrishamn: TransEdit HB.

Just, M., & Carpenter, P. (1980). A theory of the reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329–354.

Karamitroglou, F. (1998). A Proposed Set of Subtitling Standards in Europe. *Translation Journal*, *2*(2). Retrieved from http://translationjournal.net/journal/04stndrd.htm

Karchmer, M., & Mitchell, R. E. (2003). Demographic and achievement characteristics of deaf and hard-of-hearing students. In M. Marschark & P. E. Spencer (Eds.), *Oxford handbook of deaf studies, language and education* (pp. 21–37). New York, NY: Oxford University Press.

Krejtz, I., Szarkowska, A., & Krejtz, K. (2013). The Effects of Shot Changes on Eye Movements in Subtitling. *Journal of Eye Movement Research*, *6*(5), 1–12. https://doi.org/http://dx.doi.org/10.16910/jemr.6.5.3

Krejtz, I., Szarkowska, A., & Łogińska, M. (2016). Reading Function and Content Words in Subtitled Videos. *Journal Of Deaf Studies And Deaf Education*, *21*(2), 222–232. https://doi.org/10.1093/deafed/env061

Kruger, J.-L., & Steyn, F. (2014). Subtitles and Eye Tracking: Reading and Performance. *Reading Research Quarterly*, *49*(1), 105–120. https://doi.org/10.1002/rrq.59

Kruger, J.-L., Szarkowska, A., & Krejtz, I. (2015). Subtitles on the moving image: an overview of eye tracking studies. *Refractory*, *25*.

Levasseur, V. (2004). *Promoting gains in reading fluency by rereading phrasally-cued text*. ProQuest Dissertations Publishing. Retrieved from https://search.proquest.com/docview/305210949/?pq-origsite=primo

Luyken, G.-M., Herbst, T., Langham-Brown, J., Reid, H., & Spinhof, H. (1991). *Overcoming language barriers in television: dubbing and subtitling for the European audience*. Manchester: The European Institute for the Media.

Marschark, M. (1993). *Psychological development of deaf children*. New York, NY: Oxford Press.

Marschark, M., Lang, H. G., & Albertini, J. A. (2002). *Educating deaf students: Research into practice*. New York, NY: Oxford University Press.

Matamala, A., Perego, E., & Bottiroli, S. (2017). Dubbing versus subtitling yet again? *Babel*. https://doi.org/10.1075/babel.63.3.07mat

Media Access Australia. (2012). Captioning Guidelines. Retrieved from https://mediaaccess.org.au/practical-web-accessibility/media/caption-guidelines

Monreal, S. T., & Hernandez, R. S. (2005). Reading Levels of Spanish Deaf Students. *American Annals of the Deaf*, *150*(4), 379–387. https://doi.org/10.1353/aad.2005.0043

Moran, S. (2009). The Effect of Linguistic Variation on Subtitle Reception. In E. Perego (Ed.), *Eye Tracking in Audiovisual Translation* (p. 183−222). Roma: Aracne Editrice.

Murnane, Y. (1987). *The influences of chunking text and familiarity of text on comprehension*. Marquette University. Retrieved from https://search.proquest.com/docview/303587165/?pq-origsite=primo

Netflix. (2016). Timed Text Style Guide: General Requirements. Retrieved from https://backlothelp.netflix.com/hc/en-us/articles/215758617-Timed-Text-Style-Guide-General-Requirements

Odom, P. B., & Blanton, R. (1970). Implicit and explicit grammatical factors in reading achievement in the deaf. *Journal of Reading Behavior*, *2*, 47–55.

Ofcom. (2017). *Code on Television Access Services*. Retrieved from https://www.ofcom.org.uk/__data/assets/pdf_file/0020/97040/Access-service-code-Jan-2017.pdf

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive Load

Measurement as a Means to Advance Cognitive Load Theory. *Educational Psychologist*, *38*(1), 63–71. https://doi.org/10.1207/s15326985ep3801_8

Perego, E. (2008a). Subtitles and line-breaks: Towards improved readability. In D. Chiaro, C. Heiss, & C. Bucaria (Eds.), *Between Text and Image: Updating research in screen translation* (pp. 211–223). John Benjamins. https://doi.org/10.1075/btl.78.21per

Perego, E. (2008b). What would we read best? Hypotheses and suggestions for the location of line breaks in film subtitles. *The Sign Language Translator and Interpreter*, *2*(1), 35–63.

Perego, E., Del Missier, F., Porta, M., & Mosconi, M. (2010). The Cognitive Effectiveness of Subtitle Processing. *Media Psychology*, *13*(3), 243–272. https://doi.org/10.1080/15213269.2010.502873

Perego, E., Laskowska, M., Matamala, A., Remael, A., Robert, I. S., Szarkowska, A., … Bottiroli, S. (2016). Is subtitling equally effective everywhere? A first cross-national study on the reception of interlingually subtitled messages. *Across Languages and Cultures*, *17*(2), 205–229. https://doi.org/10.1556/084.2016.17.2.4

Qi, S., & Mitchell, R. E. (2012). Large-scale academic achievement testing of deaf and hard-of-hearing students: Past, present, and future. *Journal of Deaf Studies and Deaf Education*, (17), 1–18. https://doi.org/doi:10.1093/deafed/enr028

Quigley, S., & Paul, P. (1984). *Language and Deafness*. California: College-Hill Press.

Rajendran, D. J., Duchowski, A. T., Orero, P., Martínez, J., & Romero-Fresco, P. (2013). Effects of text chunking on subtitling: A quantitative and qualitative examination. *Perspectives*, *21*(1), 5–21. https://doi.org/10.1080/0907676X.2012.722651

Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin*, *124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K. (2015). Eye Movements in Reading. In J. D. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (2nd ed., pp. 631–634). United Kingdom: Elsevier. https://doi.org/10.1016/b978-0-08-097086-8.54008-2

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 720–732.

Rayner, K., Kambe, G., & Duffy, S. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology Section A*, *53*(4), 1061–1080. https://doi.org/10.1080/713755934

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C. J. (2012). *Psychology of reading* (2nd ed.). New York, London: Psychology Press.

Rayner, K., & Well, A. D. (1996). Effects of Contextual Constraint on Eye Movements in Reading: A Further Examination. *Psychonomic Bulletin and Review*, (3), 504–509.

Savage, R. D., Evans, L., & Savage, J. F. (1981). *Psychology and Communication in Deaf Children*. Sydney; London: Grune & Stratton.

Schirmer, B. R., & McGough, S. M. (2005). Teaching reading to children who are deaf: Do the conclusions of the National Reading Panel apply? *Review of Educational Research*, *75*, 83–117. https://doi.org/doi:10.3102/00346543075001083

Szarkowska, A., & Gerber-Morón, O. (2018). SURE Project Dataset. RepOD. Retrieved from http://dx.doi.org/10.18150/repod.4469278

Szarkowska, A., Krejtz, I., Klyszejko, Z., & Wieczorek, A. (2011). Verbatim, standard, or edited? reading patterns of different captioning styles among deaf, hard of hearing, and hearing viewers. *American Annals of the Deaf*, *156*(4).

TED. (2015). *How to break lines*. Open Translation Project. Retrieved from http://translations.ted.org/wiki/How_to_break_lines

Trezek, B. J., Wang, Y., & Paul, P. V. (2010). *Reading and deafness: theory, research, and practice*. Clifton Park, N.Y.: Delmar Cengage Learning.

Warren, P. (2012). *Introducing Psycholinguistics*. Cambridge: Cambridge University Press.

Weiss, D. S. (1983). The effects of text segmentation on children's reading comprehension. *Discourse Processes*, *6*(1), 77–89. https://doi.org/10.1080/01638538309544555

Wolbers, K. A., Dostal, H. M., & Bowers, L. M. (2012). 'I was born full deaf.' written language outcomes after 1 year of strategic and interactive writing instruction. *Journal of Deaf Studies and Deaf Education*, *17*(1), 19–38. https://doi.org/10.1093/deafed/enr018