Universitat Autònoma de Barcelona

SUBTITLE SEGMENTATION QUALITY ACROSS SCREENS

Olivia Gerber-Morón

UNIVERSITAT AUTÒNOMA DE BARCELONA

Departament de Traducció i d'Interpretació

i d'Estudis de l'Àsia Oriental

Doctorat en Traducció i Estudis Interculturals

SUBTITLE SEGMENTATION QUALITY ACROSS SCREENS

PhD thesis presented by:

Olivia Gerber-Morón

Supervised by:

Dr. Pilar Orero Clavero

Dr. Judit Castellà Mate

2018

With the support of "la Caixa" Foundation

To Nonna Laura

Acknowledgements

When we start our PhD journey, we generally lack the necessary foundations to conduct effective research. At best, we have read research papers on our PhD topic, participated in workshops on how to undertake research, attended conferences with experts in the field, and perhaps collaborated on projects. However, it will be the first time embarking on real in-depth research. We may have innovative and groundbreaking ideas, but we will be confronted with the unknown as we explore them. We will have to build our own path through failures and successes, try different approaches, and live many experiences. And so, we will achieve our ultimate goal, which is to develop our own approach to do good, useful research.

Along this road paved with significant challenges, I learnt valuable lessons that I will always carry with me. The most important one was the need to embrace stressful and difficult moments to overcome seemingly impossible obstacles. By conducting empirical studies and writing academic articles I learnt both the implications of doing scientific research and the focus I would like to give to mine. Collaboration between academia and the audiovisual industry is fundamental to implement scientific results and make audiovisual services more accessible to end users. I would not have learnt these lessons and countless others without the cornerstone of this investigation: the people. All this has been possible thanks to them.

Above all, this thesis has been inspired by all those committed to working for a society with better opportunities. "la Caixa" Foundation made this research possible with their financial support, and I am grateful to them for believing in my project. I am also thankful to TransMedia Catalonia, my research group, for teaching me the necessity to partner with experts from different disciplines to obtain more beneficial results. Among the people who fight to improve Media Accessibility services, my deepest appreciation goes to all the end users and deaf and hard-of-hearing associations who participated in my studies.

Four researchers in the Media Accessibility community have been my role models during my journey. Pilar Orero has always encouraged and pushed me to believe in myself, think out of the box, and dare to try new ways, even if eventually they fail. Judit Castellà gave me very honest and constructive comments about my work, which helped me to insist on carrying out more meticulous research. The long and fruitful discussions on subtitle segmentation with Elisa Perego contributed to developing my own experimental design. I also gained valuable insights from working side by side with Agnieszka Szarkowska, broadening my knowledge with her long experience. I owe a significant debt to all of them.

The encouragement of my beloved people has empowered me to achieve this goal. I exchanged ideas and lived many unique experiences with my fellow colleagues. My friends were all ears and managed to distract me when I really need it. All along the way, I found motivational support in my family. Thank you for being there for me.

Special thanks go to my friend Tara, for being enthusiastic about my research and proofreading my work.

And finally, Taku, I am very lucky to have shared my thoughts with you. Speaking with you about my thesis helped me to think more clearly, be more patient and feel more passion for my professional ambitions.

Table of contents

Index of Tables and Figures

Abbreviation and Acronym Glossary

AdjN	Adjective + noun
AENOR	Asociación Española de Normalización y Certificación
AOIs	Areas of Interest (in subtitles)
AuxVerb	Auxiliary + lexical verb
AVT	Audiovisual Translation
BBC	British Broadcasting Corporation
CI	Confidence Interval
CLT	Cognitive Load Theory
Comp	Compound
Conj	Conjunction
DCMP	Described and Captioned Media Program
DefArt	Definite article + noun
DVD	Digital Versatile Disk
EEG	Electroencephalography
H1	Hypothesis 1
H2	Hypothesis 2
H3	Hypothesis 3
H4	Hypothesis 4
HBB4ALL	Hybrid Broadcast Broadband for All
IndArt	Indefinite article + noun
MFD	Mean fixation duration
NSS	Non-syntactically segmented (text in subtitles)
Ofcom	Office of Communications (UK)
Poss	Possessive
Prep	Preposition
SD	Standard Deviation
SentSent	Sentence + sentence
SS	Syntactically segmented (text in subtitles)
SURE	Exploring Subtitle Reading Process with Eyetracking Technology
ToInf	To + infinitive

TV	Television
UAB	Universitat Autònoma de Barcelona
UCL	University College London
UK	United Kingdom
UNE	Una Norma Española
VLC	VideoLan Client
VOD	Video on demand

Chapter 1. Introduction

1. Introduction

How can we make media more accessible to everyone? I became interested in finding solutions to this question during my master's degree in Audiovisual Translation. It was then that I realised that Media Accessibility not only refers to audio description, or to subtitling for the deaf and hard of hearing; it goes much further, encompassing the services, techniques and means that make media available for people who are unable to access content in its original form (Greco, 2016; Szarkowska, Krejtz, Krejtz, & Duchowski, 2013). Later on, I attended a talk given by TransMedia Catalonia research group[1], and their approach to research drew my attention. This group conducted empirical studies in Media Accessibility with end users and provided results that could have a direct impact on society. I also wanted to contribute to improving people's quality of life, so I prepared a project on Media Accessibility and submitted it to "la Caixa" Foundation[2], a private Spanish institution that promotes social initiatives for a society with better opportunities. I was successfully awarded one of the fellowships from this foundation [E-08-2014-1306365] to carry out a PhD thesis within the Translation and Intercultural Studies PhD programme at the Department of Translation, Interpreting and East Asian Studies of the Universitat Autònoma de Barcelona. The Catalan Government [2017SGR113], the Hybrid Broadcast Broadband for All[3] project [FP7 CIP-ICT-PSP.2013.5.1], and the Exploring Subtitle Reading Process with Eyetracking Technology[4] project [H2020-MSCA-IF-2015] also partially supported the research of my PhD.

When I started my PhD thesis, I set six aims to fulfil my training as a researcher:

- To conduct empirical research and specialise in an area of Media Accessibility;
- To provide practical results that would lead to an improvement in society;

1 http://grupsderecerca.uab.cat/transmedia/

2 https://obrasociallacaixa.org/en/investigacion-y-becas/programa-de-becas-de-posgrado/doctorado-en-universidades-espanolas/descripcion-del-programa

3 https://cordis.europa.eu/project/rcn/191771_en.html

4 https://cordis.europa.eu/project/rcn/204914_it.html

- To learn how to design experiments and analyse data with statistical methods;
- To undertake academic research stays abroad to enrich my knowledge;
- To write a PhD thesis by a compendium of articles in order to understand the actual work of a researcher;
- To take part in European projects to identify the needs of the audiovisual industry.

1.1. Research focus: subtitle segmentation

As a research fellow in the field of Media Accessibility, I decided to focus my research on one of the elements of subtitling that still required further research: subtitle segmentation. Also known as "line breaks" (Gerber-Morón & Szarkowska, 2018) and "text segmentation" (Gerber-Morón, Szarkowska, & Woll, 2018), subtitle segmentation refers to the way text is divided between the two lines of a subtitle. There are normally two approaches to dividing subtitle lines. One approach considers syntactic rules to divide text, keeping units of sense in the same line (syntactic segmentation). For instance:

She told me
that <u>her aunt</u> was on holidays.

The other approach prioritises geometry and aims for more balanced lines, resulting in a rectangular shape (geometric segmentation):

She told me that <u>her</u>
<u>aunt</u> was on holidays.

1.2. Research background on subtitle segmentation

The reason behind a thorough exploration of subtitle segmentation was to determine its impact on the reading process. For over two decades, experts in Audiovisual Translation included subtitle segmentation in their handbooks, guidelines and articles, claiming — without providing empirical evidence — that syntactic segmentation facilitated the reading process of subtitles (Baker, Lambourne, & Rowston, 1984; Díaz Cintas & Remael, 2007; Ivarsson & Carroll, 1998; Karamitroglou, 2000; Perego, 2008a, 2008b). This recommendation may be based on the concept of parsing in reading (Rayner, Pollatsek, Ashby, & Clifton, 2012): readers interpret a text by identifying groups of words that go together in a sentence (Warren, 2012). According to Rayner (1998), readers experience difficulties when these groups of words are not placed together, and consequently they have to re-read the text (Frazier, 1979; Rayner et al., 2012). Based on this concept, the reading process in subtitling could be hindered when units of sense are not kept on the same line.

The appropriate placement of line breaks in subtitling could be more beneficial for specific profiles of viewers (e.g. the deaf or people more exposed to dubbing) who experience more difficulties reading subtitles (Perego, Orrego-Carmona, & Bottiroli, 2016; Szarkowska, Krejtz, Kłyszejko, & Wieczorek, 2011). Different reading studies on deaf people found that they have more difficulties processing definite and indefinite articles (Channon & Sayers, 2007; Wolbers, Dostal, & Bowers, 2012). Moreover, a subtitling study on function words (e.g. determiners, auxiliary verbs and prepositions) showed that deaf viewers spend more time reading these words than hard of hearing and hearing viewers (Krejtz, Szarkowska, & Łogińska, 2016). In the case of viewers unaccustomed to subtitling, Perego et al. (2016) found that viewers who are used to dubbing have more difficulties processing subtitled films, especially when subtitles are structurally complex. Therefore, syntactic segmentation could contribute to a better processing of subtitles.

To the best of my knowledge, the first empirical study on subtitle segmentation previous to this PhD thesis was carried out by Perego, Del Missier, Porta, & Mosconi (2010) and followed by another study by Rajendran, Duchowski, Orero, Martínez, & Romero-Fresco (2013). However, the results of these two studies did not provide conclusive evidence regarding the effects of subtitle segmentation on the reading process. Perego et al. (2010), in their eye-tracking study on line breaks with interlingual subtitles[5], did not find significant differences between processing syntactically segmented and non-syntactically segmented subtitles. In contrast, Rajendran et al. (2013) undertook an eye-tracking study on live subtitling segmentation and found that subtitles segmented by phrases were easier to process than subtitles segmented word-for-word (i.e. scrolling mode). The results of these studies probably differ because they did not use the same type of subtitles or methodology. Perego et al. (2010) tested segmentation in noun phrase structures, using a 15-minute video excerpt with interlingual subtitles. On the other hand, Rajendran et al. (2013) examined segmentation on live (respoken) subtitles, using a short video (49 seconds, with audio disabled) displayed four times with four different chunking methods.

For more detailed information regarding the research background on subtitle segmentation, please refer to the articles of the studies conducted for this PhD thesis in Chapters 2, 3 and 4. In addition, refer to Article 4 in Annex 1 which includes another article on subtitle segmentation (under review at the time of the thesis submission).

1.3. Elements studied in this PhD thesis

To understand the importance of syntactic segmentation in reading subtitles, further research on this topic had to be conducted. Subtitlers are confronted with some serious limitations when they apply syntactic rules to create subtitles. One of these is the loss of information from the original soundtrack, as this type of segmentation requires text reduction to comply with maximum line length limits (Gerber-Morón et al., 2018). Another limitation is the increase in the cost of subtitle production, as it demands more text editing, and consequently more human time and effort (Gerber-

5 Interlingual subtitling provides the translation of a foreign dialogue. In the study mentioned, subtitles were translated from Hungarian into Italian.

Morón et al., 2018). By conducting my PhD thesis on subtitle segmentation, I intended to shed more light on the relevance of syntactic rules for line breaks. To that end, I took into account the previous empirical studies on this subject (Perego et al., 2010; Rajendran et al., 2013), but I also included new elements — related to the current media situation — that had not been previously analysed.

Screen size was one of the new elements explored in this PhD thesis. Nowadays, the development of new technologies allows viewers to watch subtitled media content everywhere using innovative handheld devices, such as smartphones and tablets (Messerlin, Siwek, & Cocq, 2005; Palen, Salzman, & Youngs, 2000). Media service providers and broadcasters[6] also enable end users to activate and personalise subtitles across devices with different screen sizes. The effects of screen size have been analysed in Media Psychology and Human-Computer Interaction to assess the viewers' perception of mobility and content (Al-Showarah, AL-Jawad, & Sellahewa, 2014; Kim, Sundar, & Park, 2011; Lombard, Ditton, Grabe, & Reich, 1997; Maniar, Bennett, Hand, & Allan, 2008). The results from these studies indicate that large screens are key to great enjoyment. The effects of screen size have also been explored in subtitling by studying comprehension levels and reading patterns across devices (Castellà, Olivier, Gerber-Morón, & Soler-Vilageliu, 2016; Szarkowska, Laskowska, Oliver, & Pilipczuk, 2015), pointing out that devices with small screens seem to restrict the viewers' experience. Nevertheless, no empirical studies have been conducted on the influence of screen size on subtitle segmentation. This PhD investigated whether subtitle layout has a higher impact on comprehension scores on some of the devices where subtitles are displayed.

Another element taken into consideration was the profile of subtitle users. Previous studies on subtitle segmentation only tested undergraduate and postgraduate students (Perego et al., 2010; Rajendran et al., 2013). In their conclusions, both of these studies suggested to conduct further research with different groups of subtitle viewers. The cognitive processing of subtitles differs depending on the viewers' exposure to subtitles (Perego, Orrego-Carmona, et al.,

6 Such as Amazon Video, BBC iPlayer and Netflix.

2016). Consequently, the choice of line-break styles might affect more viewers who experience difficulty in processing subtitled films. Viewers' hearing status (deaf, hard of hearing or hearing) also determines the cognitive processing of subtitles (Szarkowska et al., 2011). Therefore, line-break styles might affect some profiles of viewers with hearing loss to a greater extent. In fact, different studies showed that deaf people have more difficulties in processing subtitles due to their generally lower reading levels (Cambra, Silvestre, & Leal, 2009; Monreal & Hernandez, 2005; Szarkowska et al., 2011). This PhD thesis included a broader range of profiles of subtitle users — native speakers of different languages, and viewers with different degrees of hearing loss — to establish whether syntactic segmentation is more beneficial for some of these profiles.

In order to complement the variables examined in the two previous studies on subtitle segmentation (Perego et al., 2010; Rajendran et al., 2013), I included other measures in the design of the experiments of this PhD thesis. The aforementioned studies comprised questionnaires on comprehension, word and scene recognition, and eye-tracking measures. Perego et al., (2010) manipulated various types of noun phrases, and Rajendran et al., (2013) tested four subtitling styles related to live (respoken) subtitles (no segmentation, word-for-word, chunked by phrase, and chunked by sentence). The studies conducted for this PhD thesis evaluated the relevance of syntactic segmentation in various linguistic units (e.g. definite article + noun, or adjective + noun) that have not been previously studied. It also compared preferences regarding two segmentation styles that had not yet been examined: syntactic and geometric segmentation. More cognitive measures were evaluated through self-reports and eye-tracking measures, and viewers' preference scores were measured with an exhaustive questionnaire. Short interviews were also conducted in some of these studies to gather more information about the participants' views on subtitle segmentation.

1.4. Research objectives

The main question I examined for my PhD thesis was: *Is subtitle segmentation a key element in Media Accessibility?* I set two objectives to provide an answer to this question.

Objective 1 studied the impact of syntactic segmentation in order to determine whether:

- the use of syntactic segmentation facilitates the reading process of subtitling;
- the implementation of syntactic segmentation in subtitling is more beneficial for specific profiles of viewers;
- some viewer profiles prefer more syntactic segmentation than others.

Objective 2 investigated the effects of screen size on subtitling, specifically on subtitle layout preferences and comprehension, so as to establish:

- the experience of watching subtitled videos across devices;
- viewers' preferences on subtitle layout (specifically focusing on line-break styles) in relation to the device;
- the device that has the most unsatisfactory results.

1.5. Hypotheses

For each of these objectives, I formulated a series of hypotheses about the impact of subtitle segmentation on the reading process.

1.5.1. Hypotheses for Objective 1

The following hypotheses for Objective 1 were formulated in relation to the impact of syntactic segmentation:

- H1: non-syntactically segmented will be more difficult to read, with a higher cognitive load and lower comprehension scores;
- H2: users will spend more time reading non-syntactically segmented subtitles, with higher mean fixation duration and more revisits to the subtitle area;
- H3: comprehension will be lower, whereas cognitive load and time spent on the subtitle will be higher for some of the viewer profiles, particularly those who are less experienced with subtitling or the deaf;
- H4: syntactically segmented subtitles will be preferred by viewers.

1.5.2. Hypotheses for Objective 2

As for the effects of screen size on subtitle layout preferences and comprehension, the following hypotheses were put forward:

- H1: viewers will evaluate subtitles differently depending on the screen size;
- H2: the smallest screen devices (i.e. smartphones) will obtain more negative results regarding subtitle layout.

1.6. Empirical studies conducted for this PhD thesis

The research objectives for this PhD thesis resulted in three empirical studies. In what follows, I briefly present the studies and the methodology used for each of them according to the research objectives. For more information, see Chapters 2, 3 and 4, which describe the methodology used for each study in detail.

Article 4 in Annex 1 is an empirical study on subtitle segmentation, but it was not included in the main manuscript as it has not yet been accepted. This study represents the ultimate verification of syntactic segmentation, considering different profiles of users and testing the effects of syntactic segmentation on smartphone screens.

1.6.1. Studies on the impact of syntactic segmentation (Objective 1)

The first two eye-tracking studies included more profiles of users, as already suggested by Perego et al. (2010) and Rajendran et al. (2013). On the one hand, cognitive load and comprehension levels were measured using eye tracking to compare the difference between syntactic and non-syntactic subtitles. On the other hand, viewers' perceptions of line breaks in subtitles were evaluated thoroughly to understand preferences and whether viewers are conscious of subtitle segmentation in subtitles.

Study 1 (Article 1, Gerber-Morón et al., 2018) explored the impact of text segmentation on subtitle processing among different groups of viewers: 74 hearing people with different mother tongues (English, Polish, and Spanish) and 46 deaf, hard of hearing, and hearing people with English as a first language. Participants watched two self-contained scenes of 1-minute duration each dubbed into Hungarian — an unknown language to all of the participants — with subtitles displayed in English. Participants watched one of the clips with syntactically segmented subtitles, and the other clip with non-syntactically segmented subtitles. The order of the clips was randomised and the segmentation conditions were counterbalanced. Segmentation (syntactically segmented vs. non-syntactically segmented subtitles) was the within-subject independent variable, and language (English, Polish, and Spanish) or hearing loss was the between-subject factor. The dependent variables included for this study were: comprehension and self-assessment cognitive load questionnaires, eye-tracking measures, and a brief interview at the end of the experiment.

Study 2 (Article 2, Gerber-Morón & Szarkowska, 2018) investigated viewers' preferences on line breaks in various linguistic units. The participants were the same as for **Study 1**, although numbers slightly differed among studies, as some participants were excluded from the eye-tracking results when their ratio was below

80%. Altogether, 68 hearing people with different mother tongues (English, Polish, and Spanish) and 40 deaf, hard of hearing and hearing people with English as a first language were tested. They were presented with 30 pairs of screenshots — in a randomised order — with syntactically segmented and non-syntactically segmented subtitles. Three eye-tracking measures were taken during the experiment (dwell time, mean fixation duration and revisits). Participants had to select the screenshot with their preferred subtitle segmentation, then answer a question on the segmentation style preference, and finally undertake a brief interview on their views regarding subtitle segmentation and their experience with subtitles. Language (English, Polish, Spanish) or hearing loss (hearing, hard of hearing and deaf), and the type of segmentation (syntactically segmented subtitles vs. non-syntactically segmented subtitles) were the independent between-subject variables. The dependent variables were preferences on line breaks (syntactically segmented subtitles vs. non-syntactically segmented subtitles) and eye-tracking measures (revisits, dwell time, and mean fixation duration).

1.6.2. Study on the effects of screen size on subtitle layout preferences and comprehension (Objective 2)

Before examining the effects of syntactic segmentation across devices, it was important to undertake a general study on the viewers' preferences on subtitle layout and their comprehension scores depending on the screen size. This study aimed to provide empirical evidence on the screen device that requires more improvement, where different line-break styles could be tested (syntactic and geometric) to determine their impact on subtitle processing.

Study 3 (Article 3, Gerber-Morón, Soler-Vilageliu, & Castellà, forthcoming) evaluated the effects of screen size on subtitle layout using three devices with different screen sizes: monitor, tablet, and smartphone. Thirty native Spanish or Catalan-Spanish speakers watched three complete scenes with a Norwegian soundtrack and Spanish subtitles. Each video fragment had an average duration of three minutes and was displayed in a randomised order on a different device according to a within-subject design. After watching each clip, participants answered questionnaires on reading and layout preferences for subtitles, as well as a comprehension questionnaire on the specific content of the scenes. The independent variable was the screen size of the devices (monitor, tablet, and smartphone), and the main dependent variables were preferences and comprehension scores.

1.7. PhD thesis structure

The following four chapters consist of the articles of the studies carried out for this PhD research. Chapter 5 includes a summary of the thesis in English, Catalan and Spanish, as required to all theses by a compendium of academic publications. Chapter 6 presents the conclusions drawn from the results of these studies, and shows the final considerations of my experience carrying out a PhD thesis. It also includes new ideas and approaches for future research in Media Accessibility. Chapter 7 contains an extensive bibliography of all the references used in this PhD thesis, and Chapter 8 comprises the annexes with all the complementary information of this research, including a forth article on the topic which was under review at the time of the thesis submission.

This PhD thesis is intended to serve not only as a contribution to improve the quality of media accessibility services, but also as food for thought on how to conduct useful and beneficial empirical research in Media Accessibility. At the end of my PhD journey, I hope to find the right resources and strategies to continue to contribute towards a better society, both in terms of technological development to aid end users and greater accessibility for the deaf and hard of hearing.

Chapter 2. Article 1

The impact of text segmentation on subtitle reading

2. Article 1

Gerber-Morón, O., Szarkowska, A. & Woll, B. (2018). The impact of text segmentation on subtitle reading. *Journal of Eye Movement Research*, *11*(4):1-18. DOI: 10.16910/11.4.2

Abstract

Understanding the way people watch subtitled films has become a central concern for subtitling researchers in recent years. Both subtitling scholars and professionals generally believe that in order to reduce cognitive load and enhance readability, line breaks in two-line subtitles should follow syntactic units. However, previous research has been inconclusive as to whether syntactic-based segmentation facilitates comprehension and reduces cognitive load. In this study, we assessed the impact of text segmentation on subtitle processing among different groups of viewers: hearing people with different mother tongues (English, Polish, and Spanish) and deaf, hard of hearing, and hearing people with English as a first language. We measured three indicators of cognitive load (difficulty, effort, and frustration) as well as comprehension and eye tracking variables. Participants watched two video excerpts with syntactically and non-syntactically segmented subtitles. The aim was to determine whether syntactic-based text segmentation as well as the viewers' linguistic background influence subtitle processing. Our findings show that non-syntactically segmented subtitles induced higher cognitive load, but they did not adversely affect comprehension. The results are discussed in the context of cognitive load, audiovisual translation, and deafness.

Keywords: eye movement, reading, region of interest, subtitling, audiovisual translation, media accessibility, cognitive load, segmentation, line breaks, revisits

2.1. Introduction

In the modern world, we are surrounded by screens, captions, and moving images more than ever before. Technological advancements and accessibility legislation, such as the United Nations Convention on the Rights of Persons with Disabilities (2006), Audiovisual Media Services Directive or the European Accessibility Act, have empowered different types of viewers across the globe in accessing multilingual audiovisual content. Viewers who do not know the language of the original production or people who are deaf or hard of hearing can follow film dialogues thanks to subtitles (Gernsbacher, 2015).

Because watching subtitled films requires viewers to follow the action, listen to the soundtrack and read the subtitles, it is important for subtitles to be presented in a way that facilitates rather than hampers reading (Díaz Cintas & Remael, 2007; Karamitroglou, 1998). Some typographical subtitle parameters, such as small font size, illegible typeface or optical blur, have been shown to impede reading (Allen, Garman, Calvert, & Murison, 2011; Thorn & Thorn, 1996). In this study, we examine whether segmentation, i.e. the way text is divided across lines in a two-line subtitle, affects the subtitle reading process. We predict that segmentation not aligned with grammatical structure may have a detrimental effect on the processing of subtitles.

2.1.1. Readability and syntactic segmentation in subtitles

The general consensus among scholars in audiovisual translation, media regulation, and television broadcasting is that to enhance readability, linguistic phrases in two-line subtitles should not be split across lines (BBC, 2017; Díaz Cintas & Remael, 2007; Ivarsson & Carroll, 1998; Karamitroglou, 1998; Ofcom, 2015). For instance, subtitle (1a) below is an example of correct syntactic-based line segmentation, whereas in (1b) the indefinite article “a” is incorrectly separated from the accompanying noun phrase (BBC, 2017).

(1a)
We are aiming to get
a better television service.

(1b)
We are aiming to get a
better television service.

The underlying assumption is that more cognitive effort is required to process text when it is not segmented according to syntactic rules (Perego, 2008a). However, segmentation rules are not always respected in the subtitling industry. One of the reasons for this might be the cost: editing text in subtitles requires human time and effort, and as such is not always cost-effective. Another reason is that syntactic-based segmentation may require substantial text reduction in order to comply with maximum line length limits. As a result, when applying syntactic rules to segmentation of subtitles, some information might be lost. Following this line of thought, BBC subtitling guidelines (BBC, 2017) stress that well-edited text and synchronisation should be prioritized over syntactically-based line breaks.

The widely held belief that words "intimately connected by logic, semantics, or grammar" should be kept in the same line whenever possible (Ivarsson & Carroll, 1998, p. 77) may be rooted in the concept of parsing in reading (Rayner et al., 2012, p. 216). Parsing, i.e. the process of identifying which groups of words go together in a sentence (Warren, 2012), allows a text to be interpreted incrementally as it is read. It has been reported that "line breaks, like punctuation, may have quite profound effects on the reader's segmentation strategies" (Kennedy, Murray, Jennings, & Reid, 1989, p. 56). Insight into these strategies can be obtained through studies of readers' eye movements, which reflect the process of parsing: longer fixation durations, higher frequency of regressions, and longer reading time may be indicative of processing difficulties (Rayner, 1998). An inappropriately placed line break may lead a reader to incorrectly interpret the meaning and structure, luring the reader into a parse that turns out to be a dead end or yield a clearly unintended reading – a so-called "garden path" experience (Frazier, 1979; Rayner et al., 2012).

The reader must then reject their initial interpretation and re-read the text. This takes extra time and, as such, is unwanted in subtitling, which is supposed to be as unobtrusive as possible and should not interfere with the viewer's enjoyment of the moving images (Díaz Cintas & Remael, 2007).

Despite a substantial body of experimental research on subtitling (Bisson, Van Heuven, Conklin, & Tunney, 2014; d'Ydewalle & De Bruycker, 2007; d'Ydewalle, Praet, Verfaillie, & Van Rensbergen, 1991; Koolstra, Van Der Voort, & D'Ydewalle, 1999; Kruger, Hefer, & Matthew, 2013; Kruger & Steyn, 2014; Perego, Laskowska, et al., 2016; Szarkowska, Krejtz, Pilipczuk, Dutka, & Kruger, 2016), the question of whether text segmentation affects subtitle processing (Perego, 2008a) still remains unanswered. Previous research is inconclusive as to whether linguistically segmented text facilitates subtitle processing and comprehension. Contrary to arguments underpinning professional subtitling recommendations, (Perego et al., 2010), who used eye-tracking to examine subtitle comprehension and processing, found no disruptive effect of "syntactically incoherent" segmentation of noun phrases on the effectiveness of subtitle processing in Italian. In their study, the number of fixations and saccadic crossovers (i.e. gaze jumps between the image and the subtitle) did not differ between the syntactically segmented and non-segmented conditions. In contrast, in a study on live subtitling, (Rajendran et al., 2013) showed benefits of linguistically-based segmentation by phrase, which induced fewer fixations and saccadic crossovers, and resulted in shortest mean fixation duration, together indicating less effortful processing.

Ivarsson & Carroll (1998) noted that "matching line breaks with sense blocks is especially important for viewers with any kind of linguistic disadvantage, e.g. immigrants or young children learning to read or the deaf with their acknowledged reading problems" (p. 78). Indeed, early deafness is strongly associated with reading difficulties (Mayberry, del Giudice, & Lieberman, 2011; Musselman, 2000). Researchers investigating subtitle reading by deaf viewers have demonstrated processing difficulties resulting in lower comprehension and more time spent by deaf viewers on reading subtitles (Krejtz, Szarkowska, & Krejtz, 2013; Krejtz et al., 2016; Szarkowska et al., 2011). Lack of familiarity with subtitling is another aspect which may affect the way people read subtitles. In a recent study, Perego et al.

(2016) found that subtitling can hinder viewers accustomed to dubbing from fully processing film images, especially in the case of structurally complex subtitles.

2.1.2. Cognitive load

Watching a subtitled video is a complex task: not only do viewers need to follow the dynamically unfolding on-screen actions, accompanied by various sounds, but they also need to read the subtitles (Kruger, Szarkowska, & Krejtz, 2015). This complex processing task may be hindered by poor quality subtitles, possibly including aspects such as non-syntactic segmentation. The processing of subtitles has been previously studied in association with the concept of cognitive load (Kruger & Doherty, 2016), rooted in cognitive load theory (CLT) and instructional design (Sweller, 2011). Drawing on the central tenet of CLT, the design of materials should aim at reducing any unnecessary load to free the processing capacity for task-related activities (Sweller, Van Merrienboer, & Paas, 1998).

In the initial formulation of CLT, two types of cognitive load were distinguished: intrinsic and extraneous (Chandler & Sweller, 1991). Intrinsic cognitive load is related to the complexity and characteristics of the task (Schmeck, Opfermann, van Gog, Paas, & Leutner, 2014). Extraneous load relates to how the information is presented; if presentation is inefficient, learning can be hindered (Sweller, Ayres, & Kalyuga, 2011). For instance, too many colours or blinking headlines in a lecture presentation can distract students rather than help them focus, wasting attentional resources on task-irrelevant details (Schmeck et al., 2014). Later studies in CLT also distinguish the concept of ‘germane cognitive load’ and, more recently, ‘germane resources’ (Schmeck et al., 2014; Sweller et al., 2011). It is believed that germane load is not imposed by the characteristics of the materials and germane resources should be “high enough to deal with the intrinsic cognitive load caused by the content” (Schmeck et al., 2014). In this paper, we set out to test whether non-syntactically segmented text may strain working memory capacity and prevent viewers from efficiently processing subtitled videos. It is our contention that just as the goal of instructional designers is to foster learning by keeping extraneous cognitive load as low as possible (Schmeck et al., 2014), so it is the task of subtitlers to reduce the extraneous load on viewers, enabling them to focus on what is important during the film-watching experience.

The concept of cognitive load encompasses different categories (Sweller et al., 1998; Wang & Duff, 2016). Mental effort is understood, following Paas, Tuovinen, Tabbers, & Van Gerven (2003, p. 64) and Sweller et al. (2011, p. 73), as "the aspect of cognitive load that refers to the cognitive capacity that is actually allocated to accommodate the demands imposed by the task". As mental effort invested in a task is not necessarily equal to the difficulty of the task, difficulty is a construct distinct from effort (van Gog & Paas, 2008). Drawing on the multidimensional NASA Task Load Index (Hart & Staveland, 1988), some researchers also included other aspects of cognitive load, such as temporal demand, performance, and frustration with the task (Sweller et al., 2011). Apart from effort, difficulty and frustration, of particular importance in the present study is performance, operationalised here as comprehension score, which demonstrates how well a person carried out the task. Performance may be positively affected by lower cognitive load, as there is more unallocated processing capacity to carry out the task. As the task complexity increases, more effort needs to be expended to keep the performance at the same level (Paas et al., 2003).

Cognitive load can be measured using subjective or objective methods (Kruger & Doherty, 2016; Sweller et al., 2011). Subjective cognitive load measurement is usually done indirectly using rating scales (Paas et al., 2003; Schmeck et al., 2014), where people are asked to rate their mental effort or the perceived difficulty of a task on a 7- or 9-point Likert scale, ranging from "very low" to "very high" (van Gog & Paas, 2008). Subjective rating scales have been criticised for using only one single item (usually either mental load or difficulty) in assessing cognitive load (Schmeck et al., 2014). Yet, they have been found to effectively show the correlations between the variation in cognitive load reported by people and the variation in the complexity of the task they were given (Paas et al., 2003). According to Sweller et al. (2011), "the simple subjective rating scale [...], has, perhaps surprisingly, been shown to be the most sensitive measure available to differentiate the cognitive load imposed by different instructional procedures" (p. 74). The problem with rating scales is they are applied to the task as a whole, after it has been completed. In contrast, objective methods, which include physiological tools such as eye tracking or electroencephalography (EEG), enable researchers to see fluctuations in cognitive load over time (Antonenko, Paas, Grabner, & van Gog, 2010; Van Gerven, Paas, Van Merriënboer, & Schmidt, 2004). Higher number of fixations and longer fixation durations are generally associated with higher processing effort and increased cognitive load (Holmqvist et al., 2011; Kruger, Doherty, Fox, & de Lissa, 2017). In our study, we combine subjective rating scales with objective eye-tracking measures to obtain a more reliable view on cognitive load during the task of subtitle processing.

Various types of measures have been used to evaluate cognitive load in subtitling. Some previous studies have used subjective post-hoc rating scales to assess people's cognitive load when watching subtitled audiovisual material (Kruger & Doherty, 2016; Kruger, Hefer, & Matthew, 2014; Yoon & Kim, 2011); subtitlers' cognitive load when producing live subtitles with respeaking (Szarkowska, Krejtz, Dutka, & Pilipczuk, 2016); or the level of translation difficulty (Sun & Shreve, 2014). Some studies on subtitling have used eye tracking to examine cognitive load and attention distribution in a subtitled lecture (Kruger et al., 2014); cognitive load while reading edited and verbatim subtitles (Szarkowska et al., 2011); or the processing of native and foreign subtitles in films (Bisson et al., 2014); to mention just a few. Using both eye tracking and subjective self-report ratings, Łuczak (2017) tested the impact of the language of the soundtrack (English, Hungarian, or no audio) on viewers' cognitive load. Kruger, Doherty, Fox, et al. (2017) combined eye tracking, EEG and self-reported psychometrics in their examination of the effects of language and subtitle placement on cognitive load in traditional intralingual subtitling and experimental integrated titles. For a critical overview of eye tracking measures used in empirical research on subtitling, see (Doherty & Kruger, 2018a), and of the applications of cognitive load theory to subtitling research, see Kruger & Doherty (2016).

2.1.3. Overview of the current study

The main goal of this study is to test the impact of segmentation on subtitle processing. With this goal in mind, we showed participants two videos: one with syntactically segmented text in the subtitles (SS) and one where text was not syntactically segmented (NSS). In order to compensate for any differences in the knowledge of source language and accessibility of the soundtrack to deaf and hearing participants, we used videos where the soundtrack was in Hungarian – a language that participants could not understand.

All subtitles in this study were shown in English. The reason for this is threefold. First, the non-compliance with the subtitling guidelines with regard to text segmentation and line breaks is particularly visible on British television in English-to-English subtitling. Although the UK is the leader in subtitling when it comes to the quantity of subtitle provision, with many TV channels having 100% subtitling to its programmes, the quality of pre-recorded subtitles is often below professional subtitling standards with regard to subtitle segmentation. Another reason for using English – as opposed to showing participants subtitles in their respective mother tongues – was to ensure identical linguistic structures in the subtitles. A final reason for using English is that, as participants live in the UK, they are able to watch English subtitles on television. The choice of English subtitles is therefore ecologically valid.

We measured participants' cognitive load and comprehension as well as a number of eye tracking variables. Following the established method of measuring self-reported cognitive load previously used by Kruger et al. (2014), (Szarkowska, Krejtz, Pilipczuk, et al., 2016), and Łuczak (2017), we measured three aspects of cognitive load: perceived difficulty, effort, and frustration, using subjective 1-7 rating scales (Schmeck et al., 2014). We also related viewers' cognitive load to their performance, operationalised here as comprehension score. Based on the subtitling literature (Perego, 2008b), we predicted that non-syntactically segmented text in subtitles would result in higher cognitive load and lower comprehension. We hypothesised that subtitles in the NSS condition would be more difficult to read because of increased parsing difficulties and extra cognitive resources which might be expended on additional processing.

In terms of eye tracking, we hypothesised that people would spend more time reading subtitles in the NSS condition. To measure this, we calculated the absolute reading time and proportional reading time of subtitles as well as fixation count in the subtitles. Absolute reading time is the time the viewers spent in the subtitle area, measured in milliseconds, whereas proportional reading time is a percentage of time spent in the subtitle area relative to subtitle duration (D'Ydewalle, Rensbergen, & Pollet, 1987; Koolstra et al., 1999). Furthermore, because we thought that the non-syntactically segmented text would be more difficult to process, we also expected higher mean fixation duration and more revisits to the subtitle area in the NSS condition (Holmqvist et al., 2011; Rayner, 2015; Rayner et al., 2012).

To address the contribution of hearing status and experience with subtitling to cognitive processing, our study includes British viewers with varying hearing status (deaf, hard of hearing, and hearing), and hearing native speakers of different languages: Spanish people, who grew up in a country where the dominant type of audiovisual translation is dubbing, and Polish people, who come from the tradition of voice-over and subtitling. We conducted two experiments: Experiment 1 with hearing people from the UK, Poland, and Spain, and Experiment 2 with English hearing, hard of hearing and deaf people. We predicted that for those who are not used to subtitling, cognitive load would be higher, comprehension would be lower and time spent in the subtitle would be higher, as indicated by absolute reading time, fixation count and proportional reading time.

By using a combination of different research methods, such as eye tracking, self-reports, and questionnaires, we have been able to analyse the impact of text segmentation on the processing of subtitles, modulated by different linguistic backgrounds of viewers. Examining these issues is particularly relevant from the point of view of current subtitling standards and practices.

2.2. Methods

The study took place at University College London and was part of a larger project on testing subtitle processing with eye tracking. In this paper, we report the results from two experiments using the same methodology and materials: Experiment 1 with hearing native speakers of English, Polish, and Spanish; and Experiment 2 with hearing, hard of hearing, and deaf British participants. The English-speaking hearing participants are the same in both experiments. In each of the two experiments, we employed a mixed factorial design with segmentation (syntactically segmented vs. non-syntactically segmented) as the main within-subject independent variable, and language (Exp. 1) or hearing loss (Exp. 2) as a between-subject factor.

All the study materials and results are available in an open data repository RepOD hosted by the University of Warsaw (Szarkowska & Gerber-Morón, 2018).

2.2.1. Participants

Participants were recruited from the UCL Psychology pool of volunteers, social media (Facebook page of the project, Twitter), and personal networking. Hard of hearing participants were recruited with the help of the National Association of Deafened People. Deaf participants were also contacted through the UCL Deafness, Cognition, and Language Research Centre participant pool. Participants were required not to know Hungarian.

Table 2.1 Demographic information on participants

Experiment 1				
		English	Polish	Spanish
Gender				
	Male	13	5	10
	Female	14	16	16
Age				
	Mean (SD)	27.59 (7.79)	24.71 (5.68)	28.12 (5.88)
	Range	20-54	19-38	19-42

Experiment 2				
		Hearing	Hard of hearing	Deaf
Gender				
	Male	13	2	4
	Female	14	8	5
Age				
	Mean (SD)	27.59 (7.79)	46.40 (12.9)	42.33 (14.18)
	Range	20-54	22-72	24-74

Experiment 1 participants were pre-screened to be native speakers of English, Polish or Spanish, aged above 18. They were all resident in the UK. We tested 27 English, 21 Polish, and 26 Spanish speakers (see Table 2.1). At the study planning and design stage, Spanish speakers were included on the assumption that they would be unaccustomed to subtitling as they come from Spain, a country in which foreign programming is traditionally presented with dubbing. Polish participants were included as Poland is a country where voice-over and subtitling are commonly used, the former on television and VOD, and the latter in cinemas, DVDs, and VOD. The hearing English participants were used as a control group.

Despite their experiences in their native countries, when asked about the preferred type of audiovisual translation (AVT), most of the Spanish participants declared they preferred subtitling and many of the Polish participants reported that they watch films in the original (see Table 2.2).

Table 2.2 Preferred way of watching foreign films

	English	Polish	Spanish
Subtitling	24	11	22
Dubbing	0	0	1
Voice-over	1	0	0
I watch films in their original version	1	10	3
I never watch foreign films	1	0	0

We also asked the participants how often they watched English and non-English programmes with English subtitles (Figure 2.1).

The heterogeneity of participants' habits and preferences reflects the changing AVT landscape in Europe (Matamala, Perego, & Bottiroli, 2017) on the one hand, and on the other, may be attributed to the fact that participants were living in the UK and thus had different experiences of audiovisual translation than in their home countries. The participants' profiles make them not fully representative of the Spanish/Polish population, which we acknowledge here as a limitation of the study.

To determine the level of participants' education, hearing people were asked to state the highest level of education they completed (Table 2.3, see also Table 2.5 for hard of hearing and deaf participants). Overall, the sample was relatively well-educated.

Figure 2.1 Participants' subtitle viewing habits

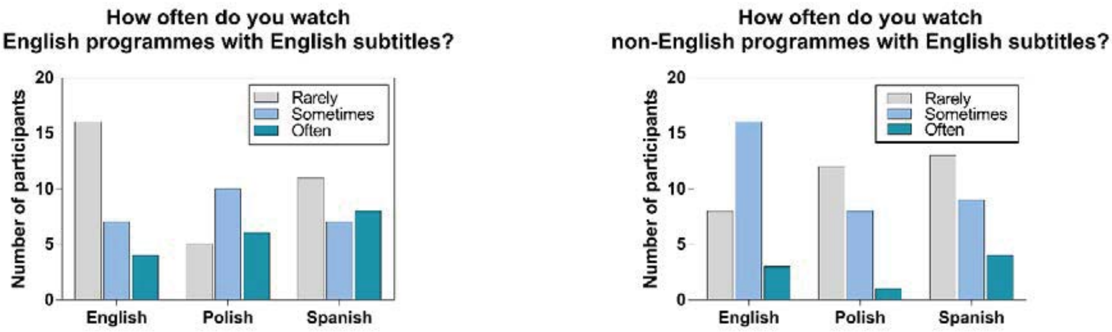


Table 2.3 Education background of hearing participants in Experiment 1

	English	Polish	Spanish
Secondary education	5	9	6
Bachelor degree	14	4	6
Master degree	8	8	13
PhD	0	0	1

As subtitles used in the experiments were in English, we asked Polish and Spanish speakers to assess their proficiency in reading English using the Common European Framework of Reference for Languages (from A1 to C2), see Table 2.4. None of the participants declared a reading level lower than B1. The difference between the proficiency in English of Polish and Spanish participants was not statistically significant, $\chi^2(3) = 5.144$, $p = .162$. Before declaring their proficiency, each participant was presented with a sheet describing the skills and competences required at each proficiency level (Szarkowska & Gerber-Morón, 2018). There is evidence that self-report correlates reasonably well with objective assessments (Marian, Blumenfeld, & Kaushanskaya, 2007).

Table 2.4 Self-reported English proficiency in reading of Polish and Spanish participants

	Polish	Spanish
B1	0	1
B2	0	4
C1	3	5
C2	18	16
Total	21	26

In Experiment 2, participants were classified as either hearing, hard of hearing, or deaf. Before taking part in the study, those with hearing impairment completed a questionnaire about the severity of their hearing impairment, age of onset of hearing impairment, communication preferences, etc. and were asked if they described themselves as deaf or hard of hearing. They were also asked to indicate their education background (see Table 2.5). We recruited 27 hearing, 10 hard of hearing, and 9 deaf participants. Of the deaf and hard of hearing participants, 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. Nine were profoundly deaf, 6 were severely deaf, and 4 had a moderate hearing loss. Seventeen of the deaf and hard of hearing participants preferred to use spoken English as their means of communication in the study and two chose to use a British Sign Language interpreter. In relation to AVT, 84.2% stated that they often watch films in English with English subtitles; 78.9% declared they could not follow a film without subtitles; 58% stated that they always or very often watch non-English films with English subtitles. Overall, deaf and hard of hearing participants in our study were experienced subtitle users, who rely on subtitles to follow audiovisual materials.

Table 2.5 Education background of deaf and hard of hearing participants

	Deaf	Hard of hearing
GCSE/O-levels	3	1
A-levels	2	4
University level	4	5

In line with UCL hourly rates for experimental participants, hearing participants received £10 for their participation in the experiment. In recognition of the greater difficulty in recruiting special populations, hard of hearing and deaf participants were paid £25. Travel expenses were reimbursed as required.

2.2.2. Materials

These comprised two self-contained 1-minute scenes from films featuring two people engaged in a conversation: one from *Philomena* (Desplat & Frears, 2013) and one from *Chef* (Bespalov & Favreau, 2014). The clips were dubbed into Hungarian – a language unknown to any of the participants and linguistically unrelated to their native languages. Subtitles were displayed in English, while the audio of the films was in Hungarian. Table 2.6 shows the number of linguistic units manipulated for each clip.

Table 2.6 Number of instances manipulated for each type of linguistic unit

Linguistic unit	Chef	Philomena
Auxiliary and lexical verb	2	2
Subject and predicate	3	3
Article and noun	3	3
Conjunction between two clauses	4	5

Subtitles were prepared in two versions: syntactically segmented and non-syntactically segmented (see Table 2.7) (SS and NSS, respectively). The SS condition was prepared in accordance with professional subtitling standards, with linguistic phrases appearing on a single line. In the NSS version, syntactic phrases were split between the first and the second line of the subtitle. Both the SS and the NSS versions had identical time codes and contained exactly the same text. The clip from Philomena contained 16 subtitles, of which 13 were manipulated for the purposes of the experiment; Chef contained 22 subtitles, of which 12 were manipulated. Four types of linguistic units were manipulated in the NSS version of both clips (see Tables 2.6 and 2.7).

Each participant watched two clips: one from *Philomena* and one from *Chef*; one in the SS and one in the NSS condition. The conditions were counterbalanced and their order of presentation was randomised using SMI Experiment Centre (see Szarkowska & Gerber-Morón, 2018).

Table 2.7 Examples of line breaks in the SS and the NSS condition

Linguistic unit	SS condition	NSS condition
Auxiliary and lexical verb	Now, should we <u>have served</u> that sandwich?	Now, should we <u>have</u> <u>served</u> that sandwich?
Subject and predicate	That's my son. Get back in there. <u>We got</u> some hungry people.	That's my son. Get back in there. <u>We</u> <u>got</u> some hungry people.
Article and noun	I've loved <u>the hotels</u>, the food and everything,	I've loved <u>the</u> <u>hotels</u>, the food and everything,
Conjunction between two clauses	Now I've made a decision <u>and</u> my mind's made up.	Now I've made a decision <u>and</u> my mind's made up.

2.2.3. Eye tracking recording

An SMI RED 250 mobile eye tracker was used in the experiment. Participants' eye movements were recorded with a sampling rate of 250Hz. The experiment was designed and conducted with the SMI software package Experiment Suite, using the velocity-based saccade detection algorithm. The minimum duration of a fixation was 80ms. The analyses used SMI BeGaze and SPSS v. 24. Eighteen participants whose tracking ratio was below 80% were excluded from the eye tracking analyses (but not from comprehension or cognitive load assessments).

2.2.4. Dependent variables

The dependent variables were: 3 indicators of cognitive load (difficulty, effort and frustration), comprehension score, and 5 eye tracking measures.

The following three indicators of cognitive load were measured using self-reports on a 1-7 scale: difficulty ("Was it difficult for you to read the subtitles in this clip?", ranging from "very easy" to "very difficult"), effort ("Did you have to put a lot of effort into reading the subtitles in this clip?", ranging from "very little effort" to "a lot of effort"), and frustration ("Did you feel annoyed when reading the subtitles in this clip?", ranging from "not annoyed at all" to "very annoyed").

Comprehension was measured as the number of correct answers to a set of five questions per clip about the content, focussing on the information from the dialogue (not the visual elements). See Szarkowska & Gerber-Morón (2018) for the details, including the exact formulations of the questions.

Table 2.8 contains a description of the eye tracking measures. We drew individual areas of interest (AOIs) on each subtitle in each clip. All eye tracking data reported here comes from AOIs on subtitles.

Table 2.8 Description of the eye tracking measures

Eye tracking measure	Description
Absolute reading time	The sum of all fixation durations and saccade durations, starting from the duration of the saccade entering the AOI, referred to in SMI software as 'glance duration'. Longer time spent on reading may be indicative of difficulties with extracting information (Holmqvist et al., 2011).
Proportional reading time	The percentage of dwell time (the sum of durations of all fixations and saccades in an AOI starting with the first fixation) a participant spent in the AOI as a function of subtitle display time. For example, if a subtitle lasted for 3 seconds and the participant spent 2.5 seconds in that subtitle, the proportional reading time was 2500/3000 ms = 83% (i.e. while the subtitle was displayed for 3 seconds, the participant was looking at that subtitle for 83% of the time). Longer proportional time spent in the AOI translates into less time available to follow on-screen action.
Mean fixation duration	The duration of a fixation in a subtitle AOI, averaged per clip per participant. Longer mean fixation duration may indicate more effortful cognitive processing (Holmqvist et al., 2011).
Fixation count	The number of fixations in the AOI, averaged per clip per participant. Higher numbers of fixations have been reported in poor readers (Holmqvist et al., 2011).
Revisits	The number of glances a participant made to the subtitle AOI after visiting the subtitle for the first time. Revisits to the AOI may indicate problems with processing, as people go back to the AOI to re-read the text.

2.2.5. Procedure

The study received full ethical approval from the UCL Research Ethics Committee. Participants were tested individually. They were informed they would take part in an eye tracking study on the quality of subtitles. The details of the experiment were not revealed until the debrief.

After reading the information sheet and signing the informed consent form, each participant underwent a 9-point calibration procedure. There was a training session, whose results were not recorded. Its aim was to familiarise the participants with the experimental procedure and the type of questions that would be asked in the experiment (comprehension and cognitive load). Participants watched the clips with the sound on. After the test, participants' views on subtitle segmentation were elicited in a brief interview.

Each experiment lasted approx. 90 minutes (including other tests not reported in this paper), depending on the time it took the participants to answer the questions and participate in the interview. Deaf participants had the option of either communicating via a British Sign Language interpreter or by using their preferred combination of spoken language, writing and lip-reading.

2.3. Results

2.3.1. Results from Experiment 1

Seventy-four participants took part in this experiment: 27 English, 21 Polish, 26 Spanish.

Cognitive load

To examine whether subtitle segmentation affects viewers' cognitive load, we conducted a 2 x 3 mixed ANOVA on three indicators of cognitive load: difficulty, effort, and frustration, with segmentation as a within-subject independent variable (SS vs. NSS) and language (English, Polish, Spanish) as a between-subject factor. We found a main effect of segmentation on all three aspects of cognitive load, which were consistently higher in the NSS condition compared to the SS one (Table 2.9).

Table 2.9 Mean cognitive load indicators for different participant groups in Experiment 1

	Language						
	English	Polish	Spanish	df	F	P	η_p^2
Difficulty				1,71	15,584	< .001*	.18
SS	2.37	2.05	1.96				
	(1.27)	(1.02)	(1.14)				
NSS	2.63	2.67	3.42				
	(1.44)	(1.46)	(1.65)				
Effort				1,71	7,788	.007*	.099
SS	2.78	1.90	2.23				
	(1.55)	(1.26)	(1.50)				
NSS	2.89	2.43	3.54				
	(1.60)	(1.16)	(2.10)				
Frustration				1,71	27,030	< .001*	.276
SS	2.15	1.38	1.62				
	(1.40)	(.80)	(.89)				
NSS	3.04	2.48	3.27				
	(1.85)	(1.91)	(2.07)				

We also found an interaction between segmentation and language in the case of difficulty, $F(2,71) = 3{,}494$, $p = .036$, $\eta_p^2 = .090$, which we separated with simple effects analyses (post-hoc tests with Bonferroni correction). We found a significant main effect of segmentation on the difficulty of reading subtitles among Spanish participants, $F(1,25) = 19{,}161$, $p < .001$, $\eta_p^2 = .434$. Segmentation did not have a statistically significant effect on the difficulty experienced by English participants, $F(1,26) = {,}855$, $p = .364$, $\eta_p^2 = .032$ or by Polish participants, $F(1,20) = 2{,}147$, $p = .158$, $\eta_p^2 = .097$. To recap, although cognitive load difficulty was declared to be higher by all participants in the NSS condition, only in the case of Spanish participants was the main effect of segmentation statistically significant.

We did not find any significant main effect of language on cognitive load (Table 2.10), which means that participants reported similar scores regardless of their linguistic background.

Table 2.10 Between-subjects results for cognitive load

Measure	df	F	p	η_p^2
Difficulty	2,71	.592	.556	.016
Effort	2,71	2.382	.100	.063
Frustration	2,71	1.850	.165	.050

Comprehension

To see whether segmentation affects viewers' performance, we conducted a 2 x 3 mixed ANOVA on segmentation (SS vs. NSS condition) with language (English, Polish, Spanish) as a between-subject factor. The dependent variable was comprehension score. There was no main effect of segmentation on comprehension $F(1,71) = .412$, $p = .523$, $\eta_p^2 = .006$. Table 2.11 shows descriptive statistics for this analysis. There were no significant interactions.

Table 2.11 Descriptive statistics for comprehension

	Language	Mean (SD)
Comprehension SS	English	4.11 (1.01)
	Polish	4.48 (.81)
	Spanish	4.08 (1.09)
	Total	4.20 (.99)
Comprehension NSS	English	4.26 (1.02)
	Polish	4.76 (.43)
	Spanish	3.88 (1.21)
	Total	4.27 (1.02)

We found a main effect of language on comprehension, $F(2,71) = 3,563$, $p = .034$, $\eta_p^2 = .091$. Pairwise comparisons with Bonferroni correction showed that Polish participants had significantly higher comprehension than Spanish participants, $p = .031$, 95% CI [.05, 1.23]. There was no difference between Polish and English, $p =.224$, 95% CI [-.15, 1.02], or Spanish and English participants, $p =1.00$, 95% CI [-.76, .35].

Eye tracking measures

Because of data quality issues, for eye tracking analyses we had to exclude 8 participants from the original sample, leaving 22 English, 19 Polish, and 25 Spanish participants. We found a significant main effect of segmentation on revisits to the subtitle area (Table 2.12). Participants went back to the subtitles more in the NSS condition (M_{NSS} = .37, *SD* = .25) compared to the SS one (M_{SS} = .25, *SD* = .22), implying potential parsing problems. There was no effect of segmentation for any other eye tracking measure (Table 2.12). There were no interactions.

Table 2.12 Mean eye tracking measures by segmentation in Experiment 1

	Language						
	English	Polish	Spanish	df	*F*	p	η_p^2
Absolute reading time (ms)				1,63	2.950	.091	.045
SS	1614	1634	1856				
NSS	1617	1529	1817				
Proportional reading time				1,63	2.128	.150	.033
SS	.65	.67	.76				
NSS	.66	.62	.74				
Mean fixation duration (ms)				1,63	2.128	.906	.000
SS	209	194	214				
NSS	211	187	218				
Fixation count				1,63	2.279	.136	.035
SS	6.41	6.68	7.27				
NSS	6.45	6.42	6.95				
Revisits				1,63	11.839	.001*	.158
SS	.28	.27	.21				
NSS	.39	.34	.36				

In relation to the between-subject factor, we found a main effect of language on absolute reading time, proportional reading time, mean fixation duration, and fixation count, but not on revisits (see Table 2.13).

Post-hoc Bonferroni analyses showed that Spanish participants spent significantly more time in the subtitle area compared to English and Polish participants. This was shown by significantly longer absolute reading time in the case of Spanish participants compared to English, $p = .027$, 95% CI [19.20, 422.73], and Polish participants, $p = .012$, 95% CI [44.61, 464.75]. Polish and English participants did not differ from each other in absolute reading time, $p = 1.00$, 95% CI [-249.88, 182.45]. There was a tendency approaching significance for fixation count to be higher among Spanish participants than English participants, $p = .077$, 95% CI [-.05, 1.41]. Spanish participants also had higher proportional reading time when compared to English participants, $p = .029$, 95% CI [.007, .189] and Polish participants, $p = .015$, 95% CI [.01, .20], i.e. the Spanish participants spent most time reading the subtitle while viewing the clip. Finally, Polish participants had a statistically lower mean fixation duration compared to English, $p = .041$, 95% CI [-38.10, -59], and Spanish, $p = .003$, 95% CI [-43.62, -7.16]. English and Spanish participants did not differ from each other in mean fixation duration, $p = 1.00$, 95% CI [-23.55, 11.47].

Table 2.13 ANOVA results for between-subject effects in Experiment 1

Measure	df	F	p	η_p^2
Absolute reading time	2,63	5.593	.006*	.151
Proportional reading time	2,63	5.398	.007*	.146
Mean fixation duration	2,63	6.166	.004*	.164
Fixation count	2,63	2.980	.058	.086
Revisits	2,63	.332	.719	.010

Overall, the results indicate that the processing of subtitles was least effortful for Polish participants and most effortful for Spanish participants.

2.3.2. Results from Experiment 2

A total of 46 participants (19 males, 27 females) took part in the experiment: 27 were hearing, 10 hard of hearing, and 9 deaf.

Cognitive load

We conducted 2 x 3 mixed ANOVAs on each indicator of cognitive load with segmentation (SS vs. NSS) as a within-subject variable and degree of hearing loss (hearing, hard of hearing, deaf) as a between-subject variable.

Similarly to Experiment 1, we found a significant main effect of segmentation on difficulty, effort, and frustration (Table 2.14). The NSS subtitles induced higher cognitive load than the SS condition in all groups of participants. There were no interactions.

Table 2.14 Mean cognitive load indicators for different participant groups in Experiment 2

	Degree of hearing loss						
	Hearing M (SD)	Hard of hearing M (SD)	Deaf M (SD)	df	*F*	*p*	η_p^2
Difficulty				1,43	6,580	.014*	.133
SS	2.37 (1.27)	1.60 (1.07)	2.56 (1.42)				
NSS	2.63 (1.44)	2.20 (1.31)	3.44 (1.59)				
Effort				1,43	4,372	.042*	.092
SS	2.78 (1.55)	1.60 (1.07)	2.78 (1.64)				
NSS	2.89 (1.60)	2.50 (1.35)	3.44 (1.42)				
Frustration				1,43	7,669	.008*	.151
SS	2.15 (1.40)	1.00 (.00)	2.56 (1.59)				
NSS	3.04 (1.85)	2.10 (1.28)	3.00 (1.58)				

There was no main effect of hearing loss on difficulty, *F*(2,43) = 2.100, *p* = .135, η_p^2 = .089 or on effort, *F*(2,43) = 1.932, *p* = .157, η_p^2 = .082, but there was an effect near to significance on frustration, *F*(2,43) = 3.100, *p* = .052, η_p^2 = .129. Post-hoc tests showed a result approaching significance: hard of hearing participants reported lower frustration levels than hearing participants, *p* = .079, 95% CI [-2.17, .09]. In general, the lowest cognitive load was reported by hard of hearing participants.

Comprehension

Expecting that non-syntactic segmentation would negatively affect comprehension, we conducted a 2 x 3 mixed ANOVA on segmentation (SS vs. NSS) and degree of hearing loss (hearing, hard of hearing, and deaf).

Table 2.15 Descriptive statistics for comprehension in Experiment 2

	Deafness	Mean (SD)
Comprehension SS	Hearing	4.11 (1.01)
	Hard of hearing	4.60 (.51)
	Deaf	4.00 (.70)
	Total	4.20 (.88)
Comprehension NSS	Hearing	4.26 (1.02)
	Hard of hearing	4.50 (.70)
	Deaf	3.44 (1.23)
	Total	4.15 (1.05)

Note*:* Maximum score was 5.

Despite our predictions, and similarly to Experiment 1, we found no main effect of segmentation on comprehension *F*(1,43) = .713, *p* = .403, η_p^2 = .016. There were no interactions.

As for between-subject effects, we found a marginally significant main effect of hearing loss on comprehension, *F*(2,43) = 3.061, *p* = .057, η_p^2 = .125. The highest comprehension scores were obtained by hard of hearing participants and the lowest by deaf participants (Table 2.15). Post-hoc analyses with Bonferroni correction showed that deaf participants differed from hard of hearing participants, p = .053, 95% CI [-1.66, .01].

Eye tracking measures

Due to problems with calibration, 10 participants had to be excluded from eye tracking analyses, leaving a total of 22 hearing, 8 hard of hearing, and 6 deaf participants.

Table 2.16 Mean eye tracking measures by segmentation in Experiment 2

	Hearing loss						
	Hearing	Hard of hearing	Deaf	Df	F	p	η_p^2
Absolute reading time (ms)				1,33	1.752	.195	.050
SS	1614	1619	1222				
NSS	1617	1519	1522				
Proportional reading time				1,33	2.270	.141	.064
SS	.65	.66	.45				
NSS	.66	.61	.62				
Mean fixation duration				1,33	.199	.659	.006
SS	209	199	214				
NSS	211	185	219				
Fixation count				1,33	2.686	.111	.075
SS	6.41	6.73	4.63				
NSS	6.45	6.45	5.90				
Revisits				1,33	.352	.557	.011
SS	.28	.20	.45				
NSS	39	.30	.15				

To examine whether the non-syntactically segmented text resulted in longer reading times, more revisits and higher mean fixation duration, we conducted an analogous mixed ANOVA. We found no main effect of segmentation on any of the eye tracking measures (Table 2.16), but a few interactions between segmentation and deafness: in absolute reading time, F(2,33) = 4,205, p = .024, η_p^2 = .203;

proportional reading time, $F(2,33) = 4{,}912$, $p = .014$, $\eta_p^2 = .229$; fixation count, $F(2,33) = 3{,}992$, $p = .028$, $\eta_p^2 = .195$; and revisits, $F(2,33) = 6{,}572$, $p = .004$, $\eta_p^2 = .285$.

We broke down the interactions with simple-effects analyses by means of post-hoc tests using Bonferroni correction. In the deaf group, we found an effect of segmentation on revisits approaching significance, $F(1,5) = 5.934$, $p = .059$, $\eta_p^2 = .543$. Deaf participants had more revisits in the SS condition than in the NSS one, $p = .059$. They also had a higher absolute reading time, proportional reading time, and fixation count in the NSS compared to the SS condition, but possibly owing to the small sample size, these differences did not reach statistical significance. In the hard of hearing group, there was no significant main effect of segmentation on any of the eye tracking measures (*ps* > .05). In the hearing group, there was no statistically significant main effect of segmentation (all *ps* > .05).

A between-subject analysis showed a close to significant main effect of degree of hearing loss on fixation count, $F(2,33) = 3.204$, $p = .054$, $\eta_p^2 = .163$. Deaf participants had fewer fixations per subtitle compared to hard of hearing, $p = .088$, 95% CI [-2.79, .14], or hearing participants, $p = .076$, 95% CI [-2.41, .08]. No other measures were significant.

2.3.3. Interviews

Following the eye tracking tests, we conducted short semi-structured interviews to elicit participants' views on subtitle segmentation, complementing the quantitative part of the study (Bazeley, 2013). We used inductive coding to identify themes reported by participants. Several Spanish, Polish, and deaf participants said that keeping units of meaning together contributed to the readability of subtitles because by creating false expectations (i.e. "garden path" sentences), NSS line-breaks can require more effort to process. These participants believed that chunking text by phrases according to "natural thoughts" allowed subtitles to be read quickly. In contrast, other participants said that NSS subtitles gave them a sense of continuity in reading the subtitles. A third theme in relation to dealing with SS and NSS subtitles was that participants adapted their reading strategies to different types of line-breaks. Finally, a number of people also admitted they had not noticed any differences in the subtitle segmentation between the clips, saying they had never paid any attention to subtitle segmentation.

2.4. Discussion

The two experiments reported in this paper examined the impact of text segmentation in subtitles on cognitive load and reading performance. We also investigated whether viewers' linguistic background (native language and hearing status) impacts on how they process syntactically and non-syntactically segmented subtitles. Drawing on the large body of literature on text segmentation in subtitling (Díaz Cintas & Remael, 2007; Ivarsson & Carroll, 1998; Perego, 2008a, 2008b; Rajendran et al., 2013) and literature on parsing and text chunking during reading (Keenan, 1984; Kennedy et al., 1989; LeVasseur, Macaruso, Palumbo, & Shankweiler, 2006; Mitchell, 1987, 1989; Rayner et al., 2012), we predicted that subtitle reading would be adversely affected by non-syntactic segmentation.

This prediction was partly upheld. One of the most important findings of this study is that participants reported higher cognitive load in non-syntactically segmented (NSS) subtitles compared to syntactically segmented (SS) ones. In both experiments, mental effort, difficulty, and frustration were reported as higher in the NSS condition. A possible explanation of this finding may be that NSS text increases extraneous load, i.e. the type of cognitive load related to the way information is presented (Sweller et al., 1998). Given the limitations of working memory capacity (Baddeley, 2007; Chandler & Sweller, 1991), NSS may leave less capacity to process the remaining visual, auditory, and textual information. This, in turn, would increase their frustration, make them expend more effort and lead them to perceive the task as more difficult.

Although cognitive load was found to be consistently higher in the NSS condition across the board in all participant groups, the mean differences between the two conditions do not differ substantially and thus the effect sizes are not large.

We believe the small effect size may stem from the fact that the clips used in this study were quite short. As cognitive fatigue increases with the length of the task, and declines simultaneously in performance (Ackerman & Kanfer, 2009; Sandry, Genova, Dobryakova, DeLuca, & Wylie, 2014; Van Dongen, Belenky, & Krueger, 2011), we might expect that in longer clips with non-syntactically segmented subtitles, the cognitive load would accumulate over time, resulting in more prominent mean differences between the two conditions. We acknowledge that the short duration of clips, necessitated by the length of the entire experiment, is an important limitation of this study. However, a number of previous studies on subtitling have also used very short clips (Jensema, 1998; Jensema, El Sharkawy, Danturthi, Burch, & Hsu, 2000; Rajendran et al., 2013; Romero-Fresco, 2015). In this study, we only examined text segmentation within a single subtitle; further research should also explore the effects of non-syntactic segmentation across two or more consecutive subtitles, where the impact of NSS subtitles on cognitive load may be even higher.

Despite the higher cognitive load and contrary to our predictions, we found no evidence that subtitles which are not segmented in accordance with professional standards result in lower comprehension. Participants coped well in both conditions, achieving similar comprehension scores regardless of segmentation. This finding is in line with the results reported by Perego et al. (2010), using Italian participants, that subtitles containing non-syntactically segmented noun phrases did not negatively affect participants' comprehension. Our research extends these findings to other linguistic units in English (verb phrases and conjunctions as well as noun phrases) and other groups of participants (hearing English, Polish, and Spanish speakers, as well as deaf and hard of hearing participants). The finding that performance in processing NSS text is not negatively affected despite the participants' extra effort (as shown by increased cognitive load) may be attributed to the short duration of the clips and also to overall high comprehension scores. As the clips were short, there were limited points that could be included in the comprehension questions. Other likely reasons for the lack of significant differences between the two conditions is the extensive experience that all the participants had of using subtitles in the UK, and that participants may have become accustomed to subtitling not adhering to professional segmentation standards. Our sample of participants was also relatively well-educated, which may have been a reason for their comprehension scores being near ceiling. Furthermore, as noted by Mitchell (1989), when interpreting the syntactic structure of sentences in reading, people use non-lexical cues such as text layout or punctuation as parsing aids, although these cues are of secondary importance when compared to words, which constitute "the central source of information" (p. 123). This is also consistent with what the participants in our study reported in the interviews. For example, one deaf participant said: "Line breaks have their value, yet when you are reading fast, most of the time it becomes less relevant."

In addition to understanding the effects of segmentation on subtitle processing, this study also found interesting results relating to differences in subtitle processing between the different groups of viewers. In Experiment 1, Spanish participants had the highest cognitive load and lowest comprehension, and spent more time reading subtitles than Polish and English participants. Although it is impossible to attribute these findings unequivocally to Spanish participants coming from a dubbing country, this finding may related to their experience of having grown up exposed more to dubbing than subtitling. In Experiment 2, we found that subtitle processing was the least effortful for the hard of hearing group: they reported the lowest cognitive effort and had the highest comprehension score. This result may be attributed to their high familiarity with subtitling (as declared in the pre-test questionnaire) compared to the hearing group. Although no data were obtained for the groups in Experiment 2 in relation to English literacy measures, as a group, individuals born deaf or deafened early in life have low average reading ages, and more effortful processing by the deaf group may be related to lower literacy.

Different viewers adopt different strategies to cope with reading NSS subtitles. In the case of hearing participants, there were more revisits to the subtitle area for NSS subtitles, which is a likely indication of parsing difficulties (Rayner et al., 2012). In the group of participants with hearing loss, deaf people spent more time reading NSS subtitles than SS ones. Given that longer reading time may indicate difficulty in extracting information (Holmqvist et al., 2011), this may also be taken to reflect parsing problems. This interpretation is also in accordance with the longer durations of fixations in the deaf group, which is another indicator of processing difficulties (Holmqvist et al., 2011; Rayner, 1998). Unlike the findings of other studies (Krejtz et al., 2016; Szarkowska et al., 2011; Szarkowska, Krejtz, Dutka, et al., 2016), in this study, deaf participants fixated less on the subtitles than hard of hearing and hearing participants. Our results, however, are in line with a recent eye tracking study (Miquel Iriarte, 2017), where deaf people also had fewer fixations than relation hearing viewers. According to Miquel Iriarte (2017), deaf viewers relate to the visual information on the screen as a whole to a greater extent than hearing viewers, reading the subtitles faster to give them more time to direct their attention towards the visual narrative.

2.5. Conclusions

Our study has shown that text segmentation influences the processing of subtitled videos: non-syntactically segmented subtitles may increase viewers' cognitive load and eye movements. This was particularly noticeable for Spanish and deaf people. In order to enhance the viewing experience, using syntactic segmentation in subtitles may facilitate the process of reading subtitles, thus giving viewers greater time to follow the visual narrative of the film. Further research is necessary to disentangle the impact of the viewers' country of origin, familiarity with subtitling, reading skills, and language proficiency on subtitle processing.

This study also provides support for the need to base subtitling guidelines on research evidence, particularly in view of the tremendous expansion of subtitling across different media and formats. The results are directly applicable to current practices in television broadcasting and video-on-demand services. They can also be adopted in subtitle personalization to improve automation algorithms for subtitle display in order to facilitate the processing of subtitles among the myriad different viewers using subtitles.

Ethics and Conflict of Interest

The author(s) declare(s) 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 (2017SGR113).

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

References

Ackerman, P. L., & Kanfer, R. (2009). Test Length and Cognitive Fatigue: An Empirical Examination of Effects on Performance and Test-Taker Reactions. *Journal of Experimental Psychology: Applied*, *15*(2), 163–181. https://doi.org/10.1037/a0015719

Allen, P., Garman, J., Calvert, I., & Murison, J. (2011). Reading in multimodal environments: assessing legibility and accessibility of typography for television. In *The proceedings of the 13th international ACM SIGACCESS conference on computers and accessibility* (pp. 275–276). https://doi.org/10.1145/2049536.2049604

Antonenko, P., Paas, F., Grabner, R., & van Gog, T. (2010). Using Electroencephalography to Measure Cognitive Load. *Educational Psychology Review*, *22*(4), 425–438. https://doi.org/10.1007/s10648-010-9130-y

Baddeley, A. (2007). *Working memory, thought, and action*. Oxford: Oxford University Press.

Bazeley, P. (2013). *Qualitative Data*. Los Angeles: Sage.

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

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.

Chandler, P., & Sweller, J. (1991). Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction*, *8*(4), 293–332. https://doi.org/10.1207/s1532690xci0804_2

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

d'Ydewalle, G., Praet, C., Verfaillie, K., & Van Rensbergen, J. (1991). Watching Subtitled Television: Automatic Reading Behavior. *Communication Research*, *18*(5), 650–666.

D'Ydewalle, G., Rensbergen, J. Van, & Pollet, J. (1987). Reading a message when the same message is available auditorily in another language: the case of subtitling. In J. K. O'Regan & A. Levy-Schoen (Eds.), *Eye movements: from physiology to cognition* (pp. 313–321). Amsterdam/New York: Elsevier.

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 T. Dwyer, C. Perkins, S. Redmond, & J. Sita (Eds.), *Seeing into Screens: Eye Tracking and the Moving Image* (pp. 46–64). London: Bloomsbury.

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

Gernsbacher, M. A. (2015). Video Captions Benefit Everyone. *Policy Insights from the Behavioral and Brain Sciences*, *2*(1), 195–202. https://doi.org/10.1177/2372732215602130

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam: North-Holland.

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.

Jensema, C. (1998). Viewer Reaction to Different Television Captioning Speeds. *American Annals of the Deaf*, *143*(4), 318–324.

Jensema, C., El Sharkawy, S., Danturthi, R. S., Burch, R., & Hsu, D. (2000). Eye Movement Patterns of Captioned Television Viewers. *American Annals of the Deaf*, *145*(3), 275–285.

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

Keenan, S. A. (1984). Effects of Chunking and Line Length on Reading Efficiency. *Visible Language*, *18*(1), 61–80.

Kennedy, A., Murray, W. S., Jennings, F., & Reid, C. (1989). Parsing Complements: Comments on the Generality of the Principle of Minimal Attachment. *Language and Cognitive Processes*, *4*(3–4), SI51-SI76. https://doi.org/10.1080/01690968908406363

Koolstra, C. M., Van Der Voort, T. H. A., & D'Ydewalle, G. (1999). Lengthening the Presentation Time of Subtitles on Television: Effects on Children's Reading Time and Recognition. *Communications*, *24*(4), 407–422. https://doi.org/10.1515/comm.1999.24.4.407

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., & Doherty, S. (2016). Measuring Cognitive Load in the Presence of Educational Video: Towards a Multimodal Methodology. *Australasian Journal of Educational Technology*, *32*(6), 19–31. https://doi.org/10.14742/ajet.3084

Kruger, J.-L., Doherty, S., Fox, W., & de Lissa, P. (2017). Multimodal measurement of cognitive load during subtitle processing: Same-language subtitles for foreign language viewers. In I. Lacruz & R. Jääskeläinen (Eds.), *New Directions in Cognitive and Empirical Translation Process Research* (pp. 267–294). London: John Benjamins.

Kruger, J.-L., Hefer, E., & Matthew, G. (2013). Measuring the impact of subtitles on cognitive load: eye tracking and dynamic audiovisual texts. *Proceedings of the 2013 Conference on Eye Tracking South Africa*. Cape Town, South Africa: ACM. https://doi.org/10.1145/2509315.2509331

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. https://doi.org/http://dx.doi.org/10.16910/jemr.7.5.4

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. M., Macaruso, P., Palumbo, L. C., & Shankweiler, D. (2006). Syntactically cued text facilitates oral reading fluency in developing readers. *Applied Psycholinguistics*, *27*(3), 423–445. https://doi.org/10.1017/s0142716406060346

Łuczak, K. (2017). *The effects of the language of the soundtrack on film comprehension, cognitive load and subtitle reading patterns. An eye-tracking study.* University of Warsaw, Warsaw.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967.

Matamala, A., Perego, E., & Bottiroli, S. (2017). Dubbing versus subtitling yet again? *Babel*, *63*(3), 423–441. https://doi.org/10.1075/babel.63.3.07mat

Mayberry, R. I., del Giudice, A. A., & Lieberman, A. M. (2011). Reading Achievement in Relation to Phonological Coding and Awareness in Deaf Readers: A Meta-analysis. *Journal of Deaf Studies and Deaf Education*, *16*(2), 164–188. https://doi.org/10.1093/deafed/enq049

Miquel Iriarte, M. (2017). *The reception of subtitling for the deaf and hard of hearing: viewers' hearing and communication profile & Subtitling speed of exposure*. Universitat Autònoma de Barcelona, Barcelona.

Mitchell, D. C. (1987). Lexical guidance in human parsing: Locus and processing characteristics. In M. Coltheart (Ed.), *Attention and Performance XII*. London: Lawrence Erlbaum Associates Ltd.

Mitchell, D. C. (1989). Verb guidance and other lexical effects in parsing. *Language and Cognitive Processes*, *4*(3–4), SI123-SI154. https://doi.org/10.1080/01690968908406366

Musselman, C. (2000). How Do Children Who Can't Hear Learn to Read an Alphabetic Script? A Review of the Literature on Reading and Deafness. *Journal of Deaf Studies and Deaf Education*, *5*(1), 9–31.

Ofcom. (2015). Ofcom's Code on Television Access Services.

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

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

Romero-Fresco, P. (2015). *The reception of subtitles for the deaf and hard of hearing in Europe*. Bern: Peter Lang.

Sandry, J., Genova, H. M., Dobryakova, E., DeLuca, J., & Wylie, G. (2014). Subjective Cognitive Fatigue in Multiple Sclerosis Depends on Task Length. *Frontiers in Neurology*, *5*, 214. https://doi.org/10.3389/fneur.2014.00214

Schmeck, A., Opfermann, M., van Gog, T., Paas, F., & Leutner, D. (2014). Measuring cognitive load with subjective rating scales during problem solving: differences between immediate and delayed ratings. *Instructional Science*, *43*(1), 93–114. https://doi.org/10.1007/s11251-014-9328-3

Sun, S., & Shreve, G. (2014). Measuring translation difficulty: An empirical study. *Target*, *26*(1), 98–127. https://doi.org/10.1075/target.26.1.04sun

Sweller, J. (2011). Cognitive Load Theory. *Psychology of Learning and Motivation*, *55*, 37–76.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). Cognitive Load Theory. *Psychology of Learning and Motivation*, *55*, 37–76.

Sweller, J., Van Merrienboer, J., & Paas, F. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, *10*(3), 251–296. https://doi.org/10.1023/a:1022193728205

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

Szarkowska, A., Krejtz, I., Kłyszejko, 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).

Szarkowska, A., Krejtz, I., Pilipczuk, O., Dutka, Ł., & Kruger, J.-L. (2016). The effects of text editing and subtitle presentation rate on the comprehension and reading patterns of interlingual and intralingual subtitles among deaf, hard of hearing and hearing viewers. *Across Languages and Cultures*, *17*(2), 183–204. https://doi.org/10.1556/084.2016.17.2.3

Szarkowska, A., Krejtz, K., Dutka, Ł., & Pilipczuk, O. (2016). Cognitive load in intralingual and interlingual respeaking-a preliminary study. *Poznan Studies in Contemporary Linguistics*, *52*(2). https://doi.org/10.1515/psicl-2016-0008

Thorn, F., & Thorn, S. (1996). Television Captions for Hearing-Impaired People: A Study of Key Factors that Affect Reading Performance. *Human Factors*, *38*(3), 452–463.

Van Dongen, H. P. A., Belenky, G., & Krueger, J. M. (2011). *Investigating the temporal dynamics and underlying mechanisms of cognitive fatigue.* American Psychological Association. https://doi.org/10.1037/12343-006

Van Gerven, P. W. M., Paas, F., Van Merriënboer, J. J. G., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology*, *41*(2), 167–174. https://doi.org/10.1111/j.1469-8986.2003.00148.x

van Gog, T., & Paas, F. (2008). Instructional Efficiency: Revisiting the Original Construct in Educational Research. *Educational Psychologist*, *43*(1), 16–26. https://doi.org/10.1080/00461520701756248

Wang, Z., & Duff, B. R. L. (2016). All Loads Are Not Equal: Distinct Influences of Perceptual Load and Cognitive Load on Peripheral Ad Processing. *Media Psychology*, *19*(4), 589–613. https://doi.org/10.1080/15213269.2015.1108204

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

Yoon, J.-O., & Kim, M. (2011). The Effects of Captions on Deaf Students' Content Comprehension, Cognitive Load, and Motivation in Online Learning. *American Annals of the Deaf*, *156*(3), 283–289. https://doi.org/10.1353/aad.2011.0026

Chapter 3. Article 2

Line breaks in subtitling:

An eye tracking study on viewer preferences

3. Article 2

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. DOI: 10.16910/jemr.11.3.2

Abstract

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

3.1. 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 et al., 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.

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.

3.1.1. 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 et al., 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.

3.1.2. 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).

3.1.3. 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.

3.1.4. 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 et al., 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 et al., 2018) 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 et al., 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.

3.1.5. 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 et al., 2014; Krejtz et al., 2013; Kruger & Steyn, 2014; Kruger et al., 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 et al., 2009; Monreal & Hernandez, 2005; Szarkowska et al., 2011). One of the difficulties experienced by deaf people when reading is related to definite and indefinite articles (Channon & Sayers, 2007; Wolbers et al., 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 et al., 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.

3.1.6. 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 et al., 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; R. 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.

3.2. 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.

3.2.1. 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 et al., 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, $\chi2(3)=5.144$, $p=.162$.

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.

3.2.2. 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).

3.2.3. 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 3.1). 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).

Figure 3.1 Stimulus example with syntactically segmented (left) and non-syntactically segmented text (right)

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 3.1 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 3.1 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.

3.2.4. 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.

3.2.5. Dependent variables

The dependent variables were the preference score and three eye tracking measures (see Table 3.2). 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 3.2).

Figure 3.2 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 3.2.

Table 3.2 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, 2018b).

3.2.6. 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.

3.3. Results

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

3.3.1. Results from 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 3.3). The SS subtitles were preferred over the NSS ones.

Figure 3.3 Preferences for SS and NSS subtitles by linguistic units in Experiment 1

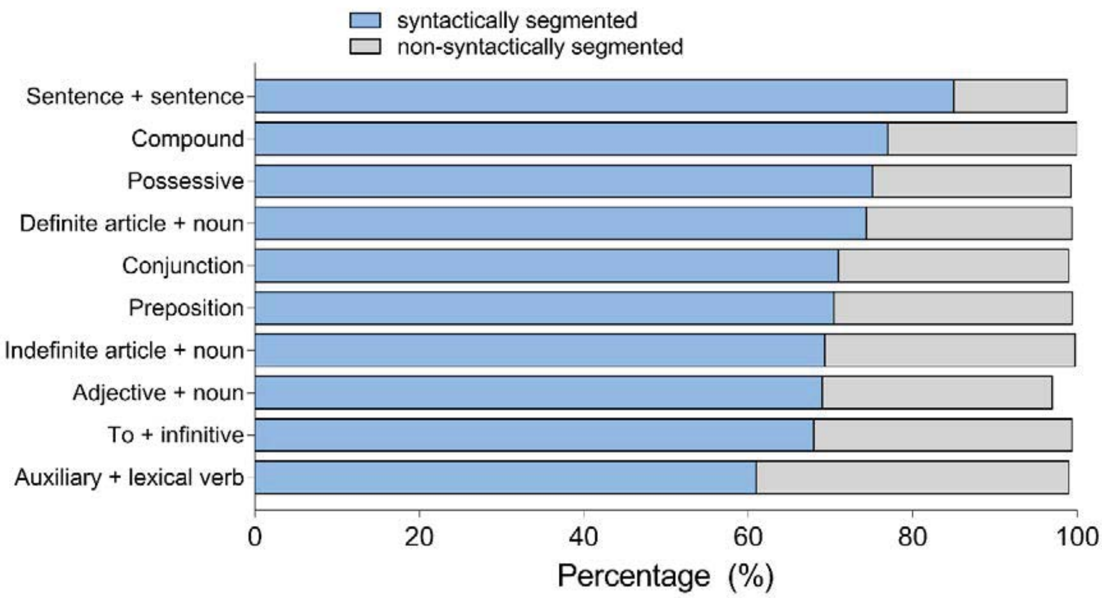


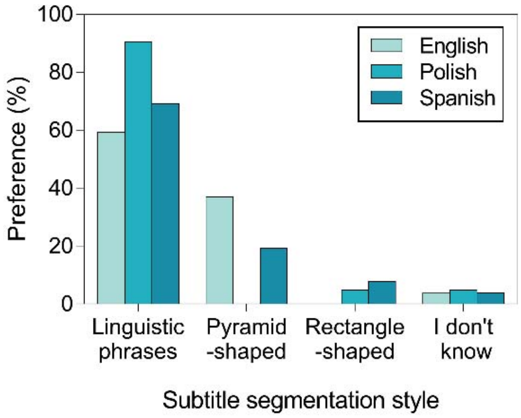
Figure 3.3 shows preferences by linguistic units and Table 3 by participant groups. There were no differences between groups in any of the linguistic conditions and no interactions. This means that regardless of their mother tongue, all participants had similar preferences.

Table 3.3 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 3.4, 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 3.4 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 3.4). 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 3.4 Dwell Time on subtitles by linguistic unit and segmentation (ms)

	Language						
Linguistic unit split	English	Polish	Spanish	*df*	*F*	*P*	η_p^2
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 3.5), 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 3.6). 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 3.5 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 3.6 ANOVA results for between-subject effects in mean fixation duration

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 3.7). 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, η_p^2=.114, and *AdjN, F*(2,53)=7.696, *p*=.001, η_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, η_p^2=.551, and *AdjN*, *F*(1,17)=5,017, *p*=.039, η_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, η_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, η_p^2=.207, but only a tendency in *AdjN*, *F*(1,21)=3.980, *p*=.059, η_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, η_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 3.7 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				

3.3.1.1. 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.

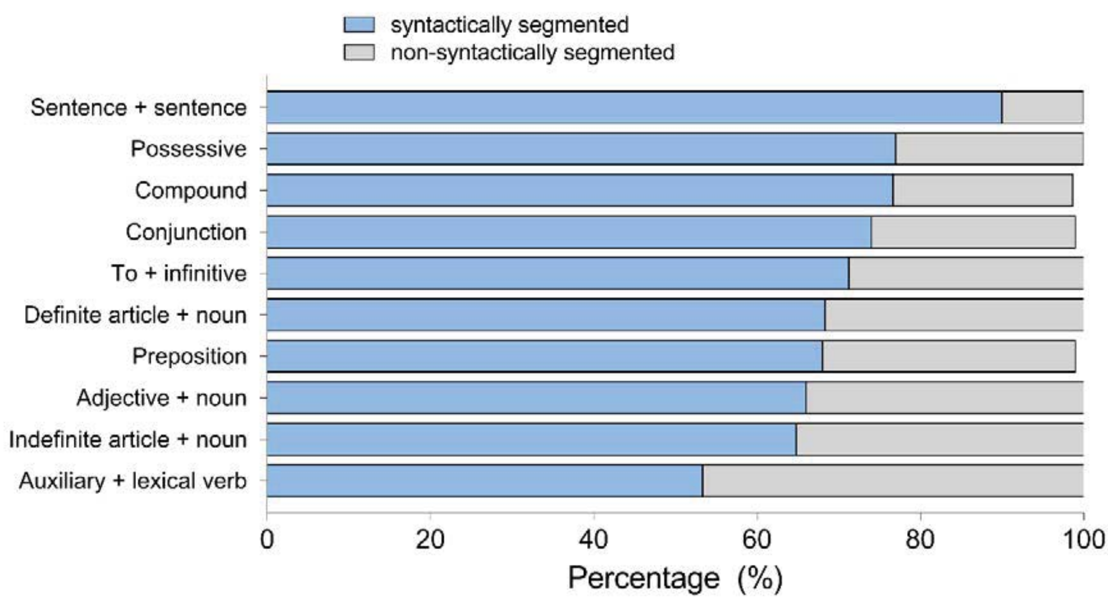
3.3.2. Results from Experiment 2

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 3.5 presents general preferences for all linguistic units and Table 3.8 shows how they differed by hearing loss.

Figure 3.5 Preferences for SS and NSS subtitles by linguistic units in Experiment 2



We found an almost significant interaction between segmentation and hearing loss in *DefArt*, *F*(2,37)=3.086, *p*=.058, η_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, η_p^2=.492, as well as for hard of hearing participants, *F*(1,9)=7.111, *p*=.026,

η_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.

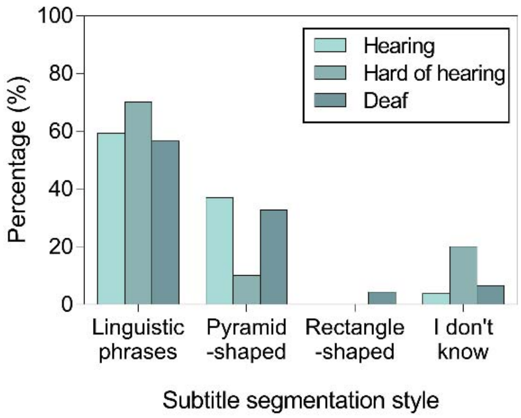
There was a main effect of hearing loss in *AdjN*, *F*(2,37)=3.469, *p*=.042, η_p^2=.158 and a tendency approaching significance in Comp, *F*(2,37)=3.063, *p*=.059, η_p^2=.142. Post-hoc Bonferroni tests showed that hearing participants tended to express higher preference for SS *AdjN* than hard of hearing 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.

Table 3.8 Percentage of Experiment 2 participants who preferred the syntactically segmented condition

	Degree of hearing loss						
Linguistic unit	Hearing	Hard of hearing	Deaf	*df*	*F*	*p*	η_p^2
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

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 3.6.

Figure 3.6 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 3.9). 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, η_p^2= .228) and a tendency approaching significance in *SentSent* (*F*(2,26)=3.241, *p*=.055, η_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 3.9 Dwell Time by linguistic unit and segmentation (ms)

	Degree of hearing loss						
Linguistic unit split	Hearing	Hard of hearing	Deaf	*df*	*F*	*p*	η_p^2
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 3.10) 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 3.10 Mean Fixation Duration by linguistic unit and segmentation

	Degree of hearing loss						
Linguistic unit split	Hearing	Hard of hearing	Deaf	*df*	*F*	*p*	η_p^2
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, η_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, η_p^2=.635. There was no effect for English or hard of hearing participants.

Table 3.11 Revisits by linguistic unit and segmentation

Linguistic unit split	Degree of hearing loss: Hearing	Hard of hearing	Deaf	*df*	*F*	*P*	η_p^2
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				

3.3.2.1. 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.

3.3.3. 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.”

3.4. 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 <u>the memory stick</u>
he managed to hide.

and

(NSS) He's looking for <u>the memory</u>
<u>stick</u> 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 et al., (2018)), 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 (2017SGR113).

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. *Journal of Deaf Studies and Deaf Education*, *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., & Woll, B. (2018). The impact of text segmentation on subtitle reading. *Journal of Eye Movement Research*, *11*(4), 1–18. DOI: http://dx.doi.org/10.16910/11.4.2

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

Chapter 4. Article 3

The effects of screen size on subtitle layout preferences and comprehension across devices

4. Article 3

Gerber-Morón, O., Soler-Vilageliu, O., & Castellà, J. (forthcoming). The effects of screen size on subtitle layout preferences and comprehension across devices. *Hermēneus*, *Revista de traducción e interpretación* 2019, *21*. Manuscript accepted for publication.

Abstract

The present study sheds light on the possible effects that screen size can have on preferences and comprehension of subtitled audiovisual material content. Thirty participants watched three subtitled video excerpts displayed on three devices with different screen size (monitor, tablet, and smartphone). After watching each excerpt, they filled out preference and comprehension questionnaires. This study aimed to provide new empirical evidence on viewers' needs and preferences concerning readability by analysing the reception of subtitles across screens. The results obtained indicate that smartphone devices had the most unsatisfactory effects, suggesting the need to undertake further research on small screens to improve subtitle readability.

Keywords: accessibility; new technologies; readability; screen size; subtitling

4.1. Introduction

The development of new technologies in the past decades has changed the way audiovisual products are consumed nowadays (Messerlin et al., 2005). Innovative handheld devices, such as tablets and smartphones, provide the mobility to consume media everywhere (Palen et al., 2000). The implementation of subtitles on these handheld devices makes video content accessible to different end-users, such as non-native speakers, deaf and hard-of-hearing viewers. Subtitles on mobile devices are also useful when sound has to be removed in public spaces. Because watching subtitled media on these devices is continually increasing in our society, it is important to present subtitles in the most effective way. This study examines the effects of screen size on different subtitle layout parameters, with a view to improving the most determining factor in subtitling: readability. The process of readability becomes more complex with subtitled media because viewers are continually switching from text to image (D'Ydewalle et al., 1987), without having control over the speed in subtitling (Romero-Fresco, 2015). We predict that subtitle readability may be hindered by the smaller screen size of handheld devices.

4.1.1. Readability and layout parameters in subtitles

Scholars have established various parameters that need to be considered to improve the readability of subtitles. Karamitroglou (1998) and Perego (2005, 2008b) distinguished three categories of parameters that affect the legibility and readability of subtitles: duration, text editing, and subtitle layout parameters. Duration parameters comprise the line length of time the subtitles are on the screen, the leading-in and lagging-out time for each subtitle, the time break between two consecutive subtitles, and camera takes and cuts (Perego, 2005, 2008b) Text editing parameters relate to punctuation and letter case, line breaks and line length, altering syntactic structures, omitting and retaining linguistic items of the original. In relation to layout parameters, Gottlieb (1992) compiled a list that included the position of subtitles on the screen, the number of lines, the number of characters per line, text alignment, typeface and distribution, and font colour and background.

Media regulators and professionals in the audiovisual industry have partly integrated these parameters in their guidelines to enhance the quality of subtitling (BBC, 2017; DCMP, 2017; Media Access Australia, 2012; Ofcom, 2017). As a case in point, BBC's subtitles guidelines (BBC, 2017) recommends the use of one-line subtitle instead of two short lines because it takes less time to read and causes less disruption to the picture. In our study, we tested some of the layout parameters listed by the scholars previously mentioned to examine how viewers perceive subtitles across devices.

4.1.2. Screen size effects across devices

To the best of our knowledge, subtitle layout parameters have not been studied across devices. Nevertheless, other studies have been conducted on the effects of screen size in the fields of Audiovisual Translation. Two eye-tracking studies on watching subtitled videos across screen devices have shown more negative results in smartphone devices (Castellà et al., 2016; Szarkowska et al., 2015). Szarkowska et al. (2015) studied reading patterns on smartphone, tablet and computer screen, and found evidence that smartphone has the lowest comprehension results, the longest mean fixation duration, and fewer fixations in comparison to tablet and monitor screens. In their eye-tracking study on watching subtitled videos on different screen devices, (Castellà et al., 2016) suggested that smartphone devices require more cognitive load when reading subtitles than tablets and monitors.

A number of studies in the fields of Media Psychology, and Human-Computer Interaction (Al-Showarah et al., 2014; Kim et al., 2011; Lombard et al., 1997; Maniar et al., 2008) have also focused their research of screen size on viewers' perception of mobility and content, and on attitudes towards technology. Lombard et al. (1997) studied the role of screen size in small and large television screens. They measured responses via a questionnaire and found that large screen televisions elicit more intense responses for some genres (commercials, action-adventure, and reality) but not for others (talk shows and drama programs). Maniar et al. (2008) looked at the effect of screen size on video-based learning by presenting videos on small, medium and large screen mobile phones. Their results from the questionnaires pointed out that larger screens induce more attention than medium and small screens. Moreover, they found that smaller screen displays may inhibit the effectiveness of the learning experience. Kim et al. (2011) carried out a study on the effects of screen size (across three different mobile phone devices) and communication modality (video format or text document) to assess through questionnaires the users' perception of mobility and content, and the degree of technology acceptance. Their results revealed that screen size does not affect the understanding of the news story or the perceived ease of use of the device. Nevertheless, it seems that larger screen size is the key to greater enjoyment for their participants. In another study, Al-

Showarah et al. (2014) evaluated the effects of screen size on smartphone and tablet usability across age groups and found that seniors show more difficulties in processing information on smartphone screens. Their eye-tracking results also showed that usability on a small screen size is more difficult for all age groups in comparison to large screen sizes. In general, all these studies indicate that large screen displays tend to contribute to a more satisfying experience.

4.1.3. Overview of the study

The main goal of this study is to analyse different subtitle layout preferences and comprehension scores across devices with different screen size (monitor, tablet, and smartphone). This study aims to offer more insight on viewers' needs and preferences by specifically testing subtitle layout parameters not covered in previous studies on subtitles across devices (Castellà et al., 2016; Szarkowska et al., 2015). This study also aims to validate (Szarkowska et al., 2015) as regards the effects of screen size on comprehension scores.

Based on the studies previously conducted on screen size (Al-Showarah et al., 2014; Castellà et al., 2016; Maniar et al., 2008; Szarkowska et al., 2015), we expected to find differences in the overall viewers' evaluation of the subtitle layout parameters depending on the device, and in particular, we expected to obtain more negative results in the smallest screen devices (i.e. smartphones).

4.2. Methods

4.2.1. Participants

The study involved 30 volunteer participants ranging from 18 to 58 years of age (16 females, 14 males, mean age=30.5, SD=7.6). They were all Spanish native speakers or Catalan-Spanish bilingual with normal or corrected-to-normal (contact lenses or glasses) vision. Most of the participants were university students from Spain or other Spanish-speaking countries. The majority of the participants reported not being habitual viewers of subtitled audiovisual material. None of the participants had any knowledge of the original language used for the film fragments (Norwegian).

4.2.2. Materials

4.2.2.1. Stimuli and apparatus

Video fragments

The stimuli were three short video fragments with Spanish subtitles taken from a Norwegian thriller (*Hodejegerne*, Tyldum, 2011). Each video fragment formed full scenes with coherent content, and the average duration of each of them was three minutes. We used a Norwegian film to expose participants to an unknown language, so that they would have to rely on the information provided by the subtitles to follow the video fragments.

Subtitles

The subtitles were created using EZTitles[7], a professional subtitle editing software. As for the technical considerations, we followed the recommendations by Díaz Cintas and Remael (2007) for synchronization and presentation, using 15 subtitling spaces per second and lines of 38 characters. On average, each video fragment contained 37 subtitles: 15 sentences occupied one line of text and 22 occupied two lines. The video fragments and synchronised subtitles were presented using the freeware VLC Media Player on the three devices tested in the experiment: a 22-inch Toshiba TV monitor, a 9.7-inch iPad 2 and a 3.5-inch iPhone 4.

7 For more detailed information, see <http://www.eztitles.com> (last accessed 30 November 2017).

4.2.2.2. Questionnaires

Questionnaire on reading and layout preferences for subtitles

The preference questionnaire was administered to check viewers' reception of subtitles and their preferences concerning the general layout according to the screen size of each device. The questions for this study were inspired by Gottlieb (1992) and Gambier (2009), who provided a list of subtitle parameters to measure and evaluate the viewers' reception of subtitle readability. The questionnaire assessed the experience of reading subtitles on each device by asking questions on the following parameters:

- the percentage of subtitles read;
- the ease of subtitle reading;
- the overall assessment on viewing the film excerpt on that device;
- the feeling of having lost essential parts of the plot (due to the fact of reading subtitles);
- the line length of the subtitles;
- the exposure time of subtitles;
- the line-break layout (i.e. division of lines on screen).

The questionnaire on reading and layout preferences for subtitles for this study included three questions on a 5-point Likert scale concerning the percentage of subtitles read ("What percentage of subtitles didn't you have time to read?" from 0% to 100%), the ease of reading subtitles on that device ("How did you find reading subtitles on this device?" from very difficult to very easy) and the experience of viewing the film excerpt on that device ("How would you rate the experience of watching a film on this device?" from very unpleasant to very pleasant). Furthermore, a yes/no question asked about the feeling of having lost essential parts of the film's action due to the fact of reading subtitles ("Do you think that you lost essential parts of the film's action due to the fact of having to read subtitles?"). Here is a sample of one of these questions on the ease of reading subtitles:

Example (1)

How did you find reading subtitles on this device?

- Very easy
- Easy
- Moderate
- Difficult
- Very difficult

In addition to these questions, participants had to answer three categorical questions on a 5-point scale about subtitle preferences for the line length and exposure time on the screen ("What do you think about the length of the subtitles for this device?" and "What do you think about the exposure time of the subtitles on this screen?", where 1=very long, 2=long, 3=appropriate, 4=short, 5=very short), and line-break layout ("What do you think about the line-break layout for this device?", where 1=unsuitable, 2=I would have preferred shorter subtitles of one line, 3=I would have preferred shorter subtitles, but two lines, 4=I would have preferred longer subtitles, but of one line, 5=appropriate). Here is a sample of one of these categorical questions on the line length of subtitles:

Example (2)

What do you think about the length of the subtitles for this device?

- 1 = Very long
- 2 = Long
- 3= Appropriate
- 4 = Short
- 5 = Very short

Comprehension questionnaire

The comprehension questionnaire included a set of multiple-choice questions to verify whether participants understood the main textual information provided by the subtitles. After watching each film fragment, participants had to answer a set of five questions about the content of the video. For each question, participants were asked to complete a statement by selecting a response from a list of four items including a correct answer, two distractors and an "I don't remember" response option. We based the design and procedure of the comprehension questionnaire on Day & Park (2005) and (Lavaur & Bairstow, 2011). This is an example of the questions asked for comprehension assessment:

Example (3)

Roger's wife showed him a painting by…

- Rembrandt
- Rubens
- Jordaens
- I don't remember

4.2.2.3. Design and Procedure

Participants watched three video excerpts with subtitles displayed on three devices with different screen sizes (monitor, tablet, and smartphones). They watched each video excerpt on a different device according to a within-subject design. We counterbalanced the order of the viewing of the film fragments following a Latin-square design. The screen size of the devices was the independent variable tested in the experiment, whereas the main dependent variables were preferences and comprehension measured through the evaluative questionnaire on subtitle reading and user preferences for the subtitle layout, as well as the multiple-choice questionnaire on general comprehension for each device. A pilot study was carried out prior to the main study to validate the experiment.

Each participant was tested individually in a laboratory. We informed participants that the study was on subtitled-film watching, but we did not provide additional information on the specific parameters tested. The participants signed a consent form, and read the experiment instructions. We did not reveal the audio language of the film to the participants. They watched one of the three-minute subtitled film excerpts on one of the devices. After viewing the video excerpt, participants were asked to fill out the preferences and comprehension questionnaires before watching the subsequent video fragments. They were asked to watch the remaining two video excerpts on the other devices and answer the questionnaires in the same way as they did for the first excerpt. After completing the last questionnaire, participants filled the demographic and control variable questionnaire on the preferred type of audiovisual translation (e.g. dubbing, subtitling, voice-over), gender, age and native language. The experiment lasted approximately 20 minutes.

4.3. Results

4.3.1. Comprehension and readability items

First, one-way ANOVA tests were performed on the comprehension scores and on the three preference rating scores on readability (the percentage of subtitles read, the ease of reading subtitles and the experience of viewing the film excerpt on that device). None of the ANOVAs showed significant differences, suggesting that the type of device did not affect comprehension ($F(2,58)=.677$; $p=.51$). Participants were able to read the same percentage of subtitles in each device ($F(2,58)=.081$; $p=.92$), the three devices were equally readable in terms of ease ($F(2,58)=.979$; $p=.38$), and the experience did not differ as a function of device either ($F(2,58)=.548$; $p=.58$). The mean rating for each question as a function of device can be seen in (Table 4.1).

Table 4.1 Mean rating (standard deviations) for comprehension and for each preference question as a function of device, values ranged from 1 to 5, where 5 indicated more positive ratings

	Comprehension	% Not read	Ease of reading	Overall experience
Smartphone	4.20 (1.03)	4.27 (.69)	4.13 (.97)	3.07 (1.05)
Tablet	4.40 (.77)	4.30 (.98)	4.27 (.98)	3.30 (1.60)
Monitor	4.13 (1.01)	4.23 (.89)	4.40 (.77)	3.40 (1.71)

Note: 1 indicates participants could not read 100% of the subtitles, whereas 5 means that they were able to read them all (0% not read).

Regarding the yes/no question about the feeling of having lost essential parts of the film's action due to the fact of reading subtitles, Chi-Square tests of independence revealed no significant differences between the percentage of "Yes" and "No" responses for the smartphone device (40% "Yes" versus 60% "No"), $\chi 2$ (1, n = 30) =1.20, p=.27. However, there was a significant difference between percentages for both tablet and monitor devices (26.7% "Yes" versus 73.3% "No"), $\chi 2$ (1, n = 30) = 6.53, p=.01 in both questions.

4.3.2. Subtitle preference items

Finally, the categorical questions about subtitle preferences for the line length, exposure time, and division of lines on screen (line breaks) were analysed. A related-samples Friedman's Two-Way Analysis of Variance by Ranks was performed in order to compare the distributions of answers across devices for these three categorical questions that involved non-normally distributed data. The distributions of percentages and frequencies did not significantly change across devices for the three categorical preference questions with p= .761 in line length, p= 1 in exposure time and p= .913 in division of lines on screen. Therefore, the null hypothesis of distribution of percentages and frequencies can be accepted: screen size did not have any significant effect on subtitle reading preferences.

However, we observed some trends in the Chi-Square tests performed for each of these categorical questions, revealing significant differences between percentages in each question within each device (all p<.001). Across the three devices, results showed that the majority of the participants found that the line length of the subtitles was appropriate, especially for the tablet device (86.7%). However, there was a minor tendency to report the line length for the monitor device as long (20%) and, at a slightly lower percentage, for the smartphone device (13.3%). These trends can be seen in (Table 4.2).

Table 4.2 Question "What do you think about the length of the subtitles for this device?", percentage of each response option and Chi-Square values, as a function of device

Length	Very long	Long	Appropriate	Short	Very short	*χ2* (2, *n*=30)
Smartphone	0%	13.3%	80%	6.7%	0%	29.60, p<.001
Tablet	0%	3.3%	86.7%	10%	0%	38.60, p<.001
Monitor	0%	20%	70%	10%	0%	18.60, p<.001

Likewise, and as can be seen in Table 4.3, the majority of the participants reported the exposure time for the subtitles as appropriate, especially for the smartphone device (86.7%). However, some of the participants stated that the exposure time for the subtitles was short for the tablet (16.7%) and monitor (13.3%) devices.

Table 4.3 Question "What do you think about the exposure time of the subtitles on this screen?", percentage of each response option and Chi-Square values, as a function of device

Exposure time	Very long	Long	Appropriate	Short	Very short	*χ2* (2, *n*=30)
Smartphone	0%	6.7%	86.7%	6.7%	0%	38.40, p<.001
Tablet	0%	6.7%	76.7%	16.7%	0%	25.80, p<.001
Monitor	0%	10%	76.7%	13.3%	0%	25.40, p<.001

As can be seen in Table 4.4, the findings highlight that the majority of the participants found the line-break layout appropriate, in particular for the tablet device (75.9% vs. 56.7% for smartphone and 66.7% for monitor), although some other minor slight tendencies were detected. In fact, 20% of the participants would have preferred longer subtitles in one line for the smartphone device, and 16.7% of the participants would have preferred shorter subtitles in two lines for the monitor device.

Table 4.4 Question "What do you think about the line-break layout for this device?", percentage of each response option and Chi-Square values, as a function of device

Line breaks	Inappropriate	Shorter, in 1 line	Shorter, in 2 lines	Longer, in 1 line	Appropriate	$\chi 2$ (2, n=30)
Smartphone	0%	10%	13.3%	20%	56.7%	16.67, p=.001
Tablet	0%	10%	6.9%	6.9%	75.9%	40.10, p<.001
Monitor	0%	10%	16.7%	6.7%	66.7%	28.40, p<.001

4.4. Discussion

The experiment in this paper examined the influence of screen size on viewers' subtitle layout preferences and comprehension scores across three devices (monitor, tablet, and smartphone). The main aim was to provide additional data to the two previous studies on watching subtitled content across devices (Castellà et al., 2016; Szarkowska et al., 2015) by analysing subtitle layout parameters that have not been previously studied. Another aim was to validate the comprehension scores by Szarkowska et al. (2015). Drawing on the previous studies on screen size (Al-Showarah et al., 2014; Castellà et al., 2016; Maniar et al., 2008; Szarkowska et al., 2015), we predicted that the smallest screen device would give the most unsatisfactory results. We also expected to see differences in the viewers' reception of subtitle layout parameters across devices.

Regarding comprehension, no differences were found across screens. The findings are in line with the *subtitle effectiveness hypothesis* (Perego et al., 2010), which suggests that viewers can adapt their reading and visual skills for any screen displays. Contrary to the findings on comprehension scores by Szarkowska et al. (2015), our results imply that screen size is not a limitation and does not have a considerable impact on viewers processing subtitles across devices.

The results on the readability items indicate that screen size does not affect the viewers' reception of subtitles across devices in terms of the percentages of subtitles read, the ease of subtitle readability, and the overall experience on each device. Moreover, the majority of the participants declared that they did not have the feeling of having lost essential parts of the film's action due to the fact of reading subtitles. However, we only found significant results for this yes/no question for tablet and monitor screens. This finding shows that viewers feel capable of perceiving the incorporation of subtitles into tablets and monitors, in such a way as not to miss information from the rest of the audiovisual components. As for smartphone screens, results are not significant regarding this yes/no question. Our interpretation is that there is a broader range of opinions for smartphone screens because viewers do not perceive these screens as optimal as other screens. They may not feel as confident reading subtitles on these small screens as on larger devices. This result is consistent with Kim et al. (2011), which suggested that larger screen size devices are the key to greater enjoyment.

Although viewers found subtitles appropriate in terms of line length, exposure time and line-break layout, and the results did not provide significant differences across screens, some tendencies were detected. Regarding the line length of subtitles, tablets seem to provide the highest satisfaction. This is probably because the tablet display offers the right balance between each subtitle line and its medium size, not forcing the eyes to move much, compared with larger screens (i.e. monitor). The exposure time of subtitles on screens was found to be slightly more appropriate for smartphones than tablets or monitors. A possible explanation of this finding may be that smaller screens minimize the tendency to focus on other elements of the scene because of the limited size. On the contrary, the inclusion of subtitles on large screens can divert viewers' attention from the rest of the audiovisual components, as mentioned in some studies (Lombard et al., 1997; Maniar et al., 2008). Data related to line-break layout did not indicate a clear preference for smartphone and monitor screens: slightly more than half of the participants stated that the line-break layout was appropriate, but a minor tendency preferred shorter subtitles of two lines for monitor screens, and another minor trend preferred longer subtitles of one line for smartphones. Tablet screens seem to be the device with the highest satisfaction ratings regarding line-break layout. The results from these three categorical questions are not consistent with our initial hypothesis about finding differences in the overall evaluation of subtitles depending on the screen size.

4.5. Conclusions

Our study represents the first piece of knowledge on the effects of screen size on subtitle layout parameters, and it validates previous findings on comprehension scores across devices (Szarkowska et al., 2015). Our main finding shows that participants adapt their viewing skills to different screen sizes to process short subtitled film clips, and are generally satisfied with the subtitle layout on the devices tested. We also found that screen size does not affect comprehension levels.

We acknowledge that the general profile in this experiment included university students, the average age was 30 years old, and all participants belonged to a dubbing country. We believe that differences in comprehension and preferences of the subtitles could be found if other user profiles with different technological and audiovisual material habits were tested in the experiment (e.g. children, the deaf or the elderly).

Based on the trends found for some of the subtitle layout parameters, we think that more empirical studies focused on smartphone devices. Participants felt more comfortable reading subtitles in the larger screens (monitor and tablet): they did not have the impression of losing visual information and were more satisfied with tablet screens regarding subtitle line length and line-break layout. Our results for smartphone screens were not conclusive in terms of subtitle layout parameters, and do not validate the comprehension results by Szarkowska et al. (2015). New subtitle experiments on smartphone devices could also validate the results by Castellà et al. (2016), who found a different exploration pattern on viewers reading subtitles on smartphone devices: when reading subtitles in smartphone screens, there are fewer fixations but longer in duration compared to the other devices. Moreover, to our best knowledge, there are no empirical studies on reading subtitles specifically on smartphone screens. Further research could explore different types of line-break layouts on these devices to measure the impact of this variable on comprehension, readability, and enjoyment of audiovisual products.

References

Al-Showarah, S., AL-Jawad, N., & Sellahewa, H. (2014). Effects of User Age on Smartphone and Tablet Use, Measured with an Eye-Tracker via Fixation Duration, Scan-Path Duration, and Saccades Proportion. In C. Stephanidis & M. Antona (Eds.), *UAHCI/HCII, Part II, LNCS 8514* (pp. 3–14). Springer International Publishing Switzerland.

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

Castellà, J., Olivier, A., Gerber-Morón, O., & Soler-Vilageliu, O. (2016). Smartphone, Tablet or PC: Which one is better to watch your favourite subtitled film? In *International Meeting of the Psychonomic Society*. Granada (Spain).

D'Ydewalle, G., Rensbergen, J. Van, & Pollet, J. (1987). Reading a message when the same message is available auditorily in another language: the case of subtitling. In J. K. O'Regan & A. Levy-Schoen (Eds.), *Eye movements: from physiology to cognition* (pp. 313–321). Amsterdam/New York: Elsevier Science.

Day, R. R., & Park, J. (2005). Developing reading comprehension questions. *Reading in Foreign Language*, *17*(1), 60–73.

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

Gambier, Y. (2009). Challenges in research on audiovisual translation. In A. Pym & A. Perekrestenko (Eds.), *Translation Research Projects 2* (pp. 17–25). Tarragona, Spain: Intercultural Studies Group.

Gottlieb, H. (1992). Subtitling. A New University Discipline. In C. Dollerup & A. Loddegaard (Eds.), *Teaching Translation and Interpreting: Training Talent and Experience: Papers from the First Language International Conference* (pp. 161–170). Amsterdam, the Netherlands: John Benjamins.

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

Kim, K. J., Sundar, S. S., & Park, E. (2011). The effects of screen-size and communication modality on psychology of mobile device users. In *Proceedings of the 2011 Annual Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 1207–1212). Vancouver, Canada. https://doi.org/10.1145/1979742.1979749

Lavaur, J. M., & Bairstow, D. (2011). Languages on the screen: is film comprehension related to the viewers' fluency level and to the language in the subtitles? *International Journal of Psychology*, *46*(6), 455–462. https://doi.org/10.1080/00207594.2011.565343

Lombard, M., Ditton, T. B., Grabe, M. E., & Reich, R. D. (1997). The role of screen size in viewer responses to television fare. *Communication Reports*, *10*(1), 95–106. https://doi.org/10.1080/08934219709367663

Maniar, N., Bennett, E., Hand, S., & Allan, G. (2008). The effect of mobile phone screen size on video based learning. *Journal of Software*, *3*(4), 51–61. https://doi.org/10.4304/jsw.3.4.51-61

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

Messerlin, P., Siwek, S., & Cocq, E. (2005). *The Audiovisual Services Sector in the GATS Negotiations*. Washington DC: American Enterprise Institute Press.

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

Palen, L., Salzman, M., & Youngs, E. (2000). Going wireless: behavior & practice of new mobile phone users. In *Proceedings of CSCW '00*. Philadelphia, PA, USA.

Perego, E. (2005). *La traduzione audiovisiva*. Roma, Italy: Carocci editore.

Perego, E. (2008). 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

Romero-Fresco, P. (2015). *The reception of subtitles for the deaf and hard of hearing in Europe*. Bern: Peter Lang AG.

Szarkowska, A., Laskowska, M., Oliver, A., & Pilipczuk, O. (2015). Reading subtitles across devices: A study into the differences in reading patterns of people watching subtitled videos on smartphone, tablet and computer screen. In *The International Conference on Eyetracking and Applied Linguistics*. Warsaw.

Chapter 6. Summary

5. Summary

5.1. Summary in English

Subtitle segmentation, i.e. the way text is divided in a two-line subtitle, is believed to be one of the features that influences the readability of subtitles. For over two decades, experts in subtitling claimed that subtitle lines should be split according to syntactic rules to facilitate the reading process. However, the subtitling industry does not always implement these syntactic rules when creating subtitles. Two reasons could explain why these rules are not always applied: human time and effort to edit subtitles, as well as considerable text reduction to keep units of meaning together in the same line. Previous empirical research on this topic has not provided conclusive evidence as to whether syntactic segmentation has a direct impact on the subtitle reading process.

This PhD thesis aims to shed more light on the impact of subtitle segmentation by conducting further research with elements that had not previously be included: a wider range of user profiles, devices with different screen size and more measures in the experimental design. Three empirical studies were carried out to determine whether subtitle segmentation is a key element in Media Accessibility. The first two studies examined the relevance of following syntactic segmentation among viewers with different native languages and hearing statuses, measuring cognitive load, comprehension scores, eye-tracking variables and preferences in line breaks. The third study assessed the reception of subtitles across devices with different screen size, analysing viewers' subtitle layout (specifically focusing on line-break styles) preferences and comprehension.

Overall, the results of these studies seem to indicate that subtitle segmentation is not a critical factor in Media Accessibility. Although non-syntactically segmented subtitles generally induce higher cognitive load and more eye movements, they do not negatively affect comprehension. Viewers are able to adapt their reading strategies regardless of the subtitle segmentation approach or the screen size. Eye tracking results demonstrate that linguistic units are processed differently depending on the way they are split on the screen, their linguistic category

and the viewers' profile. The results of this PhD thesis discuss the effects of segmentation on subtitle processing and the viewer experience in the context of today's changing audiovisual landscape. It is hoped that this thesis provides support for the need to base guidelines and current subtitling practices on empirical research evidence to enhance the quality of Media Accessibility.

5.2. Summary in Catalan

Es considera que la segmentació de subtítols, és a dir, la forma en què es divideix el text en un subtítol de dues línies, constitueix un dels paràmetres que influeix en la llegibilitat dels subtítols. Durant més de dues dècades, els experts en subtitulat han sostingut que les línies dels subtítols han de dividir-se seguint les regles sintàctiques per facilitar la lectura del text. No obstant això, la indústria audiovisual no té en compte sempre aquestes regles a l'hora de crear els subtítols. Existeixen dues raons que podrien justificar per què no sempre es posen en pràctica aquestes regles: el temps i l'esforç que requereixen els subtituladors per editar els subtítols, així com la necessitat de condensar el text per conservar les unitats de sentit en la mateixa línia. Les investigacions empíriques que s'han dut a terme fins ara no han aportat proves concloents sobre l'impacte directe que té la segmentació sintàctica en la lectura dels subtítols.

L'objectiu d'aquesta tesi doctoral és esclarir l'impacte de la segmentació de subtítols duent a terme noves investigacions amb elements que no es van prendre anteriorment en compte: un major nombre de perfils d'usuaris, dispositius amb diferents mides de pantalla i més mesures en el disseny experimental. Es van realitzar tres estudis empírics per determinar si la segmentació de subtítols és un element clau en l'accessibilitat als mitjans audiovisuals. Els dos primers estudis van investigar la importància de respectar les regles sintàctiques en espectadors amb diferents llengües maternes i diferents nivells de pèrdua auditiva. Es van prendre mesures de seguiment ocular, càrrega cognitiva, comprensió i preferències. El tercer estudi va analitzar la recepció de subtítols en dispositius amb diferents mides de pantalla. Es van avaluar els nivells de comprensió i les preferències de presentació de subtítols (centrant l'atenció en estils de segmentació) en cada dispositiu.

En general, els resultats d'aquests estudis semblen indicar que la segmentació de subtítols no és un factor determinant en l'accessibilitat dels mitjans audiovisuals. A pesar que els subtítols que no es divideixen sintàcticament augmenten en general la càrrega cognitiva i els moviments oculars, no afecten de manera negativa a la comprensió. Els espectadors aconsegueixen adaptar les estratègies de lectura independentment de l'enfocament emprat per segmentar els subtítols o de la mida de pantalla. Els resultats dels moviments oculars indiquen que les unitats lingüístiques es llegeixen de manera diferent segons la seva divisió a la pantalla, la categoria lingüística a la qual pertanyen i el perfil de l'espectador. Els resultats d'aquesta tesi doctoral analitzen els efectes de la segmentació en la interpretació dels subtítols i l'experiència de l'espectador en el panorama audiovisual actual en constant canvi. S'espera que aquesta tesi doni suport la necessitat de basar les recomanacions i les pràctiques actuals de subtitulat en investigacions empíriques per millorar la qualitat de l'accessibilitat dels mitjans audiovisuals.

5.3. Summary in Spanish

Se considera que la segmentación de subtítulos, es decir, la forma en la que se divide el texto en un subtítulo de dos líneas, constituye uno de los parámetros que influye en la legibilidad de los subtítulos. Durante más de dos décadas, los expertos en subtitulado han sostenido que las líneas de los subtítulos deben dividirse siguiendo las reglas sintácticas para facilitar la lectura del texto. Sin embargo, la industria audiovisual no tiene en cuenta siempre estas reglas a la hora de crear los subtítulos. Existen dos razones que podrían justificar por qué no siempre se ponen en práctica dichas reglas: el tiempo y el esfuerzo que requieren los subtituladores para editar los subtítulos, así como la necesidad de condensar el texto para conservar las unidades de sentido en la misma línea. Las investigaciones empíricas que se han llevado a cabo hasta ahora no han aportado pruebas concluyentes sobre el impacto directo que tiene la segmentación sintáctica en la lectura de los subtítulos.

El objetivo de esta tesis doctoral es esclarecer el impacto de la segmentación de subtítulos llevando a cabo nuevas investigaciones con elementos que no se tomaron anteriormente en cuenta: un mayor número de perfiles de usuarios, dispositivos con diferentes tamaños de pantalla y más medidas en el diseño experimental. Se realizaron tres estudios empíricos para determinar si la segmentación de subtítulos es un elemento clave en la accesibilidad a los medios audiovisuales. Los dos primeros estudios investigaron la importancia de respetar las reglas sintácticas en espectadores con distintas lenguas maternas y diferentes niveles de pérdida auditiva. Se tomaron medidas de seguimiento ocular, carga cognitiva, comprensión y preferencias. El tercer estudio analizó la recepción de subtítulos en dispositivos con diferentes tamaños de pantalla. Se evaluaron los niveles de comprensión y las preferencias de presentación de subtítulos (centrando la atención en estilos de segmentación) en cada dispositivo.

En general, los resultados de estos estudios parecen indicar que la segmentación de subtítulos no es un factor determinante en la accesibilidad de los medios audiovisuales. A pesar de que los subtítulos que no se dividen sintácticamente aumentan por lo general la carga cognitiva y los movimientos oculares, no afectan de manera negativa a la comprensión. Los espectadores consiguen adaptar las estrategias de lectura independientemente del enfoque empleado para segmentar los subtítulos o del tamaño de pantalla. Los resultados de los movimientos oculares indican que las unidades lingüísticas se leen de manera diferente según su división en la pantalla, la categoría lingüística a la que pertenecen y el perfil del espectador. Los resultados de esta tesis doctoral analizan los efectos de la segmentación en la interpretación de los subtítulos y la experiencia del espectador en el panorama audiovisual actual en constante cambio. Se espera que esta tesis respalde la necesidad de basar las recomendaciones y las prácticas actuales de subtitulado en investigaciones empíricas para mejorar la calidad de la accesibilidad de los medios audiovisuales.

Chapter 7. Conclusions

6. Conclusions

This PhD thesis aimed to investigate to what extent subtitle segmentation, i.e. the way text is divided in a two-line subtitle, is a key element to make media services more accessible to end users. Different studies were carried out to address this question.

Firstly, I undertook two eye-tracking studies to examine the relevance of following syntactic rules to break lines in subtitles (refer to Studies 1 and 2: Gerber-Morón & Szarkowska, 2018; Gerber-Morón et al., 2018). These studies intended to continue the discussion initiated by the previous studies on subtitle segmentation (Perego et al., 2010; Rajendran et al., 2013). They involved more elements in the design of their experiments to obtain updated and conclusive results on the topic. Different groups of viewers were included in these studies: hearing people with different mother tongues (English, Polish, and Spanish) and deaf, hard of hearing, and hearing people with English as a first language. **Study 1** (Gerber-Morón et al., 2018) investigated the impact of text segmentation on subtitle processing, measuring three indicators of cognitive load (difficulty, effort, and frustration), as well as comprehension and eye-tracking variables. **Study 2** (Gerber-Morón & Szarkowska, 2018) examined viewers' preferences regarding line breaks in various linguistic units, using 30 pairs of screenshots with syntactically segmented and non-syntactically segmented subtitles. It also investigated whether these preferences are affected by hearing status and previous experience with subtitling.

Secondly, I considered the current trend of media consumption through handheld devices (Messerlin et al., 2005; Palen et al., 2000) by studying the effects of screen size on subtitle layout and line-break styles. I carried out a third study to examine this element. **Study 3** (Gerber-Morón et al., forthcoming) assessed the reception of subtitles across screens, and aimed to determine viewers' needs and preferences on readability. Thirty native Spanish speakers watched three subtitled videos displayed on three devices with different screen sizes (monitor, tablet, and smartphone). After viewing each subtitled film excerpt, they filled out a reading and subtitle layout preference questionnaire (e.g. number of lines and line length), as well as a comprehension questionnaire. The goal was to detect the device that required more improvement, in order to conduct a subtitle segmentation study on that device.

The data obtained from these studies is discussed in the next section, validating or refuting the hypotheses formulated for each of the research objectives put forward for this PhD thesis. It continues with a section on the contributions of this PhD thesis to subtitle segmentation research. It finally points out the limitations of the studies conducted and the new avenues for future research.

6.1. Discussion of the studies on subtitle segmentation

6.1.1. Results on the impact of syntactic segmentation (Objective 1)

As explained in the Introduction section, experts in Audiovisual Translation have insisted for years on applying syntactic rules to divide text in a two-line subtitle, in order to facilitate the reading process (Baker et al., 1984; Díaz Cintas & Remael, 2007; Ivarsson & Carroll, 1998; Karamitroglou, 2000; Perego, 2008a, 2008b). Among them, Baker et al. (1984) and Karamitroglou (1998) have also recommended to pay attention to aesthetics by creating subtitle lines of equal length. With the exception of Perego et al. (2010) and (Rajendran et al., 2013), none of these experts have proved the relevance of syntactic segmentation in subtitling with empirical evidence.

The first objective of this PhD thesis was to study the impact of syntactic segmentation among various profiles of end users with different audiovisual backgrounds (Spanish, Polish and English native speakers) and hearing status (deaf, hard-of-hearing and hearing people). Based on the literature review on subtitle segmentation (Baker et al., 1984; Díaz Cintas & Remael, 2007; Ivarsson & Carroll, 1998; Karamitroglou, 2000; Perego, 2008a, 2008b) and reading processing (Rayner, 1998; Rayner et al., 2012; Warren, 2012), I hypothesised that syntactic rules would be important to make subtitles more readable, and consequently, more accessible to viewers.

The eye-tracking results from **Study 1** (Gerber-Morón et al., 2018) showed that there is a higher cognitive load when subtitles are not divided according to syntactic rules: mental effort, difficulty, and frustration were reported as higher in non-syntactically segmented subtitles. However, the mean differences did not differ substantially between syntactically and non-syntactically segmented subtitles, and, as a result, the effect sizes were small. A possible explanation for this may be that the clips used in this study were quite short (approximately one minute). As cognitive fatigue increases with longer tasks (Ackerman & Kanfer, 2009; Sandry et al., 2014; Van Dongen et al., 2011), it is possible that mean differences between syntactically and non-syntactically segmented subtitles would be greater in longer clips.

The eye-tracking results in **Study 2** (Gerber-Morón & Szarkowska, 2018) found that noun phrases (i.e. indefinite article + noun, definite article + noun, noun + noun, possessive article + noun) require more reading time when they are syntactically divided into a two-line subtitle. As pointed out by Karamitroglou (1998), syntactically segmented subtitles place complete phrases at the end of lines:

He's looking for the memory stick
he managed to hide.
(Gerber-Morón & Szarkowska, 2018, p. 17)

Following reading studies by Just & Carpenter (1980) and Rayner, Kambe, & Duffy, (2000), readers integrate the information from a clause at the end of the line, a process known as “clause wrap-up effect”. Viewers spend more time looking at noun phrases in syntactically segmented subtitles (than in non-syntactically segmented subtitles) because they are placed at the end of the first subtitle line. On the contrary, viewers spend less time looking at noun phrases in non-syntactically segmented subtitles because they are placed between the first and the second lines of the subtitle:

He's looking for the memory
stick he managed to hide.
(Gerber-Morón & Szarkowska, 2018, p. 17)

This finding does not confirm the results from the previous subtitle segmentation study by Perego et al. (2010), in which no fixation differences were found between ‘ill-segmented’ and ‘well-segmented’ noun phrases. As opposed to the study by Perego et al. (2010) with dynamically changing subtitles, **Study 2** was conducted with a static text of subtitles, which could explain why the results from these two studies were different. Although static text implies a limitation, this study provides a thorough analysis on the reading processing of different linguistic units in subtitles.

In terms of comprehension, **Study 1** did not find evidence that non-syntactically segmented subtitles result in lower comprehension. Participants coped well in both conditions, achieving similar comprehension scores regardless of segmentation. This finding is in line with the results reported by Perego et al. (2010), in which subtitles containing non-syntactically segmented noun phrases did not negatively affect participants’ comprehension.

The question of whether syntactic segmentation facilitates the reading processing of subtitles remains partially unanswered. The approach used to divide a two-line subtitle — whether it is syntactic or non-syntactic — does not facilitate or hinder comprehension. This finding in **Study 1** confirms previous results by Perego et al. (2010) and Rajendran et al. (2013). However, results regarding higher cognitive load in non-syntactically segmented subtitles are still inconclusive. **Study 1** found that non-syntactically segmented subtitles induce a higher cognitive load, but **Study 2** showed that syntactic segmentation in noun phrases takes a longer time to read. This study should be replicated using dynamic subtitled videos and eye-tracking measures to verify whether viewers fixate on longer noun phrases when they are segmented according to syntactic rules.

Following the suggestions by the previous studies on subtitle segmentation (Perego et al., 2010; Rajendran et al., 2013), which only tested undergraduate and postgraduate students, **Studies 1 and 2** of this PhD thesis included more profiles of end users: native speakers with different native languages, and people with different hearing statuses. I hypothesised that syntactic segmentation would be more beneficial for viewers that have more difficulties processing subtitled films, such as deaf people (Cambra et al., 2009; Monreal & Hernandez, 2005; Szarkowska et al., 2011) or those who have been raised in countries favouring dubbing (Perego, Orrego-Carmona, et al., 2016).

In **Studies 1 and 2**, no significant differences were found between processing non-syntactically segmented or syntactically segmented subtitles among the viewer profiles tested. This finding indicates that, whether subtitles are segmented according to syntactic rules or not, different types of viewers process them similarly. Regardless of the subtitle segmentation approach taken, results in both of these studies showed that subtitle processing was least effortful for Polish participants and most effortful for Spanish participants. This finding may be explained by the participants' audiovisual backgrounds: Spanish people have been exposed more to dubbing than subtitling during their lives, as opposed to Polish people. Among hearing, hard-of-hearing and deaf participants, hard-of-hearing participants showed the lowest cognitive effort and the highest comprehension level. This result could be explained by their high familiarity with subtitling (as stated in the pre-test questionnaire) compared to the hearing group.

Study 2 also investigated whether some viewer profiles have a stronger preference for syntactically segmented subtitles than others. Based on previous subtitling and reading studies (Channon & Sayers, 2007; Krejtz et al., 2016; Perego, Orrego-Carmona, et al., 2016; Szarkowska et al., 2011; Wolbers et al., 2012), I expected to find a stronger preference for syntactic segmentation among participants unaccustomed to subtitling and deaf people, who have more difficulties processing written text and subtitling. The results from this study showed that all profiles of viewers expressed a very clear preference for syntactically segmented text in subtitles. In post-test interviews, participants declared that they based their decisions more on syntactic considerations rather than on subtitle shape. This finding is in line with the assumptions made by Ivarsson & Carroll (1998) and Karamitroglou (1998). Syntactically segmented text was preferred over non-syntactically segmented text in nearly all linguistic units by all types of viewers, except for the deaf in the case of the definite article. As sign languages lack many function words, including articles, deaf people tend to have difficulties processing them (Channon & Sayers, 2007; Krejtz et al., 2016; Wolbers et al., 2012). In order to confirm the validity of **Study 2** on viewers' preferences, another experiment should be conducted on a dynamic subtitled film testing the same linguistic units.

6.1.2. Results on the effects of screen size on subtitle layout preferences and comprehension (Objective 2)

Considering the current use of handheld devices to watch audiovisual content (Messerlin et al., 2005; Palen et al., 2000), the second objective of this PhD thesis was to examine the effects of screen size on subtitling, specifically on subtitle layout preferences and comprehension. Previous studies on screen size in Media Psychology and Human-Computer Interaction (Al-Showarah et al., 2014; Kim et al., 2011; Lombard et al., 1997; Maniar et al., 2008), and in Audiovisual Translation (Castellà et al., 2016; Szarkowska et al., 2015) seemed to suggest that larger screens provide a more satisfactory experience.

The results from **Study 3** showed differences in the viewer experience of watching subtitled videos across devices. A general tendency was found towards preferring tablets: participants declared being most satisfied with the length of subtitles and subtitle layout on this type of device. This is probably due to the fact that the tablet display provides a good balance between each subtitle line and its medium size, not forcing the eyes to move much, compared with larger screens (i.e. monitor). This preference was not due to participants being able to adjust reading distance, as it was controlled. As for the other devices, a minor tendency was seen towards preferring shorter subtitles of two lines for monitor screens, and longer subtitles of one line for smartphones. Participants also found the exposure time more appropriate for smartphone screens than tablets or monitors. However, there was a higher variation in the responses for smartphone screens with regard to the question about the feeling of having lost essential parts of the film's action due to the fact of reading subtitles. One possible explanation is that viewers do not perceive these screens as being as optimal as other screens, and they do not feel as confident reading subtitles on them as on larger screen devices. Taking into account these results, and the most unsatisfactory effects of smartphone devices in Szarkowska et al. (2015) and Castellà et al. (2016), more research should be undertaken on these small devices.

Comprehension scores in **Study 3** were not significantly different on any of the devices tested. This finding is in line with the comprehension results found in the previous studies on subtitle segmentation (Perego et al., 2010; Rajendran et al., 2013), on subtitles across screens (Szarkowska et al., 2015), as well as in **Study 1** of this PhD thesis (Gerber-Morón et al., 2018). Comprehension results in **Study 3** confirm the *subtitle effectiveness hypothesis* (Perego et al., 2010) by suggesting that viewers are able to adapt their reading and visual skills for the screen display they are exposed to.

6.2. Contributions of this PhD thesis to subtitle segmentation research

The different subtitle segmentation approaches did not negatively affect comprehension scores in any of the studies conducted for this PhD thesis, even on small screen devices and regardless of the viewer's profile. Viewers were able to easily adapt their reading strategies on small screens and were generally satisfied with how subtitles were displayed on a device of any screen size. This is possibly due to the fact that viewers are accustomed to a wide variety of subtitles, many of which do not adhere to professional segmentation standards. It can also be attributed to what was noted by Mitchell (1989): when interpreting the syntactic structure of sentences in reading, people use non-lexical cues as parsing aids (e.g. text layout or punctuation), but these cues are of secondary importance when compared to words. This is consistent with what participants in **Studies 1 and 2** reported in the interviews. For example, one deaf participant said: “Line breaks have their value, yet when you are reading fast subtitles, most of the time it becomes less relevant.” Nevertheless, the difficulty in finding a reliable measure to evaluate subtitle comprehension could also explain why no significant results were found. Multiple-choice questionnaires — the most commonly used measure — might pose problems to test whether certain manipulations in subtitling, such as segmentation, affect comprehension.

Regarding the viewers' perception of subtitle segmentation, results are not conclusive. While the majority of participants in **Study 2** showed a clear preference for syntactically-segmented subtitles, the results on preferences in the fourth article (under review at the time of submission, please refer to Article 4 in Annex 1: Gerber-

Morón, submitted) showed that participants do not seem to have a clear preference for syntactically or geometrically segmented subtitles on smartphone screens. The difference in results between **Study 2** and the fourth article under review is probably linked to the design of each experiment. **Study 2** used a long series of static text with no sound to evaluate segmentation preferences, whereas the preference questionnaire in the fourth article (see Annex X) was short and based on more real-life subtitling (two videos of 10-minute duration each). In conclusion, viewers generally do not pay attention to syntax or shape in subtitles explicitly when watching media content.

In view of all the results previously described, subtitle segmentation does not seem to be a decisive factor in Media Accessibility. As described by Perego et al. (2010) in their *subtitle effectiveness hypothesis*, the approaches to subtitle segmentation — whether it is syntactic or geometric — do not substantially affect the viewers' experience, as they represent minor manipulations in subtitling. Although more studies should be carried out to analyse subtitle segmentation in a full-length film using a small screen, or to test segmentation across subtitles (i.e. intersegmentation, for more details see Article 4 in Annex 1: Gerber-Morón, submitted), all the studies conducted up until now indicate that further research should focus on other types of subtitles, such as dynamic subtitles[8] (A. Brown et al., 2015; Fox, 2016) and live subtitling (Romero-Fresco & Pérez, 2015; Romero-Fresco & Pöchhacker, 2017). On one hand, dynamic subtitles involve major variations in the layout, and end users might find it more difficult to process information, as they are more used to finding the subtitled text placed at the bottom-centre of the screen. On the other hand, nowadays many countries use respeaking to produce live subtitles[9]. This technique has the advantage of offering a shorter training period for live subtitlers, consequently increasing the production of live subtitling (Lambourne, Hewitt, Lyon, & Warren, 2004). However, it contains speech recognition errors, shows subtitles with a delay and poses challenges to the respeakers to reduce the original dialogues (Waes, Leijten, & Remael, 2013). As technological developments

8 This type of subtitles are displayed in varying areas of the screen depending on the specific visuals of the video subtitled, to avoid covering important information.

9 Speech recognition software is used by a live subtitler (i.e. a respeaker) to create and broadcast live subtitles.

are constantly improving, more research should be carried out with this type of subtitles to improve their quality.

6.3. Limitations of the studies conducted for this PhD thesis

One significant limitation was the duration of the clips in **Studies 1 and 3**. Longer clips with non-syntactically segmented subtitles in **Study 1** may have provided more significant mean differences between the syntactically and non-syntactically segmented subtitles, as cognitive load accumulates over time. It could have also provided more differences in the subtitle layout preferences across devices in **Study 3**. However, the clips selected for Article 4 in Annex 1 (under review at the time of submission: Gerber-Morón, submitted) were relatively longer and the results corroborated the findings in **Studies 1 and 3**.

Another limitation is related to the use of static text with no sound to evaluate segmentation preferences in **Study 2**. This approach allowed for the controlling of linguistic units and for participants to be presented with two clear conditions to compare. Nevertheless, participants could take as much time as they needed to complete the task, whereas in real-life subtitling viewers have no control over the presentation speed, having less time to process subtitles. The understanding of subtitled text is also context-sensitive, and this study did not allow participants to rely on the context to interpret the sentences, as the study only contained screenshots.

The difficulty in finding a reliable measure to evaluate subtitle comprehension is another limitation to take into consideration. Multiple-choice questionnaires might pose problems to test whether certain manipulations in subtitling, such as segmentation, affect comprehension. This could explain why no significant differences were found in comprehension results for any of the studies conducted in this PhD thesis.

6.4. Future paths for researchers in Media Accessibility

We live in a society where inclusion is an increasing concern and researchers in Media Accessibility have the responsibility to adapt studies to the end users' needs. New themes need to be examined, experimental methodology has to be further improved, and more innovative solutions should be offered. Here are three considerations for future research.

Different approaches need to be adopted to experiment and interpret scientific results. On the one hand, replicating existing studies is necessary to validate findings. On the other hand, not obtaining significant results does not imply that studies do not provide novelty and consequently are not publishable. They also provide information about the impact of the elements tested on viewers and point out possible lines for future research on that topic. This was the case with my thesis topic: the different studies on subtitle segmentation (Gerber-Morón, submitted; Gerber-Morón et al., forthcoming; Gerber-Morón & Szarkowska, 2018; Gerber-Morón et al., 2018; Perego et al., 2010; Rajendran et al., 2013) showed that comprehension is not affected by the way subtitle lines are divided on screen. End users are able to adapt their way of processing audiovisual information. However, the results of these studies suggest new directions for further research. The effects of subtitle segmentation on viewers' cognitive load across two or more consecutive subtitles could be explored further. Another possible line of future research could test subtitle segmentation in languages that use very different syntactic structures than the ones tested in this PhD thesis (e.g. Asian languages).

Designing experimental studies to assess subtitle accessibility is a complex task, and results are not always entirely accurate: each end user processes information in a different way, making it difficult to find groups of participants with identical features. This has been evidenced in the experiments conducted for this PhD thesis, in particular for the comprehension analysis. The question is whether the methodology used to measure comprehension is effective or whether another model should be implemented. The same applies to other variables, such as preference or cognitive load questionnaires. Focusing on different technologies, such as eye-tracking and skin-conductance measures, could provide significant results on the viewer experience.

In the field of Media Accessibility, researchers need to look for solutions that can be directly applied to media services. Applied research offers this possibility: it explores applications that can have a fast and direct impact on the viewing experience. For instance, research could go a step further in exploring the usability of glasses to watch three-dimensional and 360-degree virtual reality subtitled videos. Another possibility would be to investigate the usability of wearing special glasses to read subtitles for stage shows, live concerts and in cinema theatres. For example, the "Invisible Subtitle"[10] technology, where subtitles become visible on the screen only after putting on special glasses, could be tested.

Media Accessibility is a great instrument that enables society to come together, communicate and feel inclusive. We must continue to commit and work towards this end.

10 http://giojax.com

Chapter 8. Bibliography

7. Bibliography

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

Ackerman, P. L., & Kanfer, R. (2009). Test Length and Cognitive Fatigue: An Empirical Examination of Effects on Performance and Test-Taker Reactions. *Journal of Experimental Psychology: Applied*, *15*(2), 163–181. https://doi.org/10.1037/a0015719

Al-Showarah, S., AL-Jawad, N., & Sellahewa, H. (2014). Effects of User Age on Smartphone and Tablet Use, Measured with an Eye-Tracker via Fixation Duration, Scan-Path Duration, and Saccades Proportion. In C. Stephanidis & M. Antona (Eds.), *UAHCI/HCII, Part II, LNCS 8514* (pp. 3–14). Springer International Publishing Switzerland.

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

Allen, P., Garman, J., Calvert, I., & Murison, J. (2011). Reading in multimodal environments: assessing legibility and accessibility of typography for television. In *The proceedings of the 13th international ACM SIGACCESS conference on computers and accessibility* (pp. 275–276). https://doi.org/10.1145/2049536.2049604

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

Antonenko, P., Paas, F., Grabner, R., & van Gog, T. (2010). Using Electroencephalography to Measure Cognitive Load. *Educational Psychology Review*, *22*(4), 425–438. https://doi.org/10.1007/s10648-010-9130-y

Baddeley, A. (2007). *Working memory, thought, and action*. Oxford: Oxford University Press.

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

Bazeley, P. (2013). *Qualitative Data*. Los Angeles: Sage.

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, A., Jones, R., Crabb, M., Sandford, J., Brooks, M., Armstrong, M., & Jay, C. (2015). Dynamic Subtitles. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video - TVX '15* (pp. 103–112). Brussels, Belgium: ACM. https://doi.org/10.1145/2745197.2745204

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. *Journal of Deaf Studies and Deaf Education*, *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

Castellà, J., Olivier, A., Gerber-Morón, O., & Soler-Vilageliu, O. (2016). Smartphone, Tablet or PC: Which one is better to watch your favourite subtitled film? In *International Meeting of the Psychonomic Society*. Granada (Spain).

Chandler, P., & Sweller, J. (1991). Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction*, *8*(4), 293–332. https://doi.org/10.1207/s1532690xci0804_2

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

d'Ydewalle, G., Praet, C., Verfaillie, K., & Van Rensbergen, J. (1991). Watching Subtitled Television: Automatic Reading Behavior. *Communication Research*, *18*(5), 650–666.

D'Ydewalle, G., Rensbergen, J. Van, & Pollet, J. (1987). Reading a message when the same message is available auditorily in another language: the case of subtitling. In J. K. O'Regan & A. Levy-Schoen (Eds.), *Eye movements: from physiology to cognition* (pp. 313–321). Amsterdam/New York: Elsevier Science.

Day, R. R., & Park, J. (2005). Developing reading comprehension questions. *Reading in Foreign Language*, *17*(1), 60–73. https://doi.org/10.1080/01443410.2014.910979

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. (2018a). The development of eye tracking in empirical research on subtitling and captioning. In T. Dwyer, C. Perkins, S. Redmond, & J. Sita (Eds.), *Seeing into Screens: Eye Tracking and the Moving Image* (pp. 46–64). London: Bloomsbury.

Doherty, S., & Kruger, J.-L. (2018b). 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.

Fox, W. (2016). Integrated titles: An improved viewing experience? In S. Hansen-Schirra & S. Grucza (Eds.), *Eyetracking and Applied Linguistics* (pp. 5–29). Berlin: Language Science Press.

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

Gambier, Y. (2009). Challenges in research on audiovisual translation. In A. Pym & A. Perekrestenko (Eds.), *Translation Research Projects 2* (pp. 17–25). Tarragona, Spain: Intercultural Studies Group.

Gerber-Morón, O. (submitted). *How to divide lines in a subtitle: an exploratory study on subtitle segmentation approaches in alternative small screens*.

Gerber-Morón, O., Soler-Vilageliu, O., & Castellà, J. (forthcoming). The effects of screen size on subtitle layout preferences and comprehension across devices. *Hermēneus. Revista de Traducción e Interpretación*, *21*.

Gerber-Morón, O., & Szarkowska, A. (2018). Line breaks in subtitling: an eye tracking study on viewer preferences. *Journal of Eye Movement Research*, *11*(3), 1–22. https://doi.org/10.16910/jemr.11.3.2

Gerber-Morón, O., Szarkowska, A., & Woll, B. (2018). The impact of text segmentation on subtitle reading. *Journal of Eye Movement Research*, *11*(4), 1–18. https://doi.org/10.16910/11.4.2

Gernsbacher, M. A. (2015). Video Captions Benefit Everyone. *Policy Insights from the Behavioral and Brain Sciences*, *2*(1), 195–202. https://doi.org/10.1177/2372732215602130

Gottlieb, H. (1992). Subtitling. A New University Discipline. In C. Dollerup & A. Loddegaard (Eds.), *Teaching Translation and Interpreting : Training Talent and Experience : Papers from the First Language International Conference* (pp. 161–170). Amsterdam, the Netherlands: John Benjamins.

Greco, G. M. (2016). On Accessibility as a Human Right, with an Application to Media Accessibility. In A. Matamala & P. Orero (Eds.), *Researching Audio Description: New Approaches* (pp. 11–33). London: Palgrave Macmillan UK. https://doi.org/10.1057/978-1-137-56917-2_2

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam: North-Holland.

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.

Jensema, C. (1998). Viewer Reaction to Different Television Captioning Speeds. *American Annals of the Deaf*, *143*(4), 318–324.

Jensema, C., El Sharkawy, S., Danturthi, R. S., Burch, R., & Hsu, D. (2000). Eye Movement Patterns of Captioned Television Viewers. *American Annals of the Deaf*, *145*(3), 275–285.

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

Karamitroglou, F. (2000). *Towards a methodology for the investigation of norms in audiovisual translation : the choice between subtitling and revoicing in Greece / Fotios Karamitroglou*. *Approaches to translation studies*. Amsterdam; Atlanta, GA: Rodopi.

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, US: Oxford University Press.

Keenan, S. A. (1984). Effects of Chunking and Line Length on Reading Efficiency. *Visible Language*, *18*(1), 61–80.

Kennedy, A., Murray, W. S., Jennings, F., & Reid, C. (1989). Parsing Complements: Comments on the Generality of the Principle of Minimal Attachment. *Language and Cognitive Processes*, *4*(3–4), SI51-SI76. https://doi.org/10.1080/01690968908406363

Kim, K. J., Sundar, S. S., & Park, E. (2011). The effects of screen-size and communication modality on psychology of mobile device users. In *Proceedings of the 2011 Annual Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 1207–1212). Vancouver, Canada. https://doi.org/10.1145/1979742.1979749

Koolstra, C. M., Van Der Voort, T. H. A., & D’Ydewalle, G. (1999). Lengthening the Presentation Time of Subtitles on Television: Effects on Children’s Reading Time and Recognition. *Communications*, *24*(4), 407–422. https://doi.org/10.1515/comm.1999.24.4.407

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., & Doherty, S. (2016). Measuring Cognitive Load in the Presence of Educational Video: Towards a Multimodal Methodology. *Australasian Journal of Educational Technology*, *32*(6), 19–31. https://doi.org/10.14742/ajet.3084

Kruger, J.-L., Doherty, S., Fox, W., & de Lissa, P. (2017). Multimodal measurement of cognitive load during subtitle processing: Same-language subtitles for foreign language viewers. In I. Lacruz & R. Jääskeläinen (Eds.), *New Directions in Cognitive and Empirical Translation Process Research* (pp. 267–294). London: John Benjamins.

Kruger, J.-L., Hefer, E., & Matthew, G. (2013). Measuring the impact of subtitles on cognitive load: eye tracking and dynamic audiovisual texts. *Proceedings of the 2013 Conference on Eye Tracking South Africa*. Cape Town, South Africa: ACM. https://doi.org/10.1145/2509315.2509331

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. https://doi.org/http://dx.doi.org/10.16910/jemr.7.5.4

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*.

Lambourne, A., Hewitt, J., Lyon, C., & Warren, S. (2004). Speech-based Real Time Subtitling Services. *International Journal of Speech Technology*, *7*, 269–279.

Lavaur, J. M., & Bairstow, D. (2011). Languages on the screen: is film comprehension related to the viewers' fluency level and to the language in the subtitles? *International Journal of Psychology*, *46*(6), 455–462. https://doi.org/10.1080/00207594.2011.565343

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

LeVasseur, V. M., Macaruso, P., Palumbo, L. C., & Shankweiler, D. (2006). Syntactically cued text facilitates oral reading fluency in developing readers. *Applied Psycholinguistics*, *27*(3), 423–445. https://doi.org/10.1017/s0142716406060346

Lombard, M., Ditton, T. B., Grabe, M. E., & Reich, R. D. (1997). The role of screen size in viewer responses to television fare. *Communication Reports*, *10*(1), 95–106. https://doi.org/10.1080/08934219709367663

Łuczak, K. (2017). *The effects of the language of the soundtrack on film comprehension, cognitive load and subtitle reading patterns. An eye-tracking study.* University of Warsaw, Warsaw.

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, UK: The European Institute for the Media.

Maniar, N., Bennett, E., Hand, S., & Allan, G. (2008). The effect of mobile phone screen size on video based learning. *Journal of Software*, *3*(4), 51–61. https://doi.org/10.4304/jsw.3.4.51-61

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967.

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*, *63*(3), 423–441. https://doi.org/10.1075/babel.63.3.07mat

Mayberry, R. I., del Giudice, A. A., & Lieberman, A. M. (2011). Reading Achievement in Relation to Phonological Coding and Awareness in Deaf Readers: A Meta-analysis. *Journal of Deaf Studies and Deaf Education*, *16*(2), 164–188. https://doi.org/10.1093/deafed/enq049

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

Messerlin, P., Siwek, S., & Cocq, E. (2005). *The Audiovisual Services Sector in the GATS Negotiations*. Washington DC: American Enterprise Institute Press.

Miquel Iriarte, M. (2017). *The reception of subtitling for the deaf and hard of hearing: viewers' hearing and communication profile & Subtitling speed of exposure*. Universitat Autònoma de Barcelona, Barcelona.

Mitchell, D. C. (1987). Lexical guidance in human parsing: Locus and processing characteristics. In M. Coltheart (Ed.), *Attention and Performance XII*. London: Lawrence Erlbaum Associates Ltd.

Mitchell, D. C. (1989). Verb guidance and other lexical effects in parsing. *Language and Cognitive Processes*, *4*(3–4), SI123-SI154. https://doi.org/10.1080/01690968908406366

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

Musselman, C. (2000). How Do Children Who Can't Hear Learn to Read an Alphabetic Script? A Review of the Literature on Reading and Deafness. *Journal Of Deaf Studies And Deaf Education*, *5*(1), 9–31.

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. (2015). Ofcom's Code on Television Access Services.

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

Palen, L., Salzman, M., & Youngs, E. (2000). Going wireless: behavior & practice of new mobile phone users. In *Proceedings of CSCW '00*. Philadelphia, PA, USA.

Perego, E. (2005). *La traduzione audiovisiva*. Roma, Italy: Carocci editore.

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

Perego, E., Orrego-Carmona, D., & Bottiroli, S. (2016). *An empirical take on the dubbing vs. subtitling debate. An eye movement study. 2016*. Retrieved from http://siba-ese.unisalento.it/index.php/linguelinguaggi/article/view/16084/14310

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*. San Diego, CA: 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.

Romero-Fresco, P. (2015). *The reception of subtitles for the deaf and hard of hearing in Europe*. Bern: Peter Lang AG.

Romero-Fresco, P., & Pérez, J. M. (2015). Accuracy Rate in Live Subtitling: The NER Model. *Audiovisual Translation in a Global Context*, 28–50. https://doi.org/10.1057/9781137552891_3

Romero-Fresco, P., & Pöchhacker, F. (2017). Quality assessment in interlingual live subtitling: The NTR Model. *Linguistica Antverpiensia New Series-Themes in Translation Studies*, *16*, 149–167.

Sandry, J., Genova, H. M., Dobryakova, E., DeLuca, J., & Wylie, G. (2014). Subjective Cognitive Fatigue in Multiple Sclerosis Depends on Task Length. *Frontiers in Neurology*, *5*, 214. https://doi.org/10.3389/fneur.2014.00214

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

Schmeck, A., Opfermann, M., van Gog, T., Paas, F., & Leutner, D. (2014). Measuring cognitive load with subjective rating scales during problem solving: differences between immediate and delayed ratings. *Instructional Science*, *43*(1), 93–114. https://doi.org/10.1007/s11251-014-9328-3

Sun, S., & Shreve, G. (2014). Measuring translation difficulty: An empirical study. *Target*, *26*(1), 98–127. https://doi.org/10.1075/target.26.1.04sun

Sweller, J. (2011). Cognitive Load Theory. *Psychology of Learning and Motivation*, *55*, 37–76.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). Cognitive Load Theory. *Psychology of Learning and Motivation*, *55*, 37–76.

Sweller, J., Van Merrienboer, J., & Paas, F. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, *10*(3), 251–296. https://doi.org/10.1023/a:1022193728205

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

Szarkowska, A., Krejtz, I., Kłyszejko, 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).

Szarkowska, A., Krejtz, I., Krejtz, K., & Duchowski, A. (2013). Harnessing the Potential of Eyetracking for Media Accessibility. In S. Grucza, M. Płużyczka, & J. Zając (Eds.), *Translation studies and eye-tracking analysis*. Frankfurt am Main: Peter Lang.

Szarkowska, A., Krejtz, I., Pilipczuk, O., Dutka, Ł., & Kruger, J.-L. (2016). The effects of text editing and subtitle presentation rate on the comprehension and reading patterns of interlingual and intralingual subtitles among deaf, hard of hearing and hearing viewers. *Across Languages and Cultures*, *17*(2), 183–204. https://doi.org/10.1556/084.2016.17.2.3

Szarkowska, A., Krejtz, K., Dutka, Ł., & Pilipczuk, O. (2016). Cognitive load in intralingual and interlingual respeaking-a preliminary study. *Poznan Studies in Contemporary Linguistics*, *52*(2). https://doi.org/10.1515/psicl-2016-0008

Szarkowska, A., Laskowska, M., Oliver, A., & Pilipczuk, O. (2015). Reading subtitles across devices: A study into the differences in reading patterns of people watching subtitled videos on smartphone, tablet and computer screen. In *The International Conference on Eyetracking and Applied Linguistics*. Warsaw.

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

Thorn, F., & Thorn, S. (1996). Television Captions for Hearing-Impaired People: A Study of Key Factors that Affect Reading Performance. *Human Factors*, *38*(3), 452–463.

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

Van Dongen, H. P. A., Belenky, G., & Krueger, J. M. (2011). *Investigating the temporal dynamics and underlying mechanisms of cognitive fatigue.* American Psychological Association. https://doi.org/10.1037/12343-006

Van Gerven, P. W. M., Paas, F., Van Merriënboer, J. J. G., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology*, *41*(2), 167–174. https://doi.org/10.1111/j.1469-8986.2003.00148.x van Gog, T., & Paas, F. (2008). Instructional Efficiency: Revisiting the Original Construct in Educational Research. *Educational Psychologist*, *43*(1), 16–26. https://doi.org/10.1080/00461520701756248

Waes, L., Leijten, M., & Remael, A. (2013). Live subtitling with speech recognition. Causes and consequences of text reduction. *Across Languages and Cultures*, *14*(1), 15–46. https://doi.org/10.1556/Acr.14.2013.1.2

Wang, Z., & Duff, B. R. L. (2016). All Loads Are Not Equal: Distinct Influences of Perceptual Load and Cognitive Load on Peripheral Ad Processing. *Media Psychology*, *19*(4), 589–613. https://doi.org/10.1080/15213269.2015.1108204

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

Yoon, J.-O., & Kim, M. (2011). The Effects of Captions on Deaf Students' Content Comprehension, Cognitive Load, and Motivation in Online Learning. *American Annals of the Deaf, 156*(3), 283–289. https://doi.org/10.1353/aad.2011.0026

Chapter 9. Annexes

Annexes

Annex 1: Articles within this PhD thesis

Article 1

Gerber-Morón, O., Szarkowska, A. & Woll, B. (2018). The impact of text segmentation on subtitle reading. *Journal of Eye Movement Research*, *11*(4):1-18. DOI: http://dx.doi.org/10.16910/11.4.2

Journal of Eye Movement Research
10.16910/jemr.11.4.2

The impact of text segmentation on subtitle reading

Olivia Gerber-Morón
Universitat Autònoma de Barcelona, Spain

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

Bencie Woll
University College London, UK

Understanding the way people watch subtitled films has become a central concern for subtitling researchers in recent years. Both subtitling scholars and professionals generally believe that in order to reduce cognitive load and enhance readability, line breaks in two-line subtitles should follow syntactic units. However, previous research has been inconclusive as to whether syntactic-based segmentation facilitates comprehension and reduces cognitive load. In this study, we assessed the impact of text segmentation on subtitle processing among different groups of viewers: hearing people with different mother tongues (English, Polish, and Spanish) and deaf, hard of hearing, and hearing people with English as a first language. We measured three indicators of cognitive load (difficulty, effort, and frustration) as well as comprehension and eye tracking variables. Participants watched two video excerpts with syntactically and non-syntactically segmented subtitles. The aim was to determine whether syntactic-based text segmentation as well as the viewers' linguistic background influence subtitle processing. Our findings show that non-syntactically segmented subtitles induced higher cognitive load, but they did not adversely affect comprehension. The results are discussed in the context of cognitive load, audiovisual translation, and deafness.

Keywords: eye movement, reading, region of interest, subtitling, audiovisual translation, media accessibility, cognitive load, segmentation, line breaks, revisits

Introduction

In the modern world, we are surrounded by screens, captions, and moving images more than ever before. Technological advancements and accessibility legislation, such as the United Nations Convention on the Rights of Persons with Disabilities (2006), Audiovisual Media Services Directive or the European Accessibility Act, have empowered different types of viewers across the globe in accessing multilingual audiovisual content. Viewers who do not know the language of the original production or people who are deaf or hard of hearing can follow film dialogues thanks to subtitles (Gernsbacher, 2015).

Received March 24, 2018; Published June 30, 2018.
Citation: Gerber-Morón, O., Szarkowska, A., & Woll, B. (2018). The impact of text segmentation on subtitle reading. *Journal of Eye Movement Research*, *11*(4):2
Digital Object Identifier: 10.16910/jemr.11.4.2
ISSN: 1995-8692

Because watching subtitled films requires viewers to follow the action, listen to the soundtrack and read the subtitles, it is important for subtitles to be presented in a way that facilitates rather than hampers reading (Díaz Cintas & Remael, 2007; Karamitroglou, 1998). Some typographical subtitle parameters, such as small font size, illegible typeface or optical blur, have been shown to impede reading (Allen, Garman, Calvert, & Murison, 2011; Thorn & Thorn, 1996). In this study, we examine whether segmentation, i.e. the way text is divided across lines in a two-line subtitle, affects the subtitle reading process. We predict that segmentation not aligned with grammatical structure may have a detrimental effect on the processing of subtitles.

Readability and syntactic segmentation in subtitles

The general consensus among scholars in audiovisual translation, media regulation, and television broadcasting is that to enhance readability, linguistic phrases in two-line subtitles should not be split across lines (BBC, 2017; Díaz Cintas & Remael, 2007; Ivarsson & Carroll, 1998;

Journal of Eye Movement Research
10.16910/jemr.11.4.2

Karamitroglou, 1998; Ofcom, 2015). For instance, subtitle (1a) below is an example of correct syntactic-based line segmentation, whereas in (1b) the indefinite article "a" is incorrectly separated from the accompanying noun phrase (BBC, 2017).

(1a)
We are aiming to get
<u>a better television service</u>.

(1b)
We are aiming to get <u>a</u>
<u>better television service</u>.

The underlying assumption is that more cognitive effort is required to process text when it is not segmented according to syntactic rules (Perego, 2008a). However, segmentation rules are not always respected in the subtitling industry. One of the reasons for this might be the cost: editing text in subtitles requires human time and effort, and as such is not always cost-effective. Another reason is that syntactic-based segmentation may require substantial text reduction in order to comply with maximum line length limits. As a result, when applying syntactic rules to segmentation of subtitles, some information might be lost. Following this line of thought, BBC subtitling guidelines (BBC, 2017) stress that well-edited text and synchronisation should be prioritized over syntactically-based line breaks.

The widely held belief that words "intimately connected by logic, semantics, or grammar" should be kept in the same line whenever possible (Ivarsson & Carroll, 1998, p. 77) may be rooted in the concept of parsing in reading (Rayner, Pollatsek, Ashby, & Clifton, 2012, p. 216). Parsing, i.e. the process of identifying which groups of words go together in a sentence (Warren, 2012), allows a text to be interpreted incrementally as it is read. It has been reported that "line breaks, like punctuation, may have quite profound effects on the reader's segmentation strategies" (Kennedy, Murray, Jennings, & Reid, 1989, p. 56). Insight into these strategies can be obtained through studies of readers' eye movements, which reflect the process of parsing: longer fixation durations, higher frequency of regressions, and longer reading time may be indicative of processing difficulties (Rayner, 1998). An inappropriately placed line break may lead a reader to incorrectly interpret the meaning and structure, luring the reader into a parse that turns out to be a dead end or yield a clearly unintended reading – a so-called "garden path" experience (Frazier, 1979; Rayner et al., 2012). The reader must then reject their initial interpretation and re-read the text. This takes extra time and, as such, is unwanted in subtitling, which is supposed to be as unobtrusive as possible and should not interfere with the viewer's enjoyment of the moving images (Díaz Cintas & Remael, 2007).

Despite a substantial body of experimental research on subtitling (Bisson, Van Heuven, Conklin, & Tunney, 2012; d'Ydewalle & De Bruycker, 2007; d'Ydewalle, Praet, Verfaillie, & Van Rensbergen, 1991; Koolstra, Van Der Voort, & d'Ydewalle, 1999; Kruger, Hefer, & Matthew, 2013; Kruger & Steyn, 2014; Perego et al., 2016; Szarkowska, Krejtz, Pilipczuk, Dutka, & Kruger, 2016), the question of whether text segmentation affects subtitle processing (Perego, 2008a) still remains unanswered. Previous research is inconclusive as to whether linguistically segmented text facilitates subtitle processing and comprehension. Contrary to arguments underpinning professional subtitling recommendations, Perego, Del Missier, Porta, & Mosconi (2010), who used eye-tracking to examine subtitle comprehension and processing, found no disruptive effect of "syntactically incoherent" segmentation of noun phrases on the effectiveness of subtitle processing in Italian. In their study, the number of fixations and saccadic crossovers (i.e. gaze jumps between the image and the subtitle) did not differ between the syntactically segmented and non-segmented conditions. In contrast, in a study on live subtitling, Rajendran, Duchowski, Orero, Martínez, & Romero-Fresco (2013) showed benefits of linguistically-based segmentation by phrase, which induced fewer fixations and saccadic crossovers, and resulted in shortest mean fixation duration, together indicating less effortful processing.

Ivarsson & Carroll (1998) noted that "matching line breaks with sense blocks is especially important for viewers with any kind of linguistic disadvantage, e.g. immigrants or young children learning to read or the deaf with their acknowledged reading problems" (p. 78). Indeed, early deafness is strongly associated with reading difficulties (Mayberry, del Giudice, & Lieberman, 2011; Musselman, 2000). Researchers investigating subtitle reading by deaf viewers have demonstrated processing difficulties resulting in lower comprehension and more time spent by deaf viewers on reading subtitles (Krejtz, Szarkowska, & Krejtz, 2013; Krejtz, Szarkowska, & Łogińska, 2016; Szarkowska, Krejtz, Kłyszejko, & Wieczorek, 2011). Lack of familiarity with subtitling is another aspect which may affect the way people read subtitles. In a recent study, Perego et al.

No se puede mostrar la imagen en este momento.

No se puede mostrar la imagen en este momento.

To address the contribution of hearing status and experience with subtitling to cognitive processing, our study includes British viewers with varying hearing status (deaf, hard of hearing, and hearing), and hearing native speakers of different languages: Spanish people, who grew up in a country where the dominant type of audiovisual translation is dubbing, and Polish people, who come from the tradition of voice-over and subtitling. We conducted two experiments: Experiment 1 with hearing people from the UK, Poland, and Spain, and Experiment 2 with English hearing, hard of hearing and deaf people. We predicted that for those who are not used to subtitling, cognitive load would be higher, comprehension would be lower and time spent in the subtitle would be higher, as indicated by absolute reading time, fixation count and proportional reading time.

By using a combination of different research methods, such as eye tracking, self-reports, and questionnaires, we have been able to analyse the impact of text segmentation on the processing of subtitles, modulated by different linguistic backgrounds of viewers. Examining these issues is particularly relevant from the point of view of current subtitling standards and practices.

Methods

The study took place at University College London and was part of a larger project on testing subtitle processing with eye tracking. In this paper, we report the results from two experiments using the same methodology and materials: Experiment 1 with hearing native speakers of English, Polish, and Spanish; and Experiment 2 with hearing, hard of hearing, and deaf British participants. The English-speaking hearing participants are the same in both experiments. In each of the two experiments, we employed a mixed factorial design with segmentation (syntactically segmented vs. non-syntactically segmented) as the main within-subject independent variable, and language (Exp. 1) or hearing loss (Exp. 2) as a between-subject factor.

All the study materials and results are available in an open data repository RepOD hosted by the University of Warsaw (Szarkowska & Gerber-Morón, 2018).

Participants

Participants were recruited from the UCL Psychology pool of volunteers, social media (Facebook page of the project, Twitter), and personal networking. Hard of hearing participants were recruited with the help of the National Association of Deafened People. Deaf participants were also contacted through the UCL Deafness, Cognition, and Language Research Centre participant pool. Participants were required not to know Hungarian.

Table 1. Demographic information on participants

Experiment 1				
		English	Polish	Spanish
Gender				
	Male	13	5	10
	Female	14	16	16
Age				
	Mean	27.59	24.71	28.12
	(SD)	(7.79)	(5.68)	(5.88)
	Range	20-54	19-38	19-42
Experiment 2				
		Hearing	Hard of hearing	Deaf
Gender				
	Male	13	2	4
	Female	14	8	5
Age				
	Mean	27.59	46.40	42.33
	(SD)	(7.79)	(12.9)	(14.18)
	Range	20-54	22-72	24-74

Experiment 1 participants were pre-screened to be native speakers of English, Polish or Spanish, aged above 18. They were all resident in the UK. We tested 27 English, 21 Polish, and 26 Spanish speakers (see Table 1). At the study planning and design stage, Spanish speakers were included on the assumption that they would be unaccustomed to subtitling as they come from Spain, a country in which foreign programming is traditionally presented with dubbing. Polish participants were included as Poland is a country where voice-over and subtitling are commonly used, the former on television and VOD, and the latter in cinemas, DVDs, and VOD. The hearing English participants were used as a control group.

Despite their experiences in their native countries, when asked about the preferred type of audiovisual translation (AVT), most of the Spanish participants declared they preferred subtitling and many of the Polish participants reported that they watch films in the original (see Table 2).

Journal of Eye Movement Research
10.16910/jemr.11.4.2

Table 2. Preferred way of watching foreign films

	English	Polish	Spanish
Subtitling	24	11	22
Dubbing	0	0	1
Voice-over	1	0	0
I watch films in their original version	1	10	3
I never watch foreign films	1	0	0

We also asked the participants how often they watched English and non-English programmes with English subtitles (Fig. 1).

Fig. 1. Participants' subtitle viewing habits

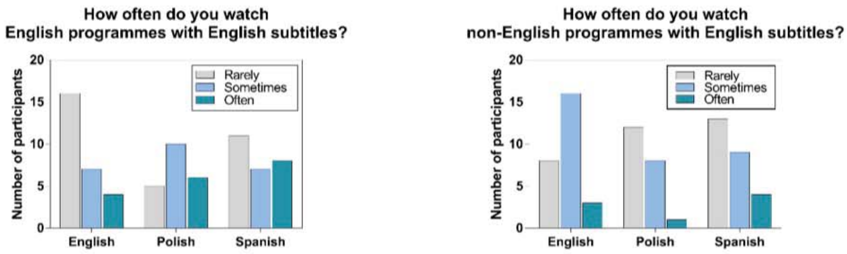


Table 3. Education background of hearing participants in Experiment 1

	English	Polish	Spanish
Secondary education	5	9	6
Bachelor degree	14	4	6
Master degree	8	8	13
PhD	0	0	1

As subtitles used in the experiments were in English, we asked Polish and Spanish speakers to assess their proficiency in reading English using the Common European Framework of Reference for Languages (from A1 to C2), see Table 4. None of the participants declared a reading level lower than B1. The difference between the proficiency in English of Polish and Spanish participants was not statistically significant, $\chi^2(3) = 5.144$, $p = .162$. Before declaring their proficiency, each participant was presented with a sheet describing the skills and competences required at each proficiency level (Szarkowska & Gerber-Morón, 2018). There is evidence that self-report correlates reasonably well with objective assessments (Marian, Blumenfeld, & Kaushanskaya, 2007).

The heterogeneity of participants' habits and preferences reflects the changing AVT landscape in Europe (Matamala, Perego, & Bottiroli, 2017) on the one hand, and on the other, may be attributed to the fact that participants were living in the UK and thus had different experiences of audiovisual translation than in their home countries. The participants' profiles make them not fully representative of the Spanish/Polish population, which we acknowledge here as a limitation of the study.

To determine the level of participants' education, hearing people were asked to state the highest level of education they completed (Table 3, see also Table 5 for hard of hearing and deaf participants). Overall, the sample was relatively well-educated.

Table 4. Self-reported English proficiency in reading of Polish and Spanish participants

	Polish	Spanish
B1	0	1
B2	0	4
C1	3	5
C2	18	16
Total	21	26

In Experiment 2, participants were classified as either hearing, hard of hearing, or deaf. Before taking part in the study, those with hearing impairment completed a questionnaire about the severity of their hearing impairment, age of onset of hearing impairment, communication preferences, etc. and were asked if they described themselves as deaf or hard of hearing. They were also asked to indicate their education background (see Table 5). We recruited 27 hearing, 10 hard of hearing, and 9 deaf participants. Of the deaf and hard of hearing participants, 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. Nine were profoundly deaf, 6 were severely deaf, and 4 had a moderate hearing loss. Seventeen of the deaf and hard of hearing participants preferred to use spoken English as their means of communication in the study and two chose to use a British Sign Language interpreter. In relation to AVT, 84.2% stated that they often watch films in English with English subtitles; 78.9% declared they could not follow a film without subtitles; 58% stated that they always or very often watch non-English films with English subtitles. Overall, deaf and hard of hearing participants in our study were experienced subtitle users, who rely on subtitles to follow audiovisual materials.

Table 5. Education background of deaf and hard of hearing participants

	Deaf	Hard of hearing
GCSE/O-levels	3	1
A-levels	2	4
University level	4	5

In line with UCL hourly rates for experimental participants, hearing participants received £10 for their participation in the experiment. In recognition of the greater difficulty in recruiting special populations, hard of hearing and deaf participants were paid £25. Travel expenses were reimbursed as required.

Materials

These comprised two self-contained 1-minute scenes from films featuring two people engaged in a conversation: one from *Philomena* (Desplat & Frears, 2013) and one from *Chef* (Bespalov & Favreau, 2014). The clips were dubbed into Hungarian – a language unknown to any of the participants and linguistically unrelated to their native languages. Subtitles were displayed in English, while the audio of the films was in Hungarian. Table 6 shows the number of linguistic units manipulated for each clip.

Table 6. Number of instances manipulated for each type of linguistic unit

Linguistic unit	Chef	Philomena
Auxiliary and lexical verb	2	2
Subject and predicate	3	3
Article and noun	3	3
Conjunction between two clauses	4	5

Subtitles were prepared in two versions: syntactically segmented and non-syntactically segmented (see Table 7) (SS and NSS, respectively). The SS condition was prepared in accordance with professional subtitling standards, with linguistic phrases appearing on a single line. In the NSS version, syntactic phrases were split between the first and the second line of the subtitle. Both the SS and the NSS versions had identical time codes and contained exactly the same text. The clip from Philomena contained 16 subtitles, of which 13 were manipulated for the purposes of the experiment; Chef contained 22 subtitles, of which 12 were manipulated. Four types of linguistic units were manipulated in the NSS version of both clips (see Tables 6 and 7).

Each participant watched two clips: one from *Philomena* and one from *Chef*; one in the SS and one in the NSS condition. The conditions were counterbalanced and their order of presentation was randomised using SMI Experiment Centre (see Szarkowska & Gerber-Morón, 2018).

Table 7. Examples of line breaks in the SS and the NSS condition

Linguistic unit	SS condition	NSS condition
Auxiliary and lexical verb	Now, should we have served that sandwich?	Now, should we have served that sandwich?
Subject and predicate	That's my son. Get back in there. We got some hungry people.	That's my son. Get back in there. We got some hungry people.
Article and noun	I've loved the hotels, the food and everything,	I've loved the hotels, the food and everything,
Conjunction between two clauses	Now I've made a decision and my mind's made up.	Now I've made a decision and my mind's made up.

Eye tracking recording

An SMI RED 250 mobile eye tracker was used in the experiment. Participants' eye movements were recorded with a sampling rate of 250Hz. The experiment was designed and conducted with the SMI software package Experiment Suite, using the velocity-based saccade detection algorithm. The minimum duration of a fixation was 80ms. The analyses used SMI BeGaze and SPSS v. 24. Eighteen participants whose tracking ratio was below 80% were excluded from the eye tracking analyses (but not from comprehension or cognitive load assessments).

Dependent variables

The dependent variables were: 3 indicators of cognitive load (difficulty, effort and frustration), comprehension score, and 5 eye tracking measures.

The following three indicators of cognitive load were measured using self-reports on a 1-7 scale: difficulty ("Was it difficult for you to read the subtitles in this clip?", ranging from "very easy" to "very difficult"), effort ("Did you have to put a lot of effort into reading the subtitles in this clip?", ranging from "very little effort" to "a lot of effort"), and frustration ("Did you feel annoyed when reading the subtitles in this clip?", ranging from "not annoyed at all" to "very annoyed").

Comprehension was measured as the number of correct answers to a set of five questions per clip about the content, focussing on the information from the dialogue (not the visual elements). See Szarkowska & Gerber-Morón (2018) for the details, including the exact formulations of the questions.

Table 8 contains a description of the eye tracking measures. We drew individual areas of interest (AOIs) on each subtitle in each clip. All eye tracking data reported here comes from AOIs on subtitles

Table 8. Description of the eye tracking measures

Eye tracking measure	Description
Absolute reading time	The sum of all fixation durations and saccade durations, starting from the duration of the saccade entering the AOI, referred to in SMI software as 'glance duration'. Longer time spent on reading may be indicative of difficulties with extracting information (Holmqvist et al., 2011).
Proportional reading time	The percentage of dwell time (the sum of durations of all fixations and saccades in an AOI starting with the first fixation) a participant spent in the AOI as a function of subtitle display time. For example, if a subtitle lasted for 3 seconds and the participant spent 2.5 seconds in that subtitle, the proportional reading time was 2500/3000 ms = 83% (i.e. while the subtitle was displayed for 3 seconds, the participant was looking at that subtitle for 83% of the time). Longer proportional time spent in the AOI translates into less time available to follow on-screen action.
Mean fixation duration	The duration of a fixation in a subtitle AOI, averaged per clip per participant. Longer mean fixation duration may indicate more effortful cognitive processing (Holmqvist et al., 2011).
Fixation count	The number of fixations in the AOI, averaged per clip per participant. Higher numbers of fixations have been reported in poor readers (Holmqvist et al., 2011).
Revisits	The number of glances a participant made to the subtitle AOI after visiting the subtitle for the first time. Revisits to the AOI may indicate problems with processing, as people go back to the AOI to re-read the text.

Procedure

The study received full ethical approval from the UCL Research Ethics Committee. Participants were tested individually. They were informed they would take part in an eye tracking study on the quality of subtitles. The details of the experiment were not revealed until the debrief.

After reading the information sheet and signing the informed consent form, each participant underwent a 9-point calibration procedure. There was a training session, whose results were not recorded. Its aim was to familiarise the participants with the experimental procedure and the type of questions that would be asked in the experiment (comprehension and cognitive load). Participants watched the clips with the sound on. After the test, participants' views on subtitle segmentation were elicited in a brief interview.

Each experiment lasted approx. 90 minutes (including other tests not reported in this paper), depending on the time it took the participants to

Journal of Eye Movement Research
10.16910/jemr.11.4.2

answer the questions and participate in the interview. Deaf participants had the option of either communicating via a British Sign Language interpreter or by using their preferred combination of spoken language, writing and lip-reading.

Results

Experiment 1

Seventy-four participants took part in this experiment: 27 English, 21 Polish, 26 Spanish.

Cognitive load

To examine whether subtitle segmentation affects viewers' cognitive load, we conducted a 2 x 3 mixed ANOVA on three indicators of cognitive load: difficulty, effort, and frustration, with segmentation as a within-subject independent variable (SS vs. NSS) and language (English, Polish, Spanish) as a between-subject factor. We found a main effect of segmentation on all three aspects of cognitive load, which were consistently higher in the NSS condition compared to the SS one (Table 9).

Table 9. Mean cognitive load indicators for different participant groups in Experiment 1

		Language					
	English	Polish	Spanish	df	*F*	*P*	η_p^2
Difficulty				1,71	15,584	< .001*	.18
SS	2.37 (1.27)	2.05 (1.02)	1.96 (1.14)				
NSS	2.63 (1.44)	2.67 (1.46)	3.42 (1.65)				
Effort				1,71	7,788	.007*	.099
SS	2.78 (1.55)	1.90 (1.26)	2.23 (1.50)				
NSS	2.89 (1.60)	2.43 (1.16)	3.54 (2.10)				
Frustration				1,71	27,030	< .001*	.276
SS	2.15 (1.40)	1.38 (.80)	1.62 (.89)				
NSS	3.04 (1.85)	2.48 (1.91)	3.27 (2.07)				

We also found an interaction between segmentation and language in the case of difficulty, $F(2,71) = 3{,}494$, $p = .036$, $\eta_p^2 = .090$, which we separated with simple effects analyses (post-hoc tests with Bonferroni correction). We found a significant main effect of segmentation on the difficulty of reading subtitles among Spanish participants, $F(1,25) = 19{,}161$, $p < .001$, $\eta_p^2 = .434$. Segmentation did not have a statistically significant effect on the difficulty experienced by English participants, $F(1,26) = ,855$, $p = .364$, $\eta_p^2 = .032$ or by Polish participants, $F(1,20) = 2{,}147$, $p = .158$, $\eta_p^2 = .097$. To recap, although cognitive load difficulty was declared to be higher by all participants in the NSS condition, only in the case of Spanish participants was the main effect of segmentation statistically significant.

We did not find any significant main effect of language on cognitive load (Table 10), which means that participants reported similar scores regardless of their linguistic background.

Table 10. Between-subjects results for cognitive load

Measure	df	*F*	*p*	η_p^2
Difficulty	2,71	.592	.556	.016
Effort	2,71	2.382	.100	.063
Frustration	2,71	1.850	.165	.050

Comprehension

To see whether segmentation affects viewers' performance, we conducted a 2 x 3 mixed ANOVA on segmentation (SS vs. NSS condition) with language (English, Polish, Spanish) as a between-subject factor. The dependent variable was comprehension score. There was no main effect of segmentation on comprehension $F(1,71) = .412$, $p = .523$, $\eta_p^2 = .006$. Table 11 shows descriptive statistics for this analysis. There were no significant interactions.

Table 11. Descriptive statistics for comprehension

	Language	Mean (SD)
Comprehension SS	English	4.11 (1.01)
	Polish	4.48 (.81)
	Spanish	4.08 (1.09)
	Total	4.20 (.99)
Comprehension NSS	English	4.26 (1.02)
	Polish	4.76 (.43)
	Spanish	3.88 (1.21)
	Total	4.27 (1.02)

We found a main effect of language on comprehension, $F(2,71) = 3{,}563$, $p = .034$, $\eta_p^2 = .091$. Pairwise comparisons with Bonferroni correction showed that Polish participants had significantly higher comprehension than Spanish participants, $p = .031$, 95% CI [.05, 1.23]. There was no difference between Polish and English, $p = .224$, 95% CI [-.15,

Journal of Eye Movement Research
10.16910/jemr.11.4.2

1.02], or Spanish and English participants, p =1.00, 95% CI [-.76, .35].

Eye tracking measures

Because of data quality issues, for eye tracking analyses we had to exclude 8 participants from the original sample, leaving 22 English, 19 Polish, and 25 Spanish participants. We found a significant main effect of segmentation on revisits to the subtitle area (Table 12). Participants went back to the subtitles more in the NSS condition (M_{NSS} = .37, SD = .25) compared to the SS one (M_{SS} = .25, SD = .22), implying potential parsing problems. There was no effect of segmentation for any other eye tracking measure (Table 12). There were no interactions.

Table 12. Mean eye tracking measures by segmentation in Experiment 1

	Language						
	English	Polish	Spanish	df	F	p	η_p^2
Absolute reading time (ms)				1,63	2.950	.091	.045
SS	1614	1634	1856				
NSS	1617	1529	1817				
Proportional reading time				1,63	2.128	.150	.033
SS	.65	.67	.76				
NSS	.66	.62	.74				
Mean fixation duration (ms)				1,63	2.128	.906	.000
SS	209	194	214				
NSS	211	187	218				
Fixation count				1,63	2.279	.136	.035
SS	6.41	6.68	7.27				
NSS	6.45	6.42	6.95				
Revisits				1,63	11.839	.001*	.158
SS	.28	.27	.21				
NSS	.39	.34	.36				

In relation to the between-subject factor, we found a main effect of language on absolute reading time, proportional reading time, mean fixation duration, and fixation count, but not on revisits (see Table 13).

Post-hoc Bonferroni analyses showed that Spanish participants spent significantly more time in the subtitle area compared to English and Polish participants. This was shown by significantly longer absolute reading time in the case of Spanish participants compared to English, p = .027, 95% CI [19.20, 422.73], and Polish participants, p = .012, 95% CI [44.61, 464.75]. Polish and English participants did not differ from each other in absolute reading time, p =1.00, 95% CI [-249.88, 182.45]. There was a tendency approaching significance for fixation count to be higher among Spanish participants than English participants, p = .077, 95% CI [-.05, 1.41]. Spanish participants also had higher proportional reading time when compared to English participants, p = .029, 95% CI [.007, .189] and Polish participants, p = .015, 95% CI [.01, .20], i.e. the Spanish participants spent most time reading the subtitle while viewing the clip. Finally, Polish participants had a statistically lower mean fixation duration compared to English, p = .041, 95% CI [-38.10, -59], and Spanish, p = .003, 95% CI [-43.62, -7.16]. English and Spanish participants did not differ from each other in mean fixation duration, p =1.00, 95% CI [-23.55, 11.47].

Table 13. ANOVA results for between-subject effects in Experiment 1

Measure	df	F	p	η_p^2
Absolute reading time	2,63	5.593	.006*	.151
Proportional reading time	2,63	5.398	.007*	.146
Mean fixation duration	2,63	6.166	.004*	.164
Fixation count	2,63	2.980	.058	.086
Revisits	2,63	.332	.719	.010

Overall, the results indicate that the processing of subtitles was least effortful for Polish participants and most effortful for Spanish participants.

Experiment 2

A total of 46 participants (19 males, 27 females) took part in the experiment: 27 were hearing, 10 hard of hearing, and 9 deaf.

Cognitive load

We conducted 2 x 3 mixed ANOVAs on each indicator of cognitive load with segmentation (SS vs. NSS) as a within-subject variable and degree of hearing loss (hearing, hard of hearing, deaf) as a between-subject variable.

Similarly to Experiment 1, we found a significant main effect of segmentation on difficulty, effort, and frustration (Table 14). The NSS subtitles induced higher cognitive load than the SS condition in all groups of participants. There were no interactions.

Table 14. Mean cognitive load indicators for different participant groups in Experiment 2

	Degree of hearing loss						
	Hearing	Hard of hearing	Deaf	df	F	p	η_p^2
	M (SD)	M (SD)	M (SD)				
Difficulty				1,43	6,580	.014*	.133
SS	2.37 (1.27)	1.60 (1.07)	2.56 (1.42)				
NSS	2.63 (1.44)	2.20 (1.31)	3.44 (1.59)				
Effort				1,43	4,372	.042*	.092
SS	2.78 (1.55)	1.60 (1.07)	2.78 (1.64)				
NSS	2.89 (1.60)	2.50 (1.35)	3.44 (1.42)				
Frustration				1,43	7,669	.008*	.151
SS	2.15 (1.40)	1.00 (.00)	2.56 (1.59)				
NSS	3.04 (1.85)	2.10 (1.28)	3.00 (1.58)				

There was no main effect of hearing loss on difficulty, $F(2,43) = 2.100$, $p = .135$, $\eta_p^2 = .089$ or on effort, $F(2,43) = 1.932$, $p = .157$, $\eta_p^2 = .082$, but there was an effect near to significance on frustration, $F(2,43) = 3.100$, $p = .052$, $\eta_p^2 = .129$. Post-hoc tests showed a result approaching significance: hard of hearing participants reported lower frustration levels than hearing participants, $p = .079$, 95% CI [-2.17, .09]. In general, the lowest cognitive load was reported by hard of hearing participants.

Comprehension

Expecting that non-syntactic segmentation would negatively affect comprehension, we conducted a 2 x 3 mixed ANOVA on segmentation (SS vs. NSS) and degree of hearing loss (hearing, hard of hearing, and deaf).

Table 15. Descriptive statistics for comprehension in Experiment 2

	Deafness	Mean (SD)
Comprehension SS	Hearing	4.11 (1.01)
	Hard of hearing	4.60 (.51)
	Deaf	4.00 (.70)
	Total	4.20 (.88)
Comprehension NSS	Hearing	4.26 (1.02)
	Hard of hearing	4.50 (.70)
	Deaf	3.44 (1.23)
	Total	4.15 (1.05)

Note: Maximum score was 5.

Despite our predictions, and similarly to Experiment 1, we found no main effect of segmentation on comprehension $F(1,43) = .713$, $p = .403$, $\eta_p^2 = .016$. There were no interactions.

As for between-subject effects, we found a marginally significant main effect of hearing loss on comprehension, $F(2,43) = 3.061$, $p = .057$, $\eta_p^2 = .125$. The highest comprehension scores were obtained by hard of hearing participants and the lowest by deaf participants (Table 15). Post-hoc analyses with Bonferroni correction showed that deaf participants differed from hard of hearing participants, $p = .053$, 95% CI [-1.66, .01].

Eye tracking measures

Due to problems with calibration, 10 participants had to be excluded from eye tracking analyses, leaving a total of 22 hearing, 8 hard of hearing, and 6 deaf participants.

To examine whether the non-syntactically segmented text resulted in longer reading times, more revisits and higher mean fixation duration, we conducted an analogous mixed ANOVA. We found no main effect of segmentation on any of the eye tracking measures (Table 16), but a few interactions between segmentation and deafness: in absolute reading time, $F(2,33) = 4,205$, $p = .024$, $\eta_p^2 = .203$; proportional reading time, $F(2,33) = 4,912$, $p = .014$, $\eta_p^2 = .229$; fixation count, $F(2,33) = 3,992$, $p = .028$, $\eta_p^2 = .195$; and revisits, $F(2,33) = 6,572$, $p = .004$, $\eta_p^2 = .285$.

Table 16. Mean eye tracking measures by segmentation in Experiment 2

	Hearing loss						
	Hearing	Hard of hearing	Deaf	Df	F	p	η_p^2
Absolute reading time (ms)				1,33	1.752	.195	.050
SS	1614	1619	1222				
NSS	1617	1519	1522				
Proportional reading time				1,33	2.270	.141	.064
SS	.65	.66	.45				
NSS	.66	.61	.62				
Mean fixation duration				1,33	.199	.659	.006
SS	209	199	214				
NSS	211	185	219				
Fixation count				1,33	2.686	.111	.075
SS	6.41	6.73	4.63				
NSS	6.45	6.45	5.90				
Revisits				1,33	.352	.557	.011
SS	.28	.20	.45				
NSS	39	.30	.15				

We broke down the interactions with simple-effects analyses by means of post-hoc tests using Bonferroni correction. In the deaf group, we found an effect of segmentation on revisits approaching significance, $F(1,5) = 5.934$, $p = .059$, $\eta_p^2 = .543$. Deaf participants had more revisits in the SS condition than in the NSS one, $p = .059$. They also had a higher absolute reading time, proportional reading time, and fixation count in the NSS compared to the SS condition, but possibly owing to the small sample size, these differences did not reach statistical significance. In the hard of hearing group, there was no significant main effect of segmentation on any of the eye tracking measures ($ps > .05$). In the hearing group, there was no statistically significant main effect of segmentation (all $ps > .05$).

A between-subject analysis showed a close to significant main effect of degree of hearing loss on fixation count, $F(2,33) = 3.204$, $p = .054$, $\eta_p^2 = .163$. Deaf participants had fewer fixations per subtitle compared to hard of hearing, $p = .088$, 95% CI [-2.79, .14], or hearing participants, $p = .076$, 95% CI [-2.41, .08]. No other measures were significant.

Interviews

Following the eye tracking tests, we conducted short semi-structured interviews to elicit participants' views on subtitle segmentation, complementing the quantitative part of the study (Bazeley, 2013). We used inductive coding to identify themes reported by participants. Several Spanish, Polish, and deaf participants said that keeping units of meaning together contributed to the readability of subtitles because by creating false expectations (i.e. "garden path" sentences), NSS line-breaks can require more effort to process. These participants believed that chunking text by phrases according to "natural thoughts" allowed subtitles to be read quickly. In contrast, other participants said that NSS subtitles gave them a sense of continuity in reading the subtitles. A third theme in relation to dealing with SS and NSS subtitles was that participants adapted their reading strategies to different types of line-breaks. Finally, a number of people also admitted they had not noticed any differences in the subtitle segmentation between the clips, saying they had never paid any attention to subtitle segmentation.

Discussion

The two experiments reported in this paper examined the impact of text segmentation in subtitles on cognitive load and reading performance. We also investigated whether viewers' linguistic background (native language and hearing status) impacts on how they process syntactically and non-syntactically segmented subtitles. Drawing on the large body of literature on text segmentation in subtitling (Díaz Cintas & Remael, 2007; Ivarsson & Carroll, 1998; Perego, 2008a, 2008b; Rajendran et al., 2013) and literature on parsing and text chunking during reading (Keenan, 1984; Kennedy et al., 1989; LeVasseur, Macaruso, Palumbo, & Shankweiler, 2006; Mitchell, 1987, 1989; Rayner et al., 2012), we predicted that subtitle reading would be adversely affected by non-syntactic segmentation.

This prediction was partly upheld. One of the most important findings of this study is that participants reported higher cognitive load in non-syntactically segmented (NSS) subtitles compared to syntactically segmented (SS) ones. In both experiments, mental effort, difficulty, and frustration were reported as higher in the NSS condition. A possible explanation of this finding may be that NSS text increases extraneous load, i.e. the type of cognitive load related to the way information is presented (Sweller et al., 1998). Given the limitations of working memory capacity (Baddeley, 2007; Chandler & Sweller, 1991), NSS may leave less capacity to process the remaining visual, auditory, and textual information. This, in turn, would increase their frustration, make them expend more effort and lead them to perceive the task as more difficult.

Although cognitive load was found to be consistently higher in the NSS condition across the board in all participant groups, the mean differences between the two conditions do not differ substantially and thus the effect sizes are not large. We believe the small effect size may stem from the fact that the clips used in this study were quite short. As cognitive fatigue increases with the length of the task, and declines simultaneously in performance (Ackerman & Kanfer, 2009; Sandry, Genova, Dobryakova, DeLuca, & Wylie, 2014; Van Dongen, Belenky, & Krueger, 2011), we might expect that in longer clips with non-syntactically segmented subtitles, the cognitive load would accumulate over time, resulting in more prominent mean differences between the two conditions. We acknowledge that the short duration of clips, necessitated by the length of the entire experiment, is an important limitation of this study. However, a number of previous studies on subtitling have also used very short clips (Jensema, 1998; Jensema, El Sharkawy, Danturthi, Burch, & Hsu, 2000; Rajendran et al., 2013; Romero-Fresco, 2015). In this study, we only examined text segmentation within a single subtitle; further research should also explore the effects of non-syntactic segmentation across two or more consecutive subtitles, where the impact of NSS subtitles on cognitive load may be even higher.

Despite the higher cognitive load and contrary to our predictions, we found no evidence that subtitles which are not segmented in accordance with professional standards result in lower comprehension. Participants coped well in both conditions, achieving similar comprehension scores regardless of segmentation. This finding is in line with the results reported by Perego et al. (2010), using Italian participants, that subtitles containing non-syntactically segmented noun phrases did not negatively affect participants' comprehension. Our research extends these findings to other linguistic units in English (verb phrases and conjunctions as well as noun phrases) and other groups of participants (hearing English, Polish, and Spanish speakers, as well as deaf and hard of hearing participants). The finding that performance in processing NSS text is not negatively affected despite the participants' extra effort (as shown by increased cognitive load) may be attributed to the short duration of the clips and also to overall high comprehension scores. As the clips were short, there were limited points that could be included in the comprehension questions. Other likely reasons for the lack of significant differences between the two conditions is the extensive experience that all the participants had of using subtitles in the UK, and that participants may have become accustomed to subtitling not adhering to professional segmentation standards. Our sample of participants was also relatively well-educated, which may have been a reason for their comprehension scores being near ceiling. Furthermore, as noted by Mitchell (1989), when interpreting the syntactic structure of sentences in reading, people use non-lexical cues such as text layout or punctuation as parsing aids, although these cues are of secondary importance when compared to words, which constitute "the central source of information" (p. 123). This is also consistent with what the participants in our study reported in the interviews. For example, one deaf participant said: "Line breaks have their value, yet when you are reading fast, most of the time it becomes less relevant."

In addition to understanding the effects of segmentation on subtitle processing, this study also found interesting results relating to differences in subtitle processing between the different groups of viewers. In Experiment 1, Spanish participants had the highest cognitive load and lowest comprehension, and spent more time reading subtitles than Polish and English participants. Although it is impossible to attribute these findings unequivocally to Spanish participants coming from a dubbing country, this finding may related to their experience of having grown up exposed more to dubbing than subtitling. In Experiment 2, we found that subtitle processing was the least effortful for the hard of hearing group: they reported the lowest cognitive effort and had the highest comprehension score. This result may be attributed to their high familiarity with subtitling (as declared in the pre-test questionnaire) compared to the hearing group. Although no data were obtained for the groups in Experiment 2 in relation to English literacy measures, as a group, individuals born deaf or deafened early in life have low average reading ages, and more effortful processing by the deaf group may be related to lower literacy.

Different viewers adopt different strategies to cope with reading NSS subtitles. In the case of hearing participants, there were more revisits to the subtitle area for NSS subtitles, which is a likely indication of parsing difficulties (Rayner et al., 2012). In the group of participants with hearing loss, deaf people spent more time reading NSS subtitles than SS ones. Given that longer reading time may indicate difficulty in extracting information (Holmqvist et al., 2011), this may also be taken to reflect parsing problems. This interpretation is also in accordance with the longer durations of fixations in the deaf group, which is another indicator of processing difficulties (Holmqvist et al., 2011; Rayner, 1998). Unlike the findings of other studies (Krejtz et al., 2016; Szarkowska et al., 2011; Szarkowska, Krejtz, Dutka, et al., 2016), in this study, deaf participants fixated less on the subtitles than hard of hearing and hearing participants. Our results, however, are in line with a recent eye tracking study (Miquel Iriarte, 2017), where deaf people also had fewer fixations than relation hearing viewers. According to Miquel Iriarte (2017), deaf viewers relate to the visual information on the screen as a whole to a greater extent than hearing viewers, reading the subtitles faster to give them more time to direct their attention towards the visual narrative.

Conclusions

Our study has shown that text segmentation influences the processing of subtitled videos: non-syntactically segmented subtitles may increase viewers' cognitive load and eye movements. This was particularly noticeable for Spanish and deaf people. In order to enhance the viewing experience, using syntactic segmentation in subtitles may facilitate the process of reading subtitles, thus giving viewers greater time to follow the visual narrative of the film. Further research is necessary to disentangle the impact of the viewers' country of origin, familiarity with subtitling, reading skills, and language proficiency on subtitle processing.

This study also provides support for the need to base subtitling guidelines on research evidence, particularly in view of the tremendous expansion of subtitling across different media and formats. The results are directly applicable to current practices in television broadcasting and video-on-demand services. They can also be adopted in subtitle personalization to improve automation algorithms for subtitle display in order to facilitate the processing of subtitles among the myriad different viewers using subtitles.

Ethics and Conflict of Interest

The author(s) declare(s) 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 (2017SGR113).

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

References

Ackerman, P. L., & Kanfer, R. (2009). Test Length and Cognitive Fatigue: An Empirical Examination of Effects on Performance and Test-Taker Reactions. *Journal of Experimental Psychology: Applied*, *15*(2), 163–181. https://doi.org/10.1037/a0015719

Allen, P., Garman, J., Calvert, I., & Murison, J. (2011). Reading in multimodal environments: assessing legibility and accessibility of typography for television. In *The proceedings of the 13th international ACM SIGACCESS conference on computers and accessibility* (pp. 275–276). https://doi.org/10.1145/2049536.2049604

Antonenko, P., Paas, F., Grabner, R., & van Gog, T. (2010). Using Electroencephalography to Measure Cognitive Load. *Educational Psychology Review*, *22*(4), 425–438. https://doi.org/10.1007/s10648-010-9130-y

Baddeley, A. (2007). *Working memory, thought, and action*. Oxford: Oxford University Press.

Bazeley, P. (2013). *Qualitative Data*. Los Angeles: Sage.

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

Bisson, M.-J., Van Heuven, W. J. B., Conklin, K., & Tunney, R. J. (2012). Processing of native and foreign language subtitles in films: An eye tracking study. *Applied Psycholinguistics*, *35*(02), 399–418. https://doi.org/10.1017/s0142716412000434

Chandler, P., & Sweller, J. (1991). Cognitive Load Theory and the Format of Instruction. *Cognition and Instruction*, *8*(4), 293–332. https://doi.org/10.1207/s1532690xci0804_2

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

d'Ydewalle, G., Praet, C., Verfaillie, K., & Van Rensbergen, J. (1991). Watching Subtitled Television: Automatic Reading Behavior. *Communication Research*, *18*(5), 650–666.

D'Ydewalle, G., Rensbergen, J. Van, & Pollet, J. (1987). Reading a message when the same message is available auditorily in another language: the case of subtitling. In J. K. O'Regan & A. Levy-Schoen (Eds.), *Eye movements: from physiology to cognition* (pp. 313–321). Amsterdam/New York: Elsevier.

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 T. Dwyer, C. Perkins, S. Redmond, & J. Sita (Eds.), *Seeing into Screens: Eye Tracking and the Moving Image* (pp. 46–64). London: Bloomsbury.

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

Gernsbacher, M. A. (2015). Video Captions Benefit Everyone. *Policy Insights from the Behavioral and Brain Sciences*, *2*(1), 195–202. https://doi.org/10.1177/2372732215602130

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam: North-Holland.

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.

Jensema, C. (1998). Viewer Reaction to Different Television Captioning Speeds. *American Annals of the Deaf*, *143*(4), 318–324.

Jensema, C., El Sharkawy, S., Danturthi, R. S., Burch, R., & Hsu, D. (2000). Eye Movement Patterns of Captioned Television Viewers. *American Annals of the Deaf*, *145*(3), 275–285.

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

Keenan, S. A. (1984). Effects of Chunking and Line Length on Reading Efficiency. *Visible Language*, *18*(1), 61–80.

Kennedy, A., Murray, W. S., Jennings, F., & Reid, C. (1989). Parsing Complements: Comments on the Generality of the Principle of Minimal Attachment. *Language and Cognitive Processes*, *4*(3–4), SI51-SI76. https://doi.org/10.1080/01690968908406363

Koolstra, C. M., Van Der Voort, T. H. A., & D'Ydewalle, G. (1999). Lengthening the Presentation Time of Subtitles on Television: Effects on Children's Reading Time and Recognition. *Communications*, *24*(4), 407–422. https://doi.org/10.1515/comm.1999.24.4.407

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., & Doherty, S. (2016). Measuring Cognitive Load in the Presence of Educational Video: Towards a Multimodal Methodology. *Australasian Journal of Educational Technology*, *32*(6), 19–31.

https://doi.org/10.14742/ajet.3084

Kruger, J.-L., Doherty, S., Fox, W., & de Lissa, P. (2017). Multimodal measurement of cognitive load during subtitle processing: Same-language subtitles for foreign language viewers. In I. Lacruz & R. Jääskeläinen (Eds.), *New Directions in Cognitive and Empirical Translation Process Research* (pp. 267–294). London: John Benjamins.

Kruger, J.-L., Hefer, E., & Matthew, G. (2013). Measuring the impact of subtitles on cognitive load: eye tracking and dynamic audiovisual texts. *Proceedings of the 2013 Conference on Eye Tracking South Africa*. Cape Town, South Africa: ACM. https://doi.org/10.1145/2509315.2509331

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. https://doi.org/http://dx.doi.org/10.16910/jemr.7.5.4

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. M., Macaruso, P., Palumbo, L. C., & Shankweiler, D. (2006). Syntactically cued text facilitates oral reading fluency in developing readers. *Applied Psycholinguistics*, *27*(3), 423–445. https://doi.org/10.1017/s0142716406060346

Łuczak, K. (2017). *The effects of the language of the soundtrack on film comprehension, cognitive load and subtitle reading patterns. An eye-tracking study.* University of Warsaw, Warsaw.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967.

Matamala, A., Perego, E., & Bottiroli, S. (2017). Dubbing versus subtitling yet again? *Babel*, *63*(3), 423–441. https://doi.org/10.1075/babel.63.3.07mat

Mayberry, R. I., del Giudice, A. A., & Lieberman, A. M. (2011). Reading Achievement in Relation to Phonological Coding and Awareness in Deaf Readers: A Meta-analysis. *Journal of Deaf Studies and Deaf Education*, *16*(2), 164–188. https://doi.org/10.1093/deafed/enq049

Miquel Iriarte, M. (2017). *The reception of subtitling for the deaf and hard of hearing: viewers' hearing and communication profile & Subtitling speed of exposure*. Universitat Autònoma de Barcelona, Barcelona.

Mitchell, D. C. (1987). Lexical guidance in human parsing: Locus and processing characteristics. In M. Coltheart (Ed.), *Attention and Performance XII*. London: Lawrence Erlbaum Associates Ltd.

Mitchell, D. C. (1989). Verb guidance and other lexical effects in parsing. *Language and Cognitive Processes*, *4*(3–4), SI123-SI154. https://doi.org/10.1080/01690968908406366

Musselman, C. (2000). How Do Children Who Can't Hear Learn to Read an Alphabetic Script? A Review of the Literature on Reading and Deafness. *Journal Of Deaf Studies And Deaf Education*, *5*(1), 9–31.

Ofcom. (2015). Ofcom's Code on Television Access Services.

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

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

Romero-Fresco, P. (2015). *The reception of subtitles for the deaf and hard of hearing in Europe*. Bern: Peter Lang.

Sandry, J., Genova, H. M., Dobryakova, E., DeLuca, J., & Wylie, G. (2014). Subjective Cognitive Fatigue in Multiple Sclerosis Depends on Task Length. *Frontiers in Neurology*, *5*, 214. https://doi.org/10.3389/fneur.2014.00214

Schmeck, A., Opfermann, M., van Gog, T., Paas, F., & Leutner, D. (2014). Measuring cognitive load with subjective rating scales during problem solving: differences between immediate and delayed ratings. *Instructional Science*, *43*(1), 93–114. https://doi.org/10.1007/s11251-014-9328-3

Sun, S., & Shreve, G. (2014). Measuring translation difficulty: An empirical study. *Target*, *26*(1), 98–127. https://doi.org/10.1075/target.26.1.04sun

Sweller, J. (2011). Cognitive Load Theory. *Psychology of Learning and Motivation*, *55*, 37–76.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). Cognitive Load Theory. *Psychology of Learning and Motivation*, *55*, 37–76.

Sweller, J., Van Merrienboer, J., & Paas, F. (1998). Cognitive Architecture and Instructional Design. *Educational Psychology Review*, *10*(3), 251–296. https://doi.org/10.1023/a:1022193728205

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

Szarkowska, A., Krejtz, I., Kłyszejko, 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).

Szarkowska, A., Krejtz, I., Pilipczuk, O., Dutka, Ł., & Kruger, J.-L. (2016). The effects of text editing and subtitle presentation rate on the comprehension and reading patterns of interlingual and intralingual subtitles among deaf, hard of hearing and hearing viewers. *Across Languages and Cultures*, *17*(2), 183–204. https://doi.org/10.1556/084.2016.17.2.3

Szarkowska, A., Krejtz, K., Dutka, Ł., & Pilipczuk, O. (2016). Cognitive load in intralingual and interlingual respeaking-a preliminary study. *Poznan Studies in Contemporary Linguistics*, *52*(2). https://doi.org/10.1515/psicl-2016-0008

Thorn, F., & Thorn, S. (1996). Television Captions for Hearing-Impaired People: A Study of Key Factors that Affect Reading Performance. *Human Factors*, *38*(3), 452–463.

Van Dongen, H. P. A., Belenky, G., & Krueger, J. M. (2011). *Investigating the temporal dynamics and underlying mechanisms of cognitive fatigue.* American Psychological Association. https://doi.org/10.1037/12343-006

Van Gerven, P. W. M., Paas, F., Van Merriënboer, J. J. G., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology*, *41*(2), 167–174. https://doi.org/10.1111/j.1469-8986.2003.00148.x

van Gog, T., & Paas, F. (2008). Instructional Efficiency: Revisiting the Original Construct in Educational Research. *Educational Psychologist*, *43*(1), 16–26. https://doi.org/10.1080/00461520701756248

Wang, Z., & Duff, B. R. L. (2016). All Loads Are Not Equal: Distinct Influences of Perceptual Load and Cognitive Load on Peripheral Ad Processing. *Media Psychology*, *19*(4), 589–

613. https://doi.org/10.1080/15213269.2015.1108204

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

Yoon, J.-O., & Kim, M. (2011). The Effects of Captions on Deaf Students' Content Comprehension, Cognitive Load, and Motivation in Online Learning. *American Annals of the Deaf*, *156*(3), 283–289. https://doi.org/10.1353/aad.2011.0026

Article 2

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. DOI: http://dx.doi.org/10.16910/jemr.11.3.2

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.

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

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

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).

No se puede mostrar la imagen en este momento.

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.

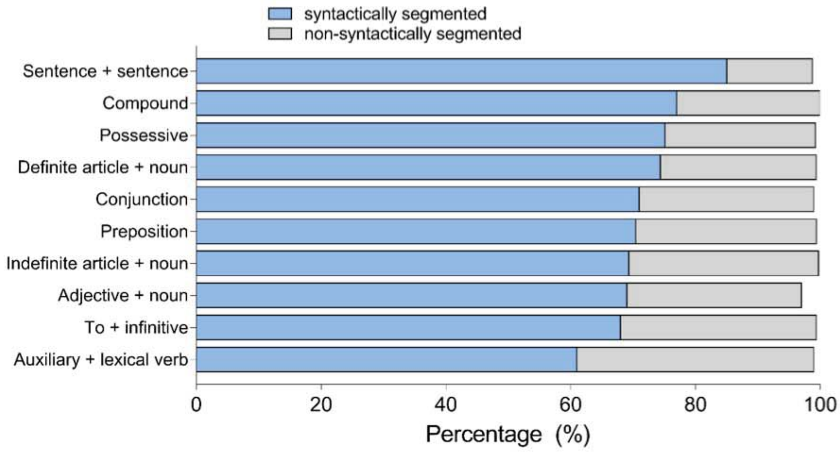


Figure 3. Preferences for SS and NSS subtitles by linguistic units in Experiment 1.

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.

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.

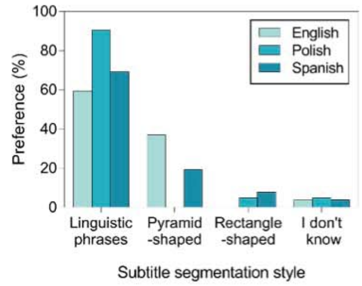


Figure 4. Segmentation preferences by group and style.

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.

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.

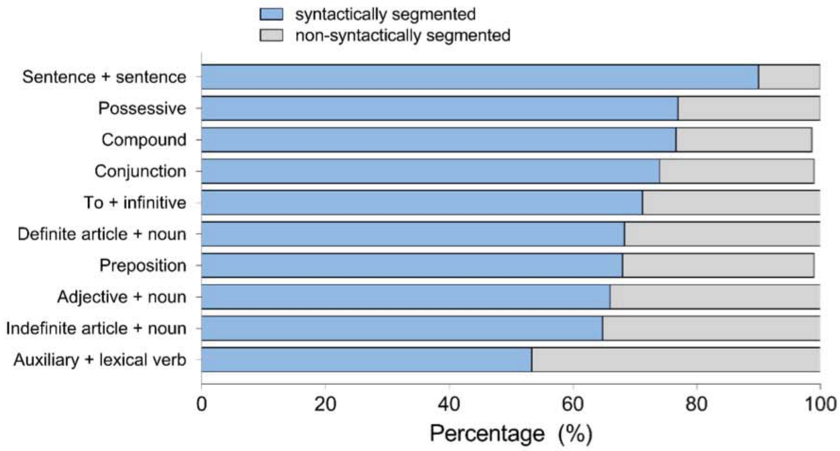


Figure 5. Preferences for SS and NSS subtitles by linguistic units in Experiment 2. Table 8. Percentage of Experiment 2 participants who preferred the syntactically segmented condition.

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.

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

Linguistic unit	Degree of hearing loss: Hearing	Hard of hearing	Deaf	*df*	*F*	*p*	η_p^2
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

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.

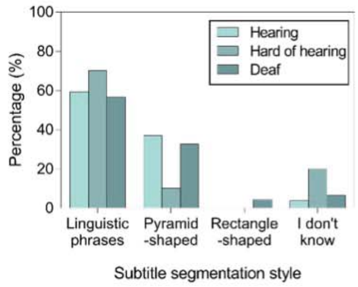


Figure 6. Segmentation preferences by group.

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, η_p^2=.378, and *Conj*, $F(2,26)$= 4.334, p=.024, η_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, η_p^2 =.819, and *Conj*, $F(1,7)$=8,306, p=.024, η_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, η_p^2= .228) and a tendency approaching significance in *SentSent* ($F(2,26)$=3.241, p=.055, η_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: Hearing	Hard of hearing	Deaf	df	F	p	η_p^2
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

No se puede mostrar la imagen en este momento.

No se puede mostrar la imagen en este momento.

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 <u>the memory stick</u>
he managed to hide.

and

(NSS) He's looking for <u>the memory</u>
<u>stick</u> 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.

No se puede mostrar la imagen en este momento.

No se puede mostrar la imagen en este momento.

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

Article 3

Gerber-Morón, O., Soler-Vilageliu, O., & Castellà, J. (forthcoming). The effects of screen size on subtitle layout preferences and comprehension across devices. *Hermēneus*, *Revista de traducción e interpretación* 2019, *21*. Manuscript accepted for publication.

EDICIONES Universidad de Valladolid

Efectos del tamaño de pantalla en las preferencias de presentación de subtítulos y en la comprensión en distintos dispositivos*

The effects of screen size on subtitle layout preferences and comprehension across devices

OLIVIA GERBER-MORÓN
UAB-Universitat Autònoma de Barcelona. Department of Translation and Interpreting & East Asian Studies. Faculty of Translation and Interpreting. Edifici MRA126, Campus UAB, 08193 Bellaterra, Spain.
Dirección de correo electrónico: Olivia.Gerber@uab.cat
ORCID: 0000-0001-6513-3662.
OLGA SOLER-VILAGELIU
UAB-Universitat Autònoma de Barcelona. Department of Basic, Educational, and Developmental Psychology. Faculty of Psychology. Edifici B, Carrer de la Fortuna, Campus UAB, 08193 Bellaterra, Spain.
Dirección de correo electrónico: Olga.Soler@uab.cat
ORCID: 0000-0001-9219-1913.
JUDIT CASTELLÀ
UAB-Universitat Autònoma de Barcelona. Department of Translation and Interpreting & East Asian Studies. Faculty of Translation and Interpreting. Edifici MRA126, Campus UAB, 08193 Bellaterra, Spain.
Dirección de correo electrónico: Judit.Castella@uab.cat
ORCID: 0000-0002-6094-3516.
Recibido: 29/1/2018. Aceptado: 17/7/2018.
Cómo citar: Gerber-Morón, Olivia, Soler-Vilageliu, Olga and Castellà, Judit, «The effects of screen size on subtitle layout preferences and comprehension across devices», Hermēneus. Revista de traducción e interpretación 21 (2019).
DOI:

Resumen: Esta investigación analiza los efectos que el tamaño de pantalla puede tener en las preferencias y en la comprensión de material audiovisual subtitulado. Treinta participantes vieron tres fragmentos de vídeo en tres dispositivos con distintos tamaños de pantalla (monitor, tableta y smartphone). Tras ver los fragmentos, los participantes respondieron a una serie de cuestionarios de preferencias y comprensión. A través del análisis de la recepción de los subtítulos en diferentes tamaños de pantalla, esta investigación tiene por objeto aportar nuevas pruebas empíricas sobre las necesidades y preferencias de los espectadores. Los resultados muestran que el dispositivo que percibe los efectos más negativos es el smartphone, lo que plantea la necesidad de seguir investigando los dispositivos con pantallas más pequeñas para mejorar la lectura de los subtítulos y adaptarlos en función del tamaño.
Palabras clave: accesibilidad; nuevas tecnologías; legibilidad; tamaño de pantalla; subtítulos.
Abstract: The present study sheds light on the possible effects that screen size can have on preferences and comprehension of subtitled audiovisual material content. Thirty participants watched three subtitled video excerpts displayed on three devices with different screen size (monitor, tablet, and smartphone). After watching each excerpt, they filled out preference and comprehension questionnaires. This study aimed to provide new empirical evidence on viewers' needs and preferences concerning readability by analysing the reception of subtitles across screens. The results obtained indicate that smartphone devices had the most unsatisfactory effects, suggesting the need to undertake further research on small screens to improve subtitle readability.
Keywords: accessibility; new technologies; readability; screen size; subtitling

*This research study has been partially supported by "La Caixa" Foundation (E-08-2014-1306365) and Transmedia Catalonia Research Group (2017SGR113).

subtítulos, 2.2.2. Cuestionario de comprensión, 2.3. Diseño y método; Resultados, 1. Análisis de comprensión y legibilidad, 2. Análisis de preferencias de subtítulos; Discusión; Conclusiones.

Summary: Introduction; 1. Readability and layout parameters in subtitles, 2. Screen size effects across devices, 3. Overview of the study; Methods; 1. Participants, 2. Materials, 2.1. Stimuli and apparatus, 2.1.1. Video fragments, 2.1.2. Subtitles, 2.2. Questionnaires, 2.2.1 Questionnaire on reading and layout preferences for subtitles, 2.2.2. Comprehension questionnaire, 2.3. Design and Procedure; Results, 1. Comprehension and readability items, 2. Subtitle preference items; Discussion; Conclusions.

INTRODUCTION

The development of new technologies in the past decades has changed the way audiovisual products are consumed nowadays (Messerlin, Siwek, & Cocq, 2005). Innovative handheld devices, such as tablets and smartphones, provide the mobility to consume media everywhere (Palen, Salzman, & Youngs, 2000). The implementation of subtitles on these handheld devices makes video content accessible to different end-users, such as non-native speakers, deaf and hard-of-hearing viewers. Subtitles on mobile devices are also useful when sound has to be removed in public spaces. Because watching subtitled media on these devices is continually increasing in our society, it is important to present subtitles in the most effective way. This study examines the effects of screen size on different subtitle layout parameters, with a view to improving the most determining factor in subtitling: readability. The process of readability becomes more complex with subtitled media because viewers are continually switching from text to image (d'Ydewalle, Van Rensbergen, & Pollet, 1987), without having control over the speed in subtitling (Romero-Fresco, 2015). We predict that subtitle readability may be hindered by the smaller screen size of handheld devices.

1. READABILITY AND LAYOUT PARAMETERS IN SUBTITLES

Scholars have established various parameters that need to be considered to improve the readability of subtitles. Karamitroglou (1998) and Perego (2005, 2008) distinguished three categories of parameters that affect the legibility and readability of subtitles: duration, text editing, and subtitle layout parameters. Duration parameters comprise the line length of time the subtitles are on the screen, the leading-in and lagging-out time for each subtitle, the time break between two consecutive subtitles, and camera takes and cuts (Perego, 2005, 2008). Text editing parameters relate to punctuation and letter case, line breaks and line length, altering syntactic structures, omitting and retaining linguistic items of the original. In relation to layout parameters, Gottlieb (1992) compiled a list that included the position of subtitles on the screen, the number of lines, the number of characters per line, text alignment, typeface and distribution, and font colour and background.

Media regulators and professionals in the audiovisual industry have partly integrated these parameters in their guidelines to enhance the quality of subtitling (BBC, 2017; Described and Captioned Media Program, 2017; Ofcom, 2015; Media Access Australia, 2012). As a case in point, BBC's subtitles guidelines (BBC, 2017) recommends the use of one-line subtitle instead of two short lines because it takes less time to read and causes less disruption to the picture. In our study, we tested some of the layout parameters listed by the scholars previously mentioned to examine how viewers perceive subtitles across devices.

ISSN: 2530-609X

2. Screen size effects across devices

To the best of our knowledge, subtitle layout parameters have not been studied across devices. Nevertheless, other studies have been conducted on the effects of screen size in the fields of Audiovisual Translation. Two eye-tracking studies on watching subtitled videos across screen devices have shown more negative results in smartphone devices (Castellà, Olivier, Gerber, & Soler, 2016; Szarkowska, Laskowska, Oliver, & Pilipczuk, 2015). Szarkowska, A. et al. (2015) studied reading patterns on smartphone, tablet and computer screen, and found evidence that smartphone has the lowest comprehension results, the longest mean fixation duration, and fewer fixations in comparison to tablet and monitor screens. In their eye-tracking study on watching subtitled videos on different screen devices, Castellà et al. (2016) suggested that smartphone devices require more cognitive load when reading subtitles than tablets and monitors.

A number of studies in the fields of Media Psychology, and Human-Computer Interaction (Al-Showarah, AL-Jawad, & Sellahewa, 2014; Kim, Sundar, & Park, 2011; Lombard, Ditton, Grabe, & Reich, 1997; Maniar, Bennett, Hand, & Allan, 2008) have also focused their research of screen size on viewers' perception of mobility and content, and on attitudes towards technology. Lombard et al. (1997) studied the role of screen size in small and large television screens. They measured responses via a questionnaire and found that large screen televisions elicit more intense responses for some genres (commercials, action-adventure, and reality) but not for others (talk shows and drama programs). Maniar et al. (2008) looked at the effect of screen size on video-based learning by presenting videos on small, medium and large screen mobile phones. Their results from the questionnaires pointed out that larger screens induce more attention than medium and small screens. Moreover, they found that smaller screen displays may inhibit the effectiveness of the learning experience. Kim et al. (2011) carried out a study on the effects of screen size (across three different mobile phone devices) and communication modality (video format or text document) to assess through questionnaires the users' perception of mobility and content, and the degree of technology acceptance. Their results revealed that screen size does not affect the understanding of the news story or the perceived ease of use of the device. Nevertheless, it seems that larger screen size is the key to greater enjoyment for their participants. In another study, Al-Showarah et al. (2014) evaluated the effects of screen size on smartphone and tablet usability across age groups and found that seniors show more difficulties in processing information on smartphone screens. Their eye-tracking results also showed that usability on a small screen size is more difficult for all age groups in comparison to large screen sizes. In general, all these studies indicate that large screen displays tend to contribute to a more satisfying experience.

3. Overview of the study

The main goal of this study is to analyse different subtitle layout preferences and comprehension scores across devices with different screen size (monitor, tablet, and smartphone). This study aims to offer more insight on viewers' needs and preferences by specifically testing subtitle layout parameters not covered in previous studies on subtitles across devices (Castellà et al., 2016; Szarkowska et al., 2015). This study also aims to validate Szarkowska et al. (2015) as regards the effects of screen size on comprehension scores.

Based on the studies previously conducted on screen size (Al-Showarah et al., 2014; Castellà et al., 2016; Maniar et al., 2008; Szarkowska, A. et al., 2015), we expected to find differences in the overall viewers' evaluation of the subtitle layout parameters depending on the device, and in particular, we expected to obtain more negative results in the smallest screen devices (i.e. smartphones).

METHODS

1. PARTICIPANTS

The study involved 30 volunteer participants ranging from 18 to 58 years of age (16 females, 14 males, mean age=30.5, SD=7.6). They were all Spanish native speakers or Catalan-Spanish bilingual with normal or corrected-to-normal (contact lenses or glasses) vision. Most of the participants were university students from Spain or other Spanish-speaking countries. The majority of the participants reported not being habitual viewers of subtitled audiovisual material. None of the participants had any knowledge of the original language used for the film fragments (Norwegian).

2. MATERIALS

2.1. Stimuli and apparatus

2.1.1. Video fragments

The stimuli were three short video fragments with Spanish subtitles taken from a Norwegian thriller (*Hodejegerne*, Tyldum, 2011). Each video fragment formed full scenes with coherent content, and the average duration of each of them was three minutes. We used a Norwegian film to expose participants to an unknown language, so that they would have to rely on the information provided by the subtitles to follow the video fragments.

2.1.2. Subtitles

The subtitles were created using EZTitles[1], a professional subtitle editing software. As for the technical considerations, we followed the recommendations by Díaz Cintas and Remael (2007) for synchronization and presentation, using 15 subtitling spaces per second and lines of 38 characters. On average, each video fragment contained 37 subtitles: 15 sentences occupied one line of text and 22 occupied two lines. The video fragments and synchronised subtitles were presented using the freeware VLC Media Player on the three devices tested in the experiment: a 22-inch Toshiba TV monitor, a 9.7-inch iPad 2 and a 3.5-inch iPhone 4.

[1] For more detailed information, see <http://www.eztitles.com> (last accessed 30 November 2017).

ISSN: 2530-609X

2.2. Questionnaires

2.2.1. Questionnaire on reading and layout preferences for subtitles

The preference questionnaire was administered to check viewers' reception of subtitles and their preferences concerning the general layout according to the screen size of each device. The questions for this study were inspired by Gottlieb (1992) and Gambier (2009), who provided a list of subtitle parameters to measure and evaluate the viewers' reception of subtitle readability. The questionnaire assessed the experience of reading subtitles on each device by asking questions on the following parameters:

- the percentage of subtitles read;
- the ease of subtitle reading;
- the overall assessment on viewing the film excerpt on that device;
- the feeling of having lost essential parts of the plot (due to the fact of reading subtitles);
- the line length of the subtitles;
- the exposure time of subtitles;
- the line-break layout (i.e. division of lines on screen).

The questionnaire on reading and layout preferences for subtitles for this study included three questions on a 5-point Likert scale concerning the percentage of subtitles read («What percentage of subtitles didn't you have time to read?» from 0% to 100%), the ease of reading subtitles on that device («How did you find reading subtitles on this device?» from very difficult to very easy) and the experience of viewing the film excerpt on that device («How would you rate the experience of watching a film on this device?» from very unpleasant to very pleasant). Furthermore, a yes/no question asked about the feeling of having lost essential parts of the film's action due to the fact of reading subtitles («Do you think that you lost essential parts of the film's action due to the fact of having to read subtitles?»). Here is a sample of one of these questions on the ease of reading subtitles:

Example (1)
How did you find reading subtitles on this device?
- Very easy
- Easy
- Moderate
- Difficult
- Very difficult

In addition to these questions, participants had to answer three categorical questions on a 5-point scale about subtitle preferences for the line length and exposure time on the screen («What do you think about the length of the subtitles for this device?» and «What do you think about the exposure time of the subtitles on this screen?», where 1=very long, 2=long, 3=appropriate, 4=short, 5=very short), and line-break layout («What do you think about the line-break layout for this device?», where 1=unsuitable, 2=I would have preferred shorter subtitles of one line, 3=I would have preferred shorter subtitles, but two lines, 4=I would have preferred longer subtitles, but of one line, 5=appropriate). Here is a sample of one of these categorical questions on the line length of subtitles:

ISSN: 2530-609X

No se puede mostrar la imagen en este momento.

No se puede mostrar la imagen en este momento.

within each device (all $p<.001$). Across the three devices, results showed that the majority of the participants found that the line length of the subtitles was appropriate, especially for the tablet device (86.7%). However, there was a minor tendency to report the line length for the monitor device as long (20%) and, at a slightly lower percentage, for the smartphone device (13.3%). These trends can be seen in (Table 2).

Table 2. Question «What do you think about the length of the subtitles for this device?» Percentage of each response option and Chi-Square values, as a function of device.

LENGTH	Very long	Long	Appropriate	Short	Very short	$\chi2$ (2, n=30)
Smartphone	0%	13.3%	80%	6.7%	0%	29.60, p<.001
Tablet	0%	3.3%	86.7%	10%	0%	38.60, p<.001
Monitor	0%	20%	70%	10%	0%	18.60, p<.001

Likewise, and as can be seen in Table 3, the majority of the participants reported the exposure time for the subtitles as appropriate, especially for the smartphone device (86.7%). However, some of the participants stated that the exposure time for the subtitles was short for the tablet (16.7%) and monitor (13.3%) devices.

Table 3. Question «What do you think about the exposure time of the subtitles on this screen?» Percentage of each response option and Chi-Square values, as a function of device.

EXPOSURE TIME	Very long	Long	Appropriate	Short	Very short	$\chi2$ (2, n=30)
Smartphone	0%	6.7%	86.7%	6.7%	0%	38.40, p<.001
Tablet	0%	6.7%	76.7%	16.7%	0%	25.80, p<.001
Monitor	0%	10%	76.7%	13.3%	0%	25.40, p<.001

As can be seen in Table 4, the findings highlight that the majority of the participants found the line-break layout appropriate, in particular for the tablet device (75.9% vs. 56.7% for smartphone and 66.7% for monitor), although some other minor slight tendencies were detected. In fact, 20% of the participants would have preferred longer subtitles in one line for the smartphone device, and 16.7% of the participants would have preferred shorter subtitles in two lines for the monitor device.

Table 4. Question «What do you think about the line-break layout for this device?» Percentage of each response option and Chi-Square values, as a function of device.

LINE BREAKS	Inappropriate	Shorter, in 1 line	Shorter, in 2 lines	Longer, in 1 line	Appropriate	$\chi2$ (2, n=30)
Smartphone	0%	10%	13.3%	20%	56.7%	16.67, p=.001
Tablet	0%	10%	6.9%	6.9%	75.9%	40.10, p<.001
Monitor	0%	10%	16.7%	6.7%	66.7%	28.40, p<.001

ISSN: 2530-609X

Discussion

The experiment in this paper examined the influence of screen size on viewers' subtitle layout preferences and comprehension scores across three devices (monitor, tablet, and smartphone). The main aim was to provide additional data to the two previous studies on watching subtitled content across devices (Castellà et al., 2016; Szarkowska et al., 2015) by analysing subtitle layout parameters that have not been previously studied. Another aim was to validate the comprehension scores by Szarkowska et al. (2015). Drawing on the previous studies on screen size (Al-Showarah et al., 2014; Castellà et al., 2016; Maniar et al., 2008; Szarkowska et al., 2015), we predicted that the smallest screen device would give the most unsatisfactory results. We also expected to see differences in the viewers' reception of subtitle layout parameters across devices.

Regarding comprehension, no differences were found across screens. The findings are in line with the *subtitle effectiveness hypothesis* (Perego, Del Missier, Porta, & Mosconi, 2010), which suggests that viewers can adapt their reading and visual skills for any screen displays. Contrary to the findings on comprehension scores by Szarkowska et al. (2015), our results imply that screen size is not a limitation and does not have a considerable impact on viewers processing subtitles across devices.

The results on the readability items indicate that screen size does not affect the viewers' reception of subtitles across devices in terms of the percentages of subtitles read, the ease of subtitle readability, and the overall experience on each device. Moreover, the majority of the participants declared that they did not have the feeling of having lost essential parts of the film's action due to the fact of reading subtitles. However, we only found significant results for this yes/no question for tablet and monitor screens. This finding shows that viewers feel capable of perceiving the incorporation of subtitles into tablets and monitors, in such a way as not to miss information from the rest of the audiovisual components. As for smartphone screens, results are not significant regarding this yes/no question. Our interpretation is that there is a broader range of opinions for smartphone screens because viewers do not perceive these screens as optimal as other screens. They may not feel as confident reading subtitles on these small screens as on larger devices. This result is consistent with Kim et al. (2011), which suggested that larger screen size devices are the key to greater enjoyment.

Although viewers found subtitles appropriate in terms of line length, exposure time and line-break layout, and the results did not provide significant differences across screens, some tendencies were detected. Regarding the line length of subtitles, tablets seem to provide the highest satisfaction. This is probably because the tablet display offers the right balance between each subtitle line and its medium size, not forcing the eyes to move much, compared with larger screens (i.e. monitor). The exposure time of subtitles on screens was found to be slightly more appropriate for smartphones than tablets or monitors. A possible explanation of this finding may be that smaller screens minimize the tendency to focus on other elements of the scene because of the limited size. On the contrary, the inclusion of subtitles on large screens can divert viewers' attention from the rest of the audiovisual components, as mentioned in some studies (Lombard et al., 1997; Maniar et al., 2008). Data related to line-break layout did not indicate a clear preference for smartphone and monitor screens: slightly more than half of the participants stated that the line-break layout was appropriate, but a minor tendency preferred shorter subtitles of two lines for monitor screens, and another minor trend preferred longer subtitles of one line for smartphones. Tablet screens seem to be the device with the highest satisfaction ratings regarding line-break layout. The results from these

ISSN: 2530-609X

three categorical questions are not consistent with our initial hypothesis about finding differences in the overall evaluation of subtitles depending on the screen size.

CONCLUSIONS

Our study represents the first piece of knowledge on the effects of screen size on subtitle layout parameters, and it validates previous findings on comprehension scores across devices (Szarkowska et al., 2015). Our main finding shows that participants adapt their viewing skills to different screen sizes to process short subtitled film clips, and are generally satisfied with the subtitle layout on the devices tested. We also found that screen size does not affect comprehension levels.

We acknowledge that the general profile in this experiment included university students, the average age was 30 years old, and all participants belonged to a dubbing country. We believe that differences in comprehension and preferences of the subtitles could be found if other user profiles with different technological and audiovisual material habits were tested in the experiment (e.g. children, the deaf or the elderly).

Based on the trends found for some of the subtitle layout parameters, we think that more empirical studies focused on smartphone devices. Participants felt more comfortable reading subtitles in the larger screens (monitor and tablet): they did not have the impression of losing visual information and were more satisfied with tablet screens regarding subtitle line length and line-break layout. Our results for smartphone screens were not conclusive in terms of subtitle layout parameters, and do not validate the comprehension results by Szarkowska et al. (2015). New subtitle experiments on smartphone devices could also validate the results by Castellà et al. (2016), who found a different exploration pattern on viewers reading subtitles on smartphone devices: when reading subtitles in smartphone screens, there are fewer fixations but longer in duration compared to the other devices. Moreover, to our best knowledge, there are no empirical studies on reading subtitles specifically on smartphone screens. Further research could explore different types of line-break layouts on these devices to measure the impact of this variable on comprehension, readability, and enjoyment of audiovisual products.

REFERENCES

Al-Showarah, S., AL-Jawad, N., & Sellahewa, H. (2014). *Effects of User Age on Smartphone and Tablet Use, Measured with an Eye-Tracker via Fixation Duration, Scan-Path Duration, and Saccades Proportion.* In C. Stephanidis & M. Antona (Eds.), UAHCI/HCII, Part II, LNCS 8514, (pp. 3-14). Springer International Publishing Switzerland.

BBC. (2017). *BBC subtitles guidelines*. Retrieved 11 October 2017, from http://bbc.github.io/subtitle-guidelines/#Line-breaks

Castellà, J., Olivier, A., Gerber, O., & Soler, O. (2016, May). *Smartphone, Tablet or PC: Which one is better to watch your favourite subtitled film?*. Paper presented at the International Meeting of the Psychonomic Society, Granada (Spain).

D'Ydewalle, G., Van Rensbergen, J., & Pollet, J. (1987). Reading a message when the same message is available auditorily in another language: The case of subtitling. In J. K. O'Regan, & A. Lévy-Schoen (Eds.), *Eye movements: From psychology to cognition* (pp.

ISSN: 2530-609X

313–321). Amsterdam, the Netherlands: Elsevier Science.

Day, R. R., & Park, J. (2005). Developing reading comprehension questions. *Reading in Foreign Language, 17*(1), 60–73.

Described and Captioned Media Program. (2017). *DCMP Captioning Key. The Described and Captioned Media Program*. Retrieved from http://www.captioningkey.org/

Gambier, Y. (2009). Challenges in research on audiovisual translation. In A. Pym & A. Perekrestenko (Eds.), *Translation Research Projects 2* (pp. 17–25), Tarragona, Spain: Intercultural Studies Group.

Gottlieb, H. (1992). Subtitling. A new university discipline. In C. Dollerup & A. Loddegaard (Eds.), *Teaching translation and interpreting 1. Training, talent and experience. Papers from the first language international conference* (pp. 161–170). Amsterdam, the Netherlands: John Benjamins.

Karamitroglou, F. (1998). A proposed set of subtitling standards in Europe. *Translation Journal, 2*. Retrieved from http://translationjournal.net/journal/04stndrd.htm

Kim, K. J., Sundar, S. S. & Park, E., (2011). The effects of screen-size and communication modality on psychology of mobile device users. In *Proceedings of the 2011 Annual Conference Extended Abstracts on Human Factors in Computing Systems,* Vancouver, Canada, 2011, pp. 1207-1212. doi: 10.1145/1979742.1979749

Lavaur, J.-M., & Bairstow, D. (2011). Languages on the screen: is film comprehension related to the viewers' fluency level and to the language in the subtitles? *International Journal of Psychology : Journal International de Psychologie, 46*(6), 455–462. doi:10.1080/00207594.2011.565343

Lombard, M., Ditton T. B., Grabe, M. E. & Reich, R. D. (1997). The role of screen size in viewer responses to television fare. *Communication Reports 10*(1), 95-106.

Maniar, N., Bennett, E., Hand, S. & Allan, G. (2008). The effect of mobile phone screen size on video based learning. *Journal of Software 3*(4), 51–61. doi:10.4304/jsw.3.4.51-61

Media Access Australia. (2012). CAPTIONING GUIDELINES. Retrieved 10 October 2017, from https://mediaaccess.org.au/practical-web-accessibility/media/caption-guidelines

Messerlin, P., Siwek, S. & Cocq, E. (2005). *The Audiovisual Services Sector in the GATS Negotiations*, Washington DC: American Enterprise Institute Press.

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

Palen, L., Salzman, M. and Youngs, E. (2000) Going wireless: behavior & practice of new mobile phone users. In *Proceedings of CSCW '00,* Philadelphia, PA, USA.

Perego, E. (2005). *La traduzione audiovisiva*, Roma, Italy: Carocci editore.

ISSN: 2530-609X

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

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

Romero-Fresco, P. (Ed.). (2015). *The Reception of Subtitles for the Deaf and Hard of Hearing in Europe.* Bern, Switzerland: Peter Lang AG.

Szarkowska, A., Laskowska, M., Oliver, A., Pilipczuk, O. (2015). *Reading subtitles across devices: A study into the differences in reading patterns of people watching subtitled videos on smartphone, tablet and computer screen*, conference paper presented at the International Conference on Eyetracking and Applied Linguistics, Warsaw, 21-22 September 2015.

ISSN: 2530-609X

Article 4

Gerber-Morón, O. (manuscript submitted for publication). *How to divide lines in a subtitle: an exploratory study on subtitle segmentation approaches in alternative small screens*.

This article was under review at the moment of the submission of this thesis. The study is relevant to this PhD thesis because it combined all the elements studied in the previous objectives, but it could not be included in the main manuscript as it has not yet been accepted. It considered different profiles of users and tested the effects of syntactic segmentation on the device that obtained the most unsatisfactory results. The study conducted for this objective represents the ultimate verification of syntactic segmentation, and adds new insight to subtitle segmentation research by considering the current technological advances and different audiences of subtitles.

It examined the effects of two different segmentation styles on the device that had the most unsatisfactory results in **Study 3** (Gerber-Morón et al., forthcoming). It involved 18 native Swiss-Italian hearing participants, and 17 deaf and hard of hearing viewers from Canton of Ticino (Switzerland). They watched two video fragments of 10 minutes each from the first episode of an American TV series dubbed into Hungarian with Italian subtitles. Each participant was exposed to two subtitle segmentation styles: they would watch one of the fragments with syntactically segmented subtitles and the other with geometrically segmented subtitles. The order of administration of the segmentation conditions was counterbalanced between subjects. After watching each fragment, participants answered comprehension and subtitle recognition questionnaires. At the end of the experiment, participants answered a questionnaire on readability and segmentation preference for the subtitles based on the fragments they watched. The independent variables were subtitle segmentation (syntactically segmented vs. geometrically segmented subtitles) and the degree of hearing loss (hearing, hard of hearing, deaf). Comprehension, subtitle recognition, perceived readability and segmentation preference were the dependent variables.

How to divide lines in a subtitle: an exploratory study on subtitle segmentation approaches in alternative small screens

Abstract

One of the features that is believed to influence how viewers read subtitles is subtitle segmentation, that is the division of text between a two-line subtitle. This study examines the factors influencing subtitle segmentation, and the approaches to divide lines in the production of subtitles. The aim was to assess the impact of subtitle segmentation on subtitle processing and viewer experience. We tested two approaches to subtitle segmentation: syntactic and geometric. The former followed the syntactic rules to keep units of sense in the same subtitle line. The latter followed aesthetic considerations distributing the text in balanced lines. Three groups of participants from different subtitle viewing backgrounds took part in the study: deaf, hard of hearing and hearing viewers. We took into consideration today's technological advances by examining these approaches in smartphone devices. Our findings show that the manipulation of segmentation does not affect subtitle processing or the viewer experience.

Keywords: subtitle segmentation; syntactic segmentation; geometric segmentation; cognitive processing; screen size

Introduction

One the functions of subtitling is to provide viewers with accessibility to audiovisual content to follow dialogues on the screen. Subtitles can be helpful for viewers who do not understand the original language of the film dialogues or people with hearing loss (Gambier, 2006). Subtitles can also become handy to viewers in noisy environments (Skoog, 2016) or to watch videos without the sound (Patel, 2016).

Subtitle features include: layout, synchronization, reading speed, and division of subtitle lines (Díaz Cintas & Remael, 2007; Gambier, 2006; Ivarsson & Carroll, 1998). The way these features are set up can influence how viewers read subtitles.

For instance, viewers might not be able to read subtitles if they are displayed too fast, the font is very small or the number of lines obscure important picture information. Among these features, this paper focuses on subtitle segmentation, i.e. the way text is divided across lines. There are two levels of subtitle segmentation (Díaz Cintas & Remael, 2007): a text may be divided between the two lines of the same subtitle (henceforth referred to as intrasegmentation), or it may be distributed across two or more subtitles (henceforth referred to as intersegmentation). In our paper, we focus on intrasegmentation.

There are different approaches to divide lines in a subtitle when the text does not fit into a single line. One approach is to divide the text keeping units of sense in the same line:

Peter went home
and told his family about Portugal.

In this example, the units of sense "Peter went home" and and "told his family [...]" are not split. Lines can also be divided according to geometry:

Peter went home and told
his family about Portugal.

In this instance, the upper and lower lines are more balanced, resulting in a rectangular shape.

The approach to segmentation affects the composition of subtitles. For instance, the text might have to be reduced, paraphrased or adapted to keep units of sense in the same line or to obtain more geometrically balanced lines. This article examines the impact of subtitle segmentation on cognitive processing and viewer experience. The first part of the article is a literature review on the external factors influencing subtitle segmentation, the approaches to divide lines in the production of subtitles, the empirical studies on the topic and the subtitle audiences. It then moves to the experimental part of the study, where two approaches to subtitle segmentation are tested on smartphone screens. The article concludes with the results of the tests and considerations for future research on subtitle segmentation.

External factors in subtitle segmentation

A number of external factors are involved in the distribution of subtitled audiovisual content. These affect how subtitles are produced and how subtitle segmentation is performed. The first influencing factor are standards or norms on television access services drafted by media regulators. Subtitle lines should be broken at grammatical breaks or punctuation marks, as well as when there is a pause in the dialogues or a silence (AENOR, 2012). Line breaks in subtitles should end "at natural linguistic breaks" (Broadcasting Authority of Ireland, 2012, p. 3; Ofcom, 2017, p. 21), "ideally at clause or phrase boundaries" (Broadcasting Authority of Ireland, 2012, p. 3), forming "an understandable segment" (Ofcom, 2017, p. 21). However, when lines are split over more than one subtitle, (Ofcom, 2017) recommends ending the first subtitle with a conjunction to indicate that there is more text to come, which goes against the linguistic breaks previously mentioned.

The second factor are broadcaster guidelines such as the BBC (BBC, 2017). These guidelines based on Ofcom (2017) develop further the two Ofcom recommendations with a comprehensive list of parts of speech that should not be split across lines, such as article and noun (e.g. the + table; a + book) and pronoun and verb (e.g. he + is; they + will come; it + comes), among others (BBC, 2017).

Although BBC subtitle guidelines stress the importance to follow syntactic rules to break lines, they also recommend finding a balance between linguistic and geometric considerations. Line breaks in justified subtitles should occur at linguistically coherent points, taking eye-movement into account. However, the example provided by the BBC on line breaks in centre justified subtitles does not follow the syntactic rules included in their guidelines, breaking parts of a complex verb (are + feeling):

We all hope you <u>are</u>
<u>feeling</u> much better.

(BBC 2017, para. 59)

BBC guidelines include contradictory recommendations regarding syntactic or geometric rules for line breaks. Neither the media regulators nor BBC subtitle guidelines explain the basis for their recommendations.

Another factor is the time available for preparing subtitles: offline subtitles or live subtitles (Díaz Cintas & Remael, 2007). Segmentation for offline subtitling depends on the instructions given by the audiovisual companies or the client, as well as on the subtitling editor used to subtitle (Díaz Cintas & Remael, 2007; Gambier, 2006). Live subtitling relies on the technology used to produce subtitles, and uses: machine translation systems, speech recognition software, velotype or stenotype. In machine translation systems, the implementation of syntactic automatic segmentation requires language processing tools for text analysis, which cannot always ensure a correct syntactic analysis (Álvarez, Arzelus, & Etchegoyhen, 2014). In live subtitles produced by speech recognition software, segmentation can be word-for-word display mode or in chunks corresponding to full sentences (Ribas Arumí & Romero-Fresco, 2008).

The linguistic parameters considered to produce subtitles are another factor that determines subtitle segmentation. Subtitles can be classified in three categories depending on the linguistic needs of end-users: intralingual, interlingual, and bilingual subtitles (Díaz Cintas & Remael, 2007). Intralingual subtitles are in the same language as the original soundtrack (e.g. English to English) and can be classified in different types: for deaf (and deaf children) and hard-of-hearing viewers, for language learning purposes, for dialects of the same language, for karaoke, and for advertising and broadcasting news. Interlingual subtitles are the translation of the original soundtrack dialogues into another language (e.g. French into English), and there are two types: for hearing people, and for deaf and hard of hearing people. Finally, bilingual subtitles are the translation of the original soundtrack into two different languages, and they are usually used in bilingual regions (e.g. English into French and German for Swiss viewers) or at international cinema festivals. The choice of these linguistic parameters has an impact on the division of lines in the subtitles. For instance, subtitles for language learning purposes include a word-for-word transcription of the original soundtrack and the text can run across two or three line subtitles. Syntactic segmentation for this type of subtitles can be more difficult to achieve, as it might be hard to keep units of sense in the same line without altering the transcription of the dialogues. Subtitles for deaf children prioritize syntactic segmentation, as provisions of television access services establish that this type of subtitles need to follow grammatical rules and have a simple syntax (DCMP, 2017; de Linde & Kay, 1999a; Ofcom, 2006; Tamayo, 2016).

Media distribution is another factor, and how subtitles are displayed changes greatly from: cinema, television, DVD or Internet (Díaz Cintas & Remael, 2007). The specific technical constraints of each medium, such as the size of the screen and the line length for each device, can affect line breaks in a subtitle. For instance, entertainment companies and broadcasters offer nowadays the option to watch content on multiple devices with different screen size (e.g. computer, tablet or smartphone)[11].

Approaches in subtitle segmentation

Scholars in audiovisual translation studies have issued guidelines on how to divide lines in subtitles, recommending three approaches to segmentation: syntactic, geometric, and rhetorical. As with regulators and broadcasters, academic guidelines are based on observation rather than research (Carroll & Ivarsson, 1998; Díaz Cintas & Remael, 2007; Ivarsson & Carroll, 1998; Karamitroglou, 1998; Luyken et al., 1991) and need to be updated to the technology available in the current information society (Baker et al., 1984).

Syntactic segmentation

Scholars encourage the use of syntactic rules to optimise reading speed (Baker et al., 1984; Díaz Cintas & Remael, 2007; Ivarsson & Carroll, 1998; Karamitroglou, 1998). Baker et al. (1984) were the first to recommend ending lines at "natural linguistic breaks" (p.12), that is at clause and phrase boundaries. Ivarsson & Carroll

11 Such as Amazon Video, BBC iPlayer and Netflix.

(1998) follow the same line of thought by explaining that subtitle segmentation should follow grammatical rules, keeping words that are intimately connected on the same line, and breaking lines at the beginning and end of phrases or clauses. They also recommend to distribute subtitle text "in sense blocks and/or grammatical units" (Carroll & Ivarsson, 1998, p. 1). For instance, rather than breaking lines in the following way:

I was very sad. She
tried to comfort me.

a syntactically-segmented subtitle would keep the pronoun and the verb in the same line:

I was very sad.
She tried to comfort me.

Karamitroglou (1998)introduces the notion of dividing lines at the highest syntactic nodes possible, in order to present the most complete linguistic information in each line to be processed by the brain. Díaz Cintas & Remael (2007) mention the same recommendations as Baker et al. (1984), Carroll & Ivarsson (1998), Ivarsson & Carroll (1998), Karamitroglou (1998), and they include a list of linguistic units that should be kept on the same line, similar to the list in BBC subtitle guidelines (BBC, 2017).

Rhetorical segmentation

This approach to subtitle segmentation takes the dramatic-prosodic perspective and aims to reflect the way characters speak (Díaz Cintas & Remael, 2007). Synchronising through line breaks to reproduce spoken language features (e.g. hesitations, pauses, rhythm, repetitions) "should respect the natural sentence structure" (Luyken et al., 1991, p. 47). In rhetorical segmentation, prosodic features of speech can be reflected by breaking sentences into more subtitles - usually one-line subtitles - that follow the rhythm of the speaker. Rather than having just one subtitle:

I don't know. Maybe you should wait
for something to happen.

A rhetorically-segmented subtitle would indicate hesitation by creating two subtitles:

I don't know.
Maybe you should wait...

for something to happen.

This type of segmentation is more related to intersegmentation and tends to overlap with syntactic segmentation, as "linguistic and paralinguistic features of speech usually collaborate" (Díaz Cintas & Remael, 2007, p. 179).

Geometric segmentation

The geometrical shape that a two-line subtitle should have is also considered from a descriptive approach to avoid "untidy formats" (Baker et al., 1984, p. 13). Lines should have lines approximately equal in length, and when it is not possible the upper line should be shorter than the lower "to keep as much of the image as free as possible" (Carroll & Ivarsson, 1998, p. 2). Geometric shape is recommended because viewers are used to reading printed text in a rectangular format (Baker et al., 1984; Karamitroglou, 1998). For example, a geometric subtitle would keep the upper and lower line of equal length:

I was scared and
I looked for help.

However, the conjunction 'and' is split from the following clause 'I looked for help', In our study, we focus on two of these three approaches to subtitle segmentation: syntactic and geometric. We do not include rhetorical segmentation, as it is closer to syntactic segmentation and it is more applied to intersegmentation.

Empirical studies in subtitle segmentation

Recent empirical studies on subtitle segmentation with eye tracking technology show no conclusive results supporting syntactic segmentation. Perego et al. (2010) showed that participants were able to process syntactically and non-syntactically segmented subtitles in the same way. Rajendran et al. (2013) found in their study on live subtitling segmentation that subtitles segmented by phrases were the easiest to process and implied fewer saccadic crossovers, as opposed to those subtitles segmented word-for-word. They did not find significant differences in preference or comprehension across the subtitling styles tested. Similarly Gerber-Morón et al. (2018) did not find differences in comprehension between syntactically and non-syntactically segmented subtitles, but the latter induced higher cognitive load. Another study by the same authors showed that viewers largely prefer syntactically-segmented subtitles (Gerber-Morón et al., 2018).

Empirical studies on screen size

The question arises whether subtitle segmentation should be adapted to the device's screen size to enhance the viewer experience and subtitle cognitive processing. To our best knowledge, there are studies on reading subtitles across devices (monitor, tablet and smartphone), but on the influence of screen size on subtitle segmentation. Gerber-Morón et al. (forthcoming) showed that viewers do not feel as confident reading subtitles on smartphone screens as on larger screens (monitor and tablet) and prefer longer one-line subtitles for smartphones. Szarkowska et al. (2015) found evidence that comprehension was lower in smartphone screens, with longer mean fixation durations and fewer fixations in comparison to tablet and monitor screens. Another similar eye tracking study on watching subtitled videos on different screen devices (Castellà et al., 2016) showed that smartphone devices require more cognitive load when reading subtitles than larger screens.

Audiences of subtitles

Research on subtitle segmentation should consider the different categories of subtitle audiences. There are several ways towards audience classification: by the function of subtitles, by the age or by the medical perspective. In the literature on subtitle reading, the approach taken to audience classification is by the function of subtitles and the expected end-user (Braun & Orero, 2010; Caimi, 2006; Nobili, 1995; Zárate, 2010). Subtitles can have a communicative function for audiences who require the translation of the original soundtrack of films (Díaz Cintas & Remael, 2007; Perego, 2005). Subtitles can also have a sensorial function for audiences who need the aural information to access to the audiovisual content: deaf and hard of hearing viewers who need the written text of the dialogues, paralinguistic information (e.g. tone, inflexion), extralinguistic sound effects and other audible cues relevant to the film (Caimi, 2006; Zárate, 2010). Likewise, subtitles can have a therapeutic function for people with language impairments (e.g. aphasia, dyslexia) or cognitive impairment (e.g. retardation, decreased concentration) (Braun & Orero, 2010; Porteiro, 2013). Subtitles can also have a learning function for end-users who want to improve their comprehension and listening skills in a foreign language they are learning (Nobili, 1995).

Audiences of subtitles can also be classified by audience age. Subtitles can be produced for children to acquire vocabulary in a foreign language (Koolstra & Beentjes, 1999), or for deaf children (de Linde & Kay, 1999a; Tamayo, 2016). Subtitles can also be produced for the elderly, by applying longer presentation times and bearing in mind legibility parameters such as black box background and larger character size (Neve & Jenniskens, 1994).

Another classification of audiences is based on the medical point of view and the linguistic perspective of the viewer. Báez Montero & Fernández Soneira (2010) classify audience by: hearing, deaf, or hard of hearing. The cognitive processing of deaf and hard of hearing viewers varies according to their mother tongue, their education and residual hearing (de Linde, 1996). Most of the hard of hearing people receive the same education as hearing people, so their reading level is high (de Linde, 1996). On the contrary, sign language is the native language for prelingual deaf people, and their understanding of subtitles can be hampered by their poor standards in reading (Albertini & Mayer, 2011; Conrad, 1979; Karchmer & Mitchell, 2003; Quigley & Paul, 1984; Savage et al., 1981; Schirmer & McGough, 2005). Deaf viewers lack specific language knowledge (semantic and syntactic) and oral skills that are necessary for reading (Monreal & Hernandez, 2005). For many deaf viewers, subtitles are displayed in a language they do not speak, as their mother tongue is sign language (de Linde & Kay, 1999a). In our study, we focus on the communicative function of subtitles, by analysing subtitle segmentation with deaf, hard of hearing and hearing viewers who need the translation of the original soundtrack to access to the audiovisual content.

Overview of the study

The study adopts a new focus on subtitle segmentation by taking into consideration two approaches to segmentation (syntactic and geometric), the expansion of small screen devices and the audiences of subtitles (hearing, hard of hearing and deaf viewers). Two styles of line breaks were contrasted to assess the impact of segmentation on comprehension and preferences of subtitles on smartphone screens. One of the styles, "syntactically segmented subtitles", followed the syntactic rules to keep units of sense together in the same subtitle line (Díaz Cintas & Remael, 2007; Ivarsson & Carroll, 1998; Karamitroglou, 1998; Perego, 2008b). The other style of segmentation, "geometrically segmented subtitles", followed aesthetic considerations to keep lines equal in length (Baker et al., 1984; Karamitroglou, 1998). We tested three groups of participants from different subtitle viewing backgrounds: deaf, hard of hearing and hearing viewers from the Italian speaking part of Switzerland. We hypothesized that geometrically segmented subtitles would have a negative effect on subtitling processing, attributed partly to the difficulty of reading text in a small screen and having units of sense split across lines. We predicted that participants would prefer syntactically segmented subtitles because they perceive them as more readable than geometrically segmented subtitles. We also expected that there would be differences between the profiles, with lower scores among participants with hearing loss, as processing written text is a more demanding task for deaf and hard of hearing viewers (de Linde & Kay, 1999b; Neves, 2005).

Methods

Participants

Thirty-five participants from Canton of Ticino in Switzerland took part in the study. Nineteen participants were female (54.3%), and sixteen were male (45.7%), with a mean age of 44 years old (*SD*=15.01). Seventeen participants of the total sample were deaf or hard of hearing participants (48.6%) recruited through two Swiss associations for deaf and hard of hearing people (52.9% females, 47.1% males; mean age = 47.35; *SD*=16.41): ATIDU[12] (Associazione per persone con problemi d'udito) and SGB-FSS[13] (Federazione Svizzera dei Sordi SGB-FSS). Of this group of participants, five were profoundly deaf, five were severely deaf, and seven had a moderate hearing loss. Eighteen hearing participants (51.4%) of the total sample formed the control group based on matching the characteristics of the deaf and hard of hearing groups (i.e. age and gender).

They were all Italian native speakers with normal or corrected-to-normal (contact lenses or glasses) vision. They were recruited from personal networking. None of the participants had any knowledge of the original language used for the film fragments (Hungarian). Participants were asked about their experience watching subtitled material: 50% of hard of hearing and 80% of deaf participants declared to watch subtitled material very often or always, compared to 5.6% in the hearing participant group.

Materials

12 http://www.atidu.ch/

13 http://www.sgb-fss.ch/it

Videos. These comprised two video fragments of 10 minutes each from the first episode of *Mad Men*, a 2007 American TV series created by Matthew Weiner and produced by Lionsgate Television[14]. The fragments were shown in their dubbed version in Hungarian with Italian subtitles. Each video fragment formed complete scenes with coherent content. Both fragments were fast-paced (between 13 to 20 camera changes per minute) and had a complex narrative structure (cf. Lang, Bolls, Potter, and Kawahara 1999; Lang, Zhou, Schwartz, Bolls, and Potter 2000): it included many characters interacting (between 9 and 11 on average per fragment) and several interweaving story lines.

Subtitles and apparatus. The subtitles were created using the freeware Subtitle Workshop, a free application for creating, editing, and converting text-based subtitle files. The font type and size used were Arial 32, with black contoured characters, as recommended by subtitle guidelines (Díaz Cintas & Remael, 2007; Karamitroglou, 1998). The reading speed chosen was of 180 words per minute[15], which is considered a high-speed rate by Díaz Cintas & Remael (2007). We created 159 subtitles for the first video fragment (99 one-liners, 60 two-liners) and 184 subtitles for the second fragment (126 one-liners, 58 two-liners), and embedded them in the Hungarian dubbed version of the videos. 27 two-line subtitles were manipulated to test subtitle segmentation for the first fragment and 33 for the second fragment.

For each video fragment, two subtitle versions were created to test line breaks on smartphones: syntactically (Figure 1) and geometrically (Figure 2) segmented versions.

14 http://www.imdb.com/title/tt0804503/

15 The equivalence is calculated between seconds/frames and spaces to estimate the approximate number of characters that can be used in a subtitle. Seventeen subtitling spaces per second and lines of 39 spaces were used to create subtitles. For further information regarding reading speed see Díaz Cintas and Remael (2007).

Figure 1. Syntactically segmented subtitles with the noun phrase "buoni sconto" (discount tickets) is not split

Figure 2. Geometrically segmented subtitles with the noun phrase "buoni sconto" split

The videos were uploaded on Vimeo[16], a video-sharing website, and sent by Wi-Fi to the device tested in the experiment: a 3.5-inch iPhone 4. Subtitles were previously burnt on each video to avoid problems with the Wi-Fi connection and synchronization of images with the written text.

Questionnaires

Comprehension questionnaire. The comprehension questionnaires included a set of multiple-choice questions to verify whether participants understood the main textual information conveyed by the subtitles. After watching each video fragment, participants had to answer a set of ten questions about the content of the fragment. For each question, they were asked to complete a statement or reply to a question by choosing a response from a list of three items including a correct answer and two

16 https://vimeo.com

distractors (e.g. "What is Greta Gutman's position in Sterling Cooper's company? Potential answers: "graphic designer", "creative director", "researcher"). The design and procedure of the questionnaire were based on the studies undertaken by Day & Park (2005), Lavaur & Bairstow (2011) and Leung (2001).

Subtitle recognition questionnaire. The subtitle recognition questionnaire was a set of multiple-choice questions to assess the ability to identify specific words or short phrases shown in the subtitles (Koolstra & Beentjes, 1999). Participants had to answer a set of ten questions about the exact wording used in the subtitles whose segmentation was manipulated. For each question, they were asked to complete a statement by choosing the exact wording from a list of three items including a correct answer and two distractors. The distractors were semantically related to the correct answer (e.g. "Midge explains to Don Draper that in the next months she will be busy drawing..." Potential answers: "puppies", "kittens", "wolf cubs"). The items selected were words or short noun phrases that appeared at the beginning, in the middle and at the end of the two-line subtitles. The reason for this choice was to control for recency and primacy memory effects, which refer to an advantage in recalling items presented at the beginning and at the end of a list (Murdock, 1962).

Questionnaire on readability and segmentation preference for subtitles. The questionnaire on readability and segmentation preference was administered to verify participants' perception of subtitle segmentation and ease of reading between the first and the second video fragment. In the first question, participants were asked to answer on a 7-point Likert scale whether they noticed any differences in the subtitle layout between the first and the second video fragment. In the second question, participants were asked to answer which video fragment was more easily read, having three possible options to select: "the first fragment", "the second fragment" or "I didn't notice any difference".

The questionnaire on segmentation preference included two series of pictures from the video fragments of the experiment. The pictures of each series were identical, except for the subtitle segmentation (geometrically and syntactically segmented) (see Figure 5.1 and Figure 5.2). Participants were asked to choose the subtitle segmentation that they preferred for each series.

Design and Procedure

Two video fragments with subtitles displayed on a 3.5-inch iPhone 4 were presented to participants. Segmentation format was manipulated within subjects, that is, each participant watched two different fragments, one with syntactically segmented subtitles and the other with geometrically segmented subtitles. The order of administration of the two subtitling conditions was counterbalanced between subjects to avoid practice or learning effects. Half of the participants were exposed to the syntactic-geometrical condition (syntactic video 1 - geometric video 2) and the other half to the geometric-syntactic condition (geometric video 1 - syntactic video 2). Subtitle segmentation and the degree of hearing loss was the independent variable tested in the experiment, whereas the dependent variables were comprehension, subtitle recognition, perceived readability and segmentation preference.

Each participant was tested individually in a laboratory. Participants were seated at a distance of 60 cm from the screen. After signing a consent form, they were given instructions on the experiment: “You will watch on a smartphone two video fragments of 10-minute each in an unknown language with Italian subtitles. Your task is to try to understand the plot of the video. After each fragment, you will answer a questionnaire related to the subtitled video.”

Participants watched the first video fragment in one of the two conditions (syntactically or geometrically segmented), and answered the set of comprehension and subtitle recognition questionnaires related to it. They were asked to watch the second video fragment in the other condition and answer the questionnaires in the same way as they did for the first fragment. After completing these questionnaires, participants were asked to answer the questionnaire on readability and segmentation preference for subtitles. Finally, they filled the demographic and control variable questionnaire. The experiment lasted approximately 40 minutes.

Results

Comprehension

A 2 x 2 mixed ANOVA was performed on segmentation (syntactically segmented and geometrically segmented subtitles) and degree of hearing loss (hearing, hard of hearing and deaf). The dependent variable was mean comprehension score (with the maximum score being 10).

There was no main effect of segmentation on comprehension ($F(1,32)=.2$, $p=.66$, $\eta_p^2=.01$). Table 1 shows descriptive statistics for this analysis. There were no interactions and no main effect of deafness in comprehension was found.

Table 1 Descriptive statistics for comprehension

	Deafness	Mean	Std. Deviation	N
Comprehension syntactically segmented subtitles	Hearing	7.44	2.007	18
	Hard of hearing	7.92	1.730	12
	Deaf	5.20	3.114	5
	Total	7.29	2.217	35
Comprehension geometrically segmented subtitles	Hearing	6.83	1.855	18
	Hard of hearing	7.17	1.749	12
	Deaf	6.00	1.871	5
	Total	6.83	1.807	35

Subtitle recognition

A 2 x 2 mixed ANOVA was performed on segmentation (syntactically segmented and geometrically segmented subtitles) and degree of hearing loss (hearing, hard of hearing and deaf). The dependent variable was mean subtitle recognition score (with the maximum score being 10).

Similarly to comprehension, no main effect of segmentation on subtitle recognition was found ($F(1,32)=.82$, $p=.37$, $\eta_p^2=.02$). No interactions and no main effect of deafness in subtitle recognition were found. Table 2 shows descriptive statistics for this analysis.

Table 2 Descriptive statistics for subtitle recognition

	Deafness	Mean	Std. Deviation	N
Subtitle recognition syntactically segmented subtitles	Hearing	7.11	1.811	18
	Hard of hearing	7.08	1.832	12
	Deaf	6.80	2.683	5
	Total	7.06	1.893	35
Subtitle recognition geometrically segmented subtitles	Hearing	7.17	1.295	18
	Hard of hearing	6.33	1.557	12
	Deaf	6.40	2.074	5
	Total	6.77	1.516	35

Readability and segmentation preference for subtitles

A likelihood-ratio chi-square test was performed to assess whether participants in the three groups - hearing, hard of hearing and deaf - noticed differences in segmentation between the syntactically and geometrically segmented subtitles in the two video excerpts. The results did not reveal any statistically significant association on the percentage of participants' responses (x^2(10, *n*=35)=13.89, *p*=.18) (see Table 3 for details on percentage).

Table 3. Question "Did you notice any difference in the subtitle layout between the first and the second clip?", percentage of each response option values for readability

Degree of deafness	A lot	Quite many	Quite a few	Some	Almost none	None
Hearing	0%	22.2%	22.2%	5.6%	11.1%	38.9%
Hard of hearing	8.3%	8.3%	0%	25%	25%	33.3%
Deaf	0%	0%	20%	0%	40%	40%

Likelihood-ratio chi-square tests were conducted to evaluate if participants read more easily subtitles in one of the two clips of the experiment. Results showed a significant association in the syntactic-geometrical condition, (x^2(4, *n*=35)=11.04, *p*=.03, and a certain trend toward significance in the geometrical-syntactic condition, $x^2(4, n=35)=8.35$, $p=.08$ (see Table 4 for details on percentage).

Table 4 Question "Which clip was easier to read?", percentage of each response option values for segmentation preference

	Syntactic-geometrical condition			Geometrical-syntactic condition		
Degree of deafness	First clip	Second clip	No difference	First clip	Second clip	No difference
Hearing	11.1%	44.4%	44.4%	44.4%	33.3%	22.2%
Hard of hearing	80%	20%	0%	0%	28.6%	71.4%
Deaf	0%	33.3%	66.7%	50%	0%	50%

A likelihood-ratio chi-square test was calculated comparing the segmentation preference (syntactically and geometrically segmented subtitles) in hearing, hard of hearing and deaf participants. No significant difference was found between degree of deafness and preferred segmentation ($x^2(2, n=35) = 2.89$, $p=.23$) (see Table 5 for details on percentage).

Table 5 Percentage of each response option values for segmentation preference

Degree of deafness	Syntactic preference	Geometrical preference
Hearing	66.7%	33.3%
Hard of hearing	41.7%	58.3%
Deaf	80%	20%

Discussion

The experiment reported in this paper investigated the effects of subtitle segmentation on subtitling cognitive processing and on the viewer experience when watching audiovisual products. It examined two approaches to subtitle segmentation (syntactic and geometric) with three audiences of subtitles (hearing, hard of hearing, and deaf viewers) in small screen devices (smartphones). We predicted that geometrically segmented segmentation would have a detrimental effect on subtitling processing, explained by the units of sense split across lines (Baker et al., 1984; Díaz Cintas & Remael, 2007; Ivarsson & Carroll, 1998; Karamitroglou, 1998) and the additional difficulty of reading text in a small screen (Castellà et al., 2016; Szarkowska et al., 2015). We also hypothesized that participants would prefer syntactically segmented subtitles and perceived them as more readable than geometrically segmented subtitles. We expected to have lower scores among participants with hearing loss, as they have more difficulties processing written text than hearing viewers (Monreal & Hernandez, 2005).

Contrary to our expectations, we did not find evidence that geometrically segmented subtitles result in lower subtitle processing. The different approaches to segmentation do not affect the cognitive subtitling processing: comprehension scores did not differ significantly in syntactically or geometrically segmented subtitles, with very similar scores among the three profiles of viewers tested in both conditions, even on small screens. The results are in line with previous studies on subtitle segmentation (Gerber-Morón et al., 2018; Perego et al., 2010), which found that participants were able to process syntactically and non-syntactically segmented subtitles in the same way. Similarly, the results on subtitle segmentation show that participants can process and recall words in subtitles regardless of the approach to segmentation, and confirm the findings on subtitle recognition by Perego et al. (2010). Our results on comprehension and subtitle recognition are consistent with the subtitle effectiveness hypothesis (Perego et al., 2010), in that viewers are able to cope well with different types of subtitles because "reading is a well-learned and partly automatized skill that will not be seriously affected by a modest increase in text integration demands" (Perego et al., 2010, p. 249). Viewers can easily adapt their reading strategies according to the subtitle segmentation approach they are exposed to. This interpretation is also in line with Mitchell (1989), who noted that readers use non-lexical cues (e.g. punctuation, subtitle layout) to process syntactic structure of sentences in reading, but these cues are of secondary importance when compared to words. In contrast to what was suggested by Karamitroglou, (1998), viewers might find it useful to read units of sense in the same subtitle line, but they are neither a determinant factor in cognitive processing, nor an element that improves comprehension. Another factor that might explain why syntactically and geometrically segmented subtitles are processed in the same way is that some linguistic units take longer to read in syntactically segmented subtitles, as demonstrated by Gerber-Morón & Szarkowska (2018).They found that participants spent more time reading syntactically segmented text in noun phrases (e.g. compound nouns, definite/indefinite article + noun) and less time reading non-syntactically text in verb phrases (e.g. auxiliary verb + participle). They attributed these reading patterns to the "clause wrap-up effect" (Just & Carpenter, 1980; Rayner et al., 2000), arguing that people process the information from the clause at the end and consequently look longer at noun phrases when they are at the end of the first subtitle line (in syntactically segmented subtitles). In our experiment, this could explain why the different approaches to subtitle segmentation do not affect comprehension or subtitle recognition, and why scores are similar in syntactically and geometrically segmented subtitles: some linguistic units were processed faster when they were split (in geometrically segmented subtitles) and others when they were kept together (in syntactically segmented subtitles). Each of these approaches to subtitle segmentation provide advantages to read and process subtitles.

In contrast to what we predicted, deaf participants did not significant score lower in the comprehension and subtitle recognition questionnaires. Although previous studies showed that deaf people have poor reading standards and lack of semantic and syntactic language knowledge (Albertini & Mayer, 2011; Conrad, 1979; Karchmer & Mitchell, 2003; Monreal & Hernandez, 2005; Quigley & Paul, 1984; Savage et al., 1981; Schirmer & McGough, 2005), our findings indicate that deaf process syntactically and geometrically segmented subtitles in the same way as the rest of the participants. These findings corroborate the study by Gerber-Morón et al. (2018), which also tested deaf participants and did not find differences in comprehension between participants. Following the *subtitle effectiveness hypothesis* (Perego et al., 2010), deaf viewers seem to be able to adapt their reading strategies to different approaches to subtitle segmentation.

Our prediction of readability on segmentation was not upheld: participants in the experiment did not notice differences in subtitle layout in either of the two conditions (syntactically and geometrically segmented), irrespective of the viewers' profile. Our results indicate that participants generally do not pay attention to syntax or shape in subtitles. This is in contradiction with previous recommendations on geometric subtitles by scholars (Baker et al., 1984; Karamitroglou, 1998), which claimed that viewers prefer lines equal in length.

As for the question on which subtitle segmentation approach was easier to read, we found significant association in the syntactic-geometrical condition, but only a trend in the geometric-syntactic condition, probably because of the small sample. The results suggest hard of hearing participants found easier to read syntactically segmented subtitles in the syntactic-geometrical condition. This could be due to fact that they are experienced with subtitles (80% of them they declared to be habitual users of subtitles) and they found more difficulties reading the geometrically segmented subtitles in the second fragment, which induced higher cognitive load (Gerber-Morón et al., 2018), after being exposed first to syntactically segmented subtitles in the first fragment. However, the majority of the hard of hearing participants did not find any of the subtitle segmentation approaches easier to read in the geometrical-syntactic condition. These differences could be explained in part by the fact that order of segmentation exposure affects the participant's perception of ease of reading subtitles. The majority of deaf participants did not find any of the subtitle segmentation approaches easier to read. This could be attributed to their lack of language knowledge (Monreal & Hernandez, 2005) and consequently none of the subtitle segmentation approaches enhance their reading strategies. As for hearing participants, the results do not show a clear perception of ease of reading in any of the conditions. Their lack of experience with subtitles (only 5.6% declared watching often or always subtitles) might explain why hearing participants, more accustomed to dubbing, perceive the same ease of reading for any of the subtitle segmentation approaches. This result is in line with the study conducted by Perego, Laskowska, et al. (2016), who found that subtitling can hinder viewers accustomed to dubbing from fully processing film images.

The results on preferences for subtitle segmentation did not show any significant difference: participants from the three groups did not have a clear preference for syntactically or geometrically segmented subtitles. This finding goes against the results obtained in the study by Gerber-Morón & Szarkowska (2018), which showed that viewers largely prefer syntactically-segmented subtitles. Differences in the results of the two studies could be attributed to the design of each experiment. Gerber-Morón & Szarkowska (2018) used static text of subtitles rather than dynamic subtitles, so participants' reading pattern was self-paced. Our study gives new evidence with regards to segmentation preferences: participants do not have a clear preference for segmentation approach after being exposed to real-life subtitling for a considerable amount of time (20 minutes in total including both conditions).

Further studies, which include eye tracking technology, will need to be undertaken to investigate possible differences in reading patterns across the three profiles of users in small screens. Eye tracking data could provide information about differences in the cognitive subtitling processing or the viewer experience. Future studies should examine a larger sample of participants to provide results with more solid data.

Conclusions

We believe that this study contributes to our understanding of processing subtitling and viewer experienced in small screens. It provides updated evidence on the effects of subtitle segmentation taking into consideration the current technological advances and different audiences of subtitles. Our findings show that different approaches to segmentation do not affect subtitling processing, regardless of the profile of viewer. Viewers do not have a clear preference for syntactically or geometrically segmented subtitles, and are able to adapt their reading and viewings strategies to these two approaches to subtitle segmentation. These findings suggest that syntactic segmentation does not determine cognitive processing and some flexibility may be allowed when creating subtitles and distributing the text over a two-line subtitle. It would be worth studying segmentation for specific linguistic units (e.g. verb and complement) in small screens, in order to understand if some categories are more separable than others. Further research with a larger sample of participants should be carried out to provide more evidence about the relevance of subtitle segmentation. More viewers with different linguistic backgrounds —coming from dubbing or subtitling countries— should be included in future studies. Research in subtitle segmentation is also needed to estimate cognitive load (i.e. difficulty, frustration and effort) in small screens, with the implementation of eye tracking technology. Other lines of research could investigate rhetorical segmentation, as well as intersegmentation.

References

AENOR. (2012). *Norma UNE 153010. Subtitulado para personas sordas y personas con discapacidad auditiva.* Madrid.

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

Álvarez, A., Arzelus, H., & Etchegoyhen, T. (2014). Towards Customized Automatic Segmentation of Subtitles. In J. L. Navarro Mesa, A. Ortega, A. Teixeira, E. Hernández Pérez, P. Quintana Morales, A. García, Ravelo, … I. Guerra Moreno (Eds.), *Advances in Speech and Language Technologies for Iberian Languages. Lecture Notes in Computer Science.*, *8854*, 1-10. Springer, Cham.

Báez Montero, I., & Fernández Soneira, A. M. (2010). Spanish deaf people as recipients of closed captioning. In A. Matamala & P. Orero (Eds.), *Listening to Subtitles. Subtitles for the Deaf and Hard of Hearing* (pp. 25–44). Bern, Switzerland: Peter Lang.

Baker, R., 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/

Braun, S., & Orero, P. (2010). Audio description with audio subtitling – an emergent modality of audiovisual localisation. *Perspectives*, *18*(3), 173–188. https://doi.org/10.1080/0907676x.2010.485687

Broadcasting Authority of Ireland. (2012). *BAI Guidelines Subtitling*. Retrieved from http://www.bai.ie/en/codes-standards

Caimi, A. (2006). Audiovisual Translation and Language Learning: The Promotion of Intralingual Subtitles. *The Journal of Specialised Translation*, *6*, 85–98.

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

Castellà, J., Olivier, A., Gerber-Morón, O., & Soler-Vilageliu, O. (2016). Smartphone, Tablet or PC: Which one is better to watch your favourite subtitled film? In *International Meeting of the Psychonomic Society*. Granada (Spain).

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

Day, R. R., & Park, J. (2005). Developing reading comprehension questions. *Reading in Foreign Language*, *17*(1), 60–73. https://doi.org/10.1080/01443410.2014.910979

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.

de Linde, Z., & Kay, N. (1999). Processing Subtitles and Film Images. *The Translator*, *5*(1), 45–60. https://doi.org/10.1080/13556509.1999.10799033

de Linde, Z., & Kay, N. (1999). *The Semiotics of Subtitling*. Manchester, UK: St. Jerome.

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

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., Soler-Vilageliu, O., & Castellà, J. (forthcoming). The effects of screen size on subtitle layout preferences and comprehension across devices. *Hermēneus. Revista de Traducción e Interpretación*, *21*.

Gerber-Morón, O., & Szarkowska, A. (2018). Line breaks in subtitling: an eye tracking study on viewer preferences. *Journal of Eye Movement Research*, *11*(3), 1–22. https://doi.org/10.16910/jemr.11.3.2

Gerber-Morón, O., Szarkowska, A., & Woll, B. (2018). The impact of text segmentation on subtitle reading. *Journal of Eye Movement Research*, *11*(4), 1–18. DOI: http://dx.doi.org/10.16910/11.4.2

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, US: Oxford University Press.

Koolstra, C. M., & Beentjes, J. (1999). Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home. *Educational Technology Research and Development*, *47*(1), 51–60. https://doi.org/10.1007/BF02299476

Lavaur, J. M., & Bairstow, D. (2011). Languages on the screen: is film comprehension related to the viewers' fluency level and to the language in the subtitles? *International Journal of Psychology*, *46*(6), 455–462. https://doi.org/10.1080/00207594.2011.565343

Leung, W. (2001). How to design a questionnaire. *Studentbmj*, (9), 187–189. Retrieved from http://student.bmj.com/student/view-article.html?id=sbmj0106187

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, UK: The European Institute for the Media.

Mitchell, D. C. (1989). Verb guidance and other lexical effects in parsing. *Language and Cognitive Processes*, *4*(3–4), SI123-SI154. https://doi.org/10.1080/01690968908406366

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

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*(5), 482–488. https://doi.org/doi:10.1037/h0045106

Neve, J. J., & Jenniskens, A. J. W. (1994). The legibility of TV-subtitles for the elderly and visually impaired. In A. C. Kooijman, P. L. Looijestijn, J. A. Welling, & G. J. van der Wildt (Eds.), *Low vision: research and new developments in rehabilitation* (pp. 243–250). Amsterdam, the Netherlands: IOS Press.

Neves, J. (2005). *Audiovisual Translation: Subtitling for the Deaf and Hard-of-Hearing*. Unpublished doctoral dissertation, Roehampton University. Retrieved from http://roehampton.openrepository.com/roehampton/handle/10142/12580

Nobili, P. (1995). Cinema(to)grafo. L'uso didattico dei sottotitoli per l'apprendimento delle lingue straniere. *SILTA - Studi Italiani Di Linguistica Teorica Ed Applicata*, *1*, 151–172.

Ofcom. (2006). *Television Access Services Review: Statement*.

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

Patel, S. (2016). 85 percent of Facebook video is watched without sound. *Digiday UK*. Retrieved from https://digiday.com/media/silent-world-facebook-video/

Perego, E. (2005). *La traduzione audiovisiva*. Roma, Italy: Carocci editore.

Perego, E. (2008). 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

Porteiro, M. (2013). The use of subtitles in speech-language therapy. *Perspectives: Studies in Translatology*, *21*(1), 57–67. https://doi.org/10.1080/0907676x.2012.722654

Quigley, S., & Paul, P. (1984). *Language and Deafness*. San Diego, CA: 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., 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

Ribas Arumí, M., & Romero-Fresco, P. (2008). A Practical Proposal for the Training of Respeakers. *The Journal of Specialised Translation*, *10*, 106–127.

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

Skoog, T. (2016). *Impact of Text in News Videos for the Viewing Experience on Smartphones*. Masters thesis. Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-192658

Szarkowska, A., Laskowska, M., Oliver, A., & Pilipczuk, O. (2015). Reading subtitles across devices: A study into the differences in reading patterns of people watching subtitled videos on smartphone, tablet and computer screen. In *The International Conference on Eyetracking and Applied Linguistics*. Warsaw.

Tamayo, A. (2016). Reading speed in subtitling for hearing impaired children: an analysis in Spanish television. *The Journal of Specialised Translation*, *26*, 275–294.

Zárate, S. (2010). Bridging the gap between Deaf Studies and AVT for Deaf children. In J. Díaz Cintas, A. Matamala, & J. Neves (Eds.), *New Insights into Audiovisual Translation and Media Accessibility. Media for All 2*. (pp. 159–173). Amsterdam: Rodopi.

Annex 2: Research documentation

Documents related to Articles 1 and 2

The experiments from Articles 1 and 2 were part of the eye-tracking study conducted for the SURE project and shared some documents reported in this annex. The complete dataset for these articles can be found in the repository RepOD: https://repod.pon.edu.pl/dataset/sure-project/resource/8866c73b-bd31-4b12-be7c-745686a7660e (refer to the last section of the files named "Experiment3_SubtitleSegmentation" in the dataset).

Recruitment form for participants

UCL

PARTICIPANTS NEEDED

Are you a native speaker of English, Spanish or Polish between 20 and 35 years old?

UCL Centre for Translation Studies
and Deafness, Cognition and Language Centre
would like to hear from you!

We are carrying out an eyetracking study exploring the quality of subtitles.

Participants will complete 90-minute testing session
involving watching subtitled videos and answering questions related to them.

We will provide £10 payment as compensation for your time.

This study has full ethical approval.
Your participation will be completely confidential.

To take part or to request more information, please email Principal Investigator, Dr Agnieszka Szarkowska at a.szarkowska@ucl.ac.uk or Olivia Gerber at olivia.gerber@uab.cat.

For more information see:
https://www.youtube.com/watch?v=yWpaapImWCA

www.facebook.com/SureProject

UCL
CENTRE FOR TRANSLATION STUDIES (CENTRAS)

Centre for Translation Studies
Recruitment Flyer version 1
Project ID: 9607/001
Project acronym: SURE

SURE
Eyetracking Subtitling

Informed Consent Form

Informed Consent Form

Please complete this form after you have read the Information Sheet.

Title of Project: *SURE – Exploring the Subtitle Reading Process with Eyetracking Technology*

This study has been approved by the UCL Research Ethics Committee (Project ID Number): 9607/001.

Thank you for your interest in taking part in this research. Before you agree to take part, the person organising the research must explain the project to you.

If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you to decide whether to join in. You will be given a copy of this Consent Form to keep and refer to at any time.

Participant's Statement

I

- have read the notes written above and the Information Sheet, and understand what the study involves.

- understand that if I decide at any time that I no longer wish to take part in this project, I can notify the researchers involved and withdraw immediately.

- consent to the processing of my personal information for the purposes of this research study.

- understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.

- agree that the research project named above has been explained to me to my satisfaction
and I agree to take part in this study.

I agree that my non-personal research data may be used by others for future research.
I am assured that the confidentiality of my personal data will be upheld through the removal of identifiers. All data will also be anonymous in the final report so that nothing can be attributed back to an individual participant.

Signed: Date:

Information Sheet

Information Sheet

You will be given a copy of this information sheet.
Title of Project: *SURE – Exploring the Subtitle Reading Process with Eyetracking Technology*
This study has been approved by the UCL Research Ethics Committee (Project ID Number): 9607/001
Name: Dr Agnieszka Szarkowska
Work Address: 50 Gordon Square London WC1H 0PQ
Contact Details: a.szarkowska@ucl.ac.uk

Details of study
We would like to invite you to participate in this eyetracking study whose goal is to study the quality of subtitling.
We are recruiting hearing native speakers of Polish, Spanish, and English as well as deaf people who can read text written in English.
The study will take place in the Centre for Translation Studies and Deafness, Cognition and Language Centre at 49-50 Gordon Square at University College London.
During the study, you will watch short film fragments displayed on a computer screen and will be asked to answer a few questions related to the videos and to the subtitles.
During the experiment, your gaze will be recorded using an eye tracker. After completing the study, all participants will be given £10 for as a thank you for their participation.
We will collect information related to your age, gender, and first language. Only the researcher will have access to this information. This data collected will be stored securely and it will not be shared with anyone.
You can withdraw from the study at any time without any consequences. If you decide to take part, you will be given this information sheet to keep and will be asked to sign a consent form.
As participation is anonymous, it will not be possible for us to withdraw your data once you have returned your questionnaire. It is up to you to decide whether to take part or not; choosing not to take part will not disadvantage you in any way. If you do decide to take part, you are still free to withdraw at any time and without giving a reason.
All data will be collected and stored in accordance with the Data Protection Act 1998.
Thank you for reading this information sheet and for considering take part in this research.

Protocol

1. Welcome the participant.
2. Ask him or her to sign the visitor's book.
3. Show the lab.
4. Give the Information Sheet and Informed Consent to sign. Each participant gets a copy of the Information Sheet.
5. Ask the participant to turn off their mobile phone.
6. Switch off your mobile phone.
7. Turn off the Internet in the eye tracking laptop.
8. Show the participant EU chart with reading and listening (except for English participants). Write it in the Experiment log file using the following abbreviations: Note small letters and no spaces. Use the same system in the Participant Information window in SMI Experiment.

Gender

f – female

m – male

Age – write the number using two digits, e.g. 25

Language

en – English

sp – Spanish

pl – Polish

df – Deaf

hh – hard of hearing and deafened

Proficiency in English (listening, reading)

a1

a2

b1

b2

c1

c2

ns for hearing English native speakers

9. Sit the participant. Explain about three parts of the experiment, a break in between, and an interview at the end.
10. Open Experiment 1 in SMI Experiment Centre. Double click on the bottom-right corner to show the participant their eyes. Instruct them how to sit. For left-handed people, move the mouse left of the keyboard.
11. Press RECORD. Complete the Participant Property Editor, using the participant code from the table (note capital and small letters).
12. Press ENTER on your keyboard and SPACE to get the calibration dot moving.
13. In the Experiment log, please note any comments related to conducting the experiment, such as problems with calibration, skipping some stimulus, etc.
14. After Experiment 1, offer the participant some water. Ask about the toilet. Let them have a short break if necessary.
15. Conduct Experiment 2. Possible break.
16. Conduct Experiment 3.
17. Interview. RECORD. Don’t forget to say the participant code.
18. Debrief them - explain what we tested, ask if they have any questions
19. Give the participant GBP10 and ask them to sign the sheet.
20. Make sure they sign off at the main door.

Instructions for deaf participants

Welcome to our eye tracking study on subtitling!

We are working to make subtitles better. In this study we look into two aspects:
1) how fast subtitles should be
2) and what two-line subtitles should look like on screen.

The study consists of three experiments. In the first one, you will watch videos dubbed into Hungarian with English subtitles (they are films originally made in English but now they have a Hungarian audio). There will be a training session, which is not recorded, to familiarise you with the procedure and the questions. After the training session, you will watch three clips of about 5 minutes each and answer a number of questions related to their content and the subtitles. In this experiment, we will ask you about the speed of subtitles: are they too slow or too fast? Do they stay on the screen long enough to read them?

The second experiment is also about the speed of subtitles, but this time the clips have audio in English. There is no training session and there are two clips.

The last experiment is not about the speed of subtitles, but about their shape and content. We want to know which line division is easier to read 2-line subtitles. We will ask you whether you prefer subtitles to have units of meaning together or whether it is more important for you that a subtitle looks like a pyramid or rectangle. At the end, there will be two clips: one with 3line subtitles and one with 2-line subtitles, and we will ask you which ones you prefer.

Before we begin, please read the Information Sheet and sign the Informed Consent Form. Take a seat and make yourself comfortable. You may want to rest your back against the chair. You will see instructions in written English on the screen. You will need the mouse to answer the questions on the screen and the SPACE BAR to move to the next stimulus. Please make sure you only press the space bar ONCE.

We start with calibration, which is telling the computer where your eyes are. You will see your eyes on the screen. They should be in the centre. If you see any arrows, they will point to how you should move (closer, further from the screen). During the calibration, you will see a pulsating dot which will move around the screen. Your task is to look at the dot for as long as it's there. Don't try to predict where it's going to go. Follow the dot with your eyes (not the whole head). Sometimes it is necessary to repeat calibration a few times to make sure the data we collect are good.

During the experiment, please try not to move too much. After each experiment, you can also have a short break if you wish. We will calibrate your eyes again before each experiment. All the clips have subtitles in English. They only contain the dialogues. There is no information about the sounds. At the end of the experiment, we will ask you what you thought about the subtitles.

Documents related to Article 3

The experiment from Article 3 was part of an eye-tracking study conducted for the HBB4ALL project and shared some documents reported in this annex, although eye-tracking measures were not taken in the study conducted for this PhD thesis.

Information Sheet (Spanish)

INFORMACIÓN A LAS PERSONAS PARTICIPANTES

....

Nombre del proyecto:

HBB4ALL. Proyecto de investigación sobre las tecnologías de accesibilidad para dispositivos desarrollados bajo los estándares HbbTv.

Transmedia CataloniaUniversitat Autònoma de Barcelona

....

El objetivo del proyecto es establecer unos estándares de calidad para la presentación de diferentes tecnologías de accesibilidad a su público final. Estas tecnologías son: subtitulado, audio descripción y traducción a lengua de signos.

Para tal fin, se realizan diferentes estudios en los que se pretende establecer cuáles son las buenas prácticas en la creación de este contenido. De esta forma, se pretende mejorar la usabilidad de estos servicios y permitir que todas las personas puedan acceder a los servicios ofrecidos por el estándar HbbTv.

En este marco, nos vamos a centrar en el subtitulado para estudiar los diferentes elementos que determinan su calidad. Al verse esta modulada por el dispositivo en el que se presenta, nuestro primer estudio va a consistir en recoger información comparativa entre el uso de subtitulado en diferentes dispositivos.

Para este fin, se va a registrar mediante Eye Tracking los movimientos oculares que los participantes realizan y se le va a sumar a esta información las valoraciones subjetivas que los diferentes participantes realicen de las diferentes presentaciones. Esta tecnología no es invasiva y no supone ningún riesgo para su salud.

La participación en el estudio será completamente voluntaria, pudiéndose interrumpir si así lo desea el participante. Los resultados obtenidos son absolutamente confidenciales y se van a usar en exclusiva para publicaciones científicas relacionadas con el proyecto en el que se enmarcan.

Los investigadores a cargo de este experimento son Andreu Oliver Moreno, Olga Soler Vilageliu y Olivia Gerber-Morón.

Para más información: Andreu.Oliver@uab.cat o llamar al Tel. 678576781; Olivia.Gerber@uab.cat o llamar 666328012.

Translation of Information Sheet

INFORMATION FOR PARTICIPANTS

....

Name of the project:

HBB4ALL. Research project on accessibility technologies for devices developed in compliance with HbbTv standards.

Transmedia Catalonia Universitat Autònoma de Barcelona

....

The aim of the project is to establish quality standards for the presentation of different accessibility technologies to end users. These technologies are: subtitling, audio description and sign language translation.

To this end, different studies are carried out to establish good practices in the creation of this content. The aim is to improve the usability of these services and allow everyone to access the services offered by the HbbTv standard.

In this context, we will focus on subtitling to study the different elements that determine its quality. As this is modulated by the device in which it is presented, our first study will consist of collecting comparative information on the use of subtitling in different devices.

To this end, the eye movements of the participants will be recorded by means of eye tracking and the subjective evaluations made by the different participants of the different presentations will be added to this information. This technology is non-invasive and poses no risk to your health.

Participation in the study will be completely voluntary and may be interrupted if the participant so wishes. The results obtained are absolutely confidential and will be used exclusively for scientific publications related to the project in which they are included.

The researchers in charge of this experiment are Andreu Oliver Moreno, Olga Soler Vilageliu and Olivia Gerber-Morón.

For more information: Andreu.Oliver@uab.cat or call 67857676781; Olivia.Gerber@uab.cat or call 666328012

Informed Consent Form (Spanish)

CONSENTIMIENTO INFORMADO

Transmedia Catalonia
Universitat Autònoma de Barcelona

Nombre del proyecto: HBB4ALL. Proyecto de investigación sobre las tecnologías de accesibilidad para dispositivos desarrollados bajo los estándares HbbTv.

Nombre del investigador: Olivia Gerber-Morón (Olivia.Gerber@uab.cat)

- He leído y comprendo la hoja de información sobre el experimento, y he tenido la oportunidad de hacer preguntas sobre posibles dudas.
- Entiendo que mi participación en el experimento es voluntaria y que mis datos personales se mantendrán en el anonimato.
- Entiendo que puedo suspender mi participación en el experimento en cualquier momento y sin justificación previa, y sin que esto tenga ninguna repercusión adicional.

Nombre del participante:___________________________

Firma del participante (o representante):

Firma del investigador:____________________________

Fecha: _____________________

Translation of Informed Consent Form

INFORMED CONSENT FORM

Transmedia Catalonia
Universitat Autònoma de Barcelona

Project name: HBB4ALL. Research project on accessibility technologies for devices developed in compliance with HbbTv standards.

Researcher's name: Olivia Gerber-Morón (Olivia.Gerber@uab.cat)

- I have read and understood the information sheet about the experiment, and have had the opportunity to ask questions about possible doubts.
- I understand that my participation in the experiment is voluntary and that my personal information will remain anonymous.
- I understand that I may withdraw from the experiment at any time and without prior justification, and without any further consequences.

Participant's Name:____

Participant's (or representative's) signature:

Researcher's signature:_____________________________________

Date: ______________________________

Questionnaires

<u>Reading and layout preference questionnaire for subtitles (Spanish)</u>

This questionnaire was the same one across clips and devices.

¿En que dispositivo has visto este clip? *

Móvil/Tablet/Monitor

¿Qué porcentaje de subtítulos crees que NO te ha dado tiempo a leer? *

	0	1	2	3	4	5	6	7	8	9	10
0%											100%

¿Cómo te ha parecido la lectura de los subtítulos? *

Valora de 1 al 10

	0	1	2	3	4	5	6	7	8	9	10
Muy difícil										Muy fácil	

¿Cómo valorarías tu experiencia viendo una película en este dispositivo? *

Valora de 1 al 10

	0	1	2	3	4	5	6	7	8	9	10
Placentera/cómoda								No placentera/incomoda			

¿Cree que ha perdido partes esenciales de la acción por leer los subtítulos?

*Sí

No

¿Cómo te ha parecido la longitud de los subtítulos para este dispositivo? *

Cada línea de texto te ha parecido...

- muy larga
- larga
- la longitud era la adecuada
- corta
- muy corta

¿Qué opinas de la duración de los subtítulos en la pantalla? *

- Muy larga
- Larga
- Adecuada
- Corta
- Muy corta

Para este dispositivo, ¿qué opinas sobre la presentación de los subtítulos?

- Adecuada
- Hubiera preferido subtítulos más largos, pero solo en una línea
- Hubiera preferido subtítulos más cortos, pero presentados en dos líneas
- Hubiera preferido subtítulos más cortos y solo en una línea

¿Cómo te ha parecido la calidad de los subtítulos? *

- He perdido mucho tiempo leyendo los subtítulos y no he apreciado adecuadamente el resto del contenido audiovisual en pantalla
- He perdido un poco de tiempo leyendo los subtítulos y no apreciado todos los detalles del resto del contenido audiovisual en pantalla
- He leído cómodamente los subtítulos y me ha dado tiempo de apreciar bastante el resto del contenido audiovisual en pantalla
- He leído cómodamente los subtítulos, que me han ayudado a apreciar el resto del contenido audiovisual en pantalla

Translation of reading and layout preference questionnaire for subtitles

What device have you seen this clip on? *

Smartphone/Tablet/Monitor

What percentage of subtitles do you think you haven't had time to read? *

0	1	2	3	4	5	6	7	8	9	10
0%										100%

How did you like the reading of the subtitles? *

Values from 1 to 10

0	1	2	3	4	5	6	7	8	9	10
Very difficult										Very easy

How would you rate your experience watching a movie on this device? *

0	1	2	3	4	5	6	7	8	9	10
Pleasant/comfortable								Unpleasant/comfortable		

Do you think you have lost essential parts of the action by reading the subtitles? *

- Yes
- No

How did you like the length of the subtitles for this device? *

Each line of text has seemed to you...

- Very long
- Quiet long
- The length was the right one
- Short films
- Very short

What do you think of the duration of the subtitles on the screen? *

- Very long
- Long
- Adequate
- Short
- Very short

For this device, what do you think about the presentation of the subtitles?

- Adequate
- I would have preferred longer subtitles, but only on one line
- I would have preferred shorter subtitles, but presented in two lines
- I would have preferred shorter subtitles and only on one line

What did you think of the quality of the subtitles? *

- I've spent a lot of time reading the subtitles and didn't properly appreciate the rest of the audiovisual content on screen.
- I've spent a little time reading the subtitles and didn't appreciate all the details of the rest of the audiovisual content on screen.
- I have read the subtitles comfortably and have had time to appreciate the rest of the audiovisual content on screen.
- I have read the subtitles comfortably, which have helped me to appreciate the rest of the audiovisual content on screen.

Comprehension questionnaires (Spanish)

SPN Headhunters Clip 1

*Obligatorio

ID *

1. La mujer de Roger le enseñó una pintura de... *
    a. Rembrandt
    b. Rubens
    c. Jordaens
    d. No lo recuerdo
2. La pintura está valorada en hasta... *
    a. 10 millones
    b. 50 millones
    c. 100 millones
    d. No lo recuerdo
3. La mujer de Roger quiere mover el cuadro a... *
    a. un museo
    b. su galería
    c. una caja de seguridad
    d. No lo recuerdo
4. La abuela de Clas... *
    a. recibió la pintura de un oficial alemán
    b. robó la pintura a un oficial alemán
    c. compró una pintura a un oficial alemán
    d. No lo recuerdo
5. Roger quiere celebrar... *
    a. su relación y la inauguración de la galería de su mujer
    b. la fortuna que ganarán gracias al cuadro
    c. su aniversario de bodas
    d. No lo recuerdo

SPN Headhunters Clip 2

*Obligatorio

ID *

1. Roger dice que el currículum de Lander es... *
    a. insuficiente
    b. como muchos otros
    c. impresionante
    d. No lo recuerdo
2. Roger le pregunta a Lander si tiene alguna pintura en el mismo rango de precios que la suya. Lander responde que... *
    a. Sí, tiene una litografía de Much
    b. Sí, tiene una litografía de Munch
    c. No, pero su mujer tiene una litografía de Munch en su galería
    d. No lo recuerdo
3. La mujer de Lander trabaja en... *
    a. un bufete de abogados
    b. un hospital
    c. una galería
    d. No lo recuerdo
4. Según Roger, Lander cometió un error porque... *
    a. encontró a alguien que le recomendara en lugar de presentarse él mismo
    b. no encontró a alguien que le recomendara en lugar de presentarse él mismo
    c. se sorprendió cuando contactaron con él
    d. No lo recuerdo
5. Lander piensa que si sigue el consejo de Roger... *
    a. no va a conseguir el trabajo
    b. pensarán que es ambicioso
    c. pensarán que no es serio
    d. No lo recuerdo

SPN Headhunters Clip 3

*Obligatorio

ID *

1. Clas Greve se mudó a Oslo porque... *
   a. quiere buscar un trabajo aquí
   b. tiene un trabajo bien pagado aquí
   c. quiere decorar la casa de su bisabuela
   d. No lo recuerdo
2. Roger piensa que... *
   a. Clas debería presentarse a un puesto como directivo de Pathfinder
   b. la compañía de Clas debería comprar Pathfinder
   c. Clas debería recibir un ascenso
   d. No lo recuerdo
3. ¿De dónde sacó Clas Greve el bolígrafo con el logo de la compañía? *
   a. Fue un regalo de la compañía para la que trabajó
   b. Lo robó de la compañía para la que trabajó
   c. Es un regalo del nuevo jefe de Clas
   d. No lo recuerdo
4. ¿Desde dónde se mudó Clas Greve a Oslo? *
   a. Bélgica
   b. Holanda
   c. Otra parte de Noruega
   d. No lo recuerdo
5. Roger invita a Clar a... *
   a. almorzar
   b. cenar
   c. tomar un café
   d. No lo recuerdo

<u>Translation of comprehension questionnaires 1</u>

SPN Headhunters Clip 1

* Mandatory

ID *

Comprehension questions

1. Roger's wife showed him a painting of.... *
   a. Rembrandt
   b. Rubens
   c. Jordaens
   d. I don't remember.
2. The painting is valued at up to... *
   a. 10 million
   b. 50 million
   c. 100 million
   d. I don't remember.
3. Roger's wife wants to move the painting to a.... *
   a. a museum
   b. her gallery
   c. a safe deposit box
   d. I don't remember.
4. Clas's grandmother... *
   a. received the painting from a German officer
   b. stole the painting from a German officer
   c. bought the painting to a German officer
   d. I don't remember.
5. Roger wants to celebrate.... *
   a. his relationship and the opening of his wife's gallery
   b. the fortune they'll make from the painting
   c. his wedding anniversary
   d. I don't remember.

SPN Headhunters Clip 2

* Mandatory

ID *

1. Roger says Lander's CV is... *

    a. insufficient

    b. like many others

    c. impressive

    d. I don't remember.

2. Roger asks Lander if he has any paintings within the same price range as his own. Lander answers that... *

    b. Yes, he's got a lithograph of Much

    c. Yes, he's got a lithograph of Munch

    d. No, but his wife has a lithograph of Munch in her gallery.

    e. I don't remember.

3. Lander's wife works in the... *

    a. a law firm

    b. a hospital

    c. a gallery

    d. I don't remember

4. According to Roger, Lander made a mistake because.... *

    a. He found someone to recommend him instead of introducing himself.

    b. He couldn't find anyone to recommend him instead of introducing himself.

    c. He was surprised when they contacted him.

    d. I don't remember.

5. Lander thinks that if he follows Roger's advice... *

    a. he's not going get the job.

    b. they will think that it is ambitious

    c. they will think that it's not serious

    d. I don't remember.

SPN Headhunters Clip 3

* Mandatory

ID *

1. Clas Greve moved to Oslo because... *

    a. he wants to look for a job here

    b. he has a well-paid job here

    c. he wants to decorate his great-grandmother's house

    d.I don't remember.

2. Roger thinks that... *

    a. Clas should apply for a position as a Pathfinder manager

    b. Clas's company should buy Pathfinder

    c. Clas should get a promotion

    d. I don't remember.

3. Where did Clas Greve get the pen with the company logo? *

    a. It's a gift from the company he worked for.b. He stole it from the company he worked for

    c. It's a gift from the new head of class.

    d. I don't remember.

4. Where did Clas Greve move to Oslo from? *

    a. Belgium

    b. Netherlands

    c. Other part of Norway

    d. I don't remember.

5. Roger invites Clar to... *

    a. have lunch

    b. have dinner

    c. have coffee

    d. I don't remember.

Documents related to Article 4

The following documents are part of the research documentation used for Article 4 (Gerber-Morón, submitted), which under review at the time of submission.

Cold-calling information sheet for the recruitment of deaf and hard-of-hearing participants (Italian)

ATTIVITÀ PER ADATTARE I SOTTOTITOLI

È fondamentale adattare i sottotitoli alle nuove tecnologie perché ci permetteranno di comunicare meglio e di essere tutti integrati in questa società dell'informazione. Non si è fatta ancora ricerca sulla lettura e la presentazione dei sottotitoli visionati su schermi più piccoli, come gli smartphone, tablet e altri prodotti simili.

L'ESPERIMENTO

L'obiettivo è capire come migliorare la qualità dei sottotitoli per gli schermi di smartphone per tutti i possibili utenti di sottotitoli: giovani, persone della terza età e sordi. Questa ricerca fa parte del progetto europeo "Hybrid Broadband Broadcast For All" (HBB4ALL).

Si guarderà la puntata di una serie televisiva di 20 minuti sottotitolata in italiano. La puntata sarà visionata su uno schermo di smartphone.

Successivamente, il partecipante risponderà a un questionario sulla comprensione, la sottotitolazione e l'esperienza globale. Tutti i dati ricavati rimarranno anonimi. L'esperimento dura circa un'ora.

LA RICERCATRICE

Olivia Gerber Morón è una dottoranda spagnola/svizzera italiana che lavora all'Università Autonoma di Barcellona. Per ulteriori informazioni: oliviagerb@hotmail.com.

Siete interessati a partecipare? Iscrivetevi entro 26 giugno 2015 a info-i@sgb-fss.ch e poi ci metteremo d'accordo tramite doodle quando faremo.

Olivia Gerber Morón in collaborazione con

SGB-FSS
Federazione Svizzera
dei Sordi

via Besso 5 6900 Lugano
Telefono 091 950 05 48 Fax 091 960 04 38
info-i@sgb-fss.ch www.sgb-fss.ch

CCP 65-752278-9

Translation of the cold-calling information sheet for the recruitment of deaf and hard-of-hearing participants

WORK TO ADAPT SUBTITLES

It is essential to adapt subtitles to new technologies because they will allow us to communicate better and be integrated into the information society. No research has yet been done on subtitling reading and presentation on smaller screens, such as smartphones, tablets and other similar products.

THE EXPERIMENT

The aim is to understand how to improve the quality of subtitles for smartphone screens for all possible users of subtitles: young people, older people and deaf people. This research is part of the European project "Hybrid Broadband Broadcast For All" (HBB4ALL).

You will watch the episode of a 20-minute television series subtitled in Italian. The episode will be viewed on a smartphone screen.

Subsequently, the participant will answer a questionnaire on comprehension, subtitling and the global experience. All data obtained will remain anonymous. The experiment lasts about an hour.

THE RESEARCHER

Olivia Gerber Morón is a Spanish/Swiss Italian PhD student working at the Autonomous University of Barcelona. For further information: oliviagerb@hotmail.com.

Are you interested in participating? Register by 26 June 2015 at info-i@sgb-fss.ch and then we will agree through doodle when we do.

Olivia Gerber Morón in collaboration with

SGB-FSS
Federazione Svizzera dei Sordi

via Besso 5 6900 Lugano
Telefono 091 950 05 48 Fax 091 960 04 38
info-i@sgb-fss.ch www.sgb-fss.ch

CCP 65-752278-9

Informed Consent Form (Italian)

CONSENSO INFORMATO

Stiamo conducendo una ricerca finalizzata a migliorare la qualità dei sottotitoli per schermi piccoli nell'ambito del progetto europeo "Hybrid Broadcast Broadband For All" (HBB4ALL). Se Lei deciderà di prendere parte a questo esperimento, Le verrà chiesto di vedere due brevi filmati e di rispondere a un questionario.

È importante che sappia che **non ci sono rischi associati** a questo esperimento. Tutte le Sue risposte saranno strettamente confidenziali e saranno codificate al fine di mantenere il Suo **anonimato**. La Sua partecipazione dipende totalmente da Lei, che potrà sentirsi **libero in qualsiasi momento di abbandonare l'esperimento senza alcuna penalizzazione**.

Se vorrà saperne di più, per cortesia contatti la dottoranda Olivia Gerber (Olivia.Gerber@uab.cat).

***

Io sono d'accordo a prendere parte a questo esperimento. Sono cosciente che partecipare dipende totalmente da me e che potrò abbandonare quando vorrò.

NOME E COGNOME	DATA	FIRMA

Translation of Informed Consent Form

INFORMED CONSENT FORM

We are conducting a research that aims to improve the quality of subtitles for secondary screens in the framework of the European project "Hybrid Broadcast Broadband For All" (HBB4ALL). If you decide to take part in this experiment, you will be asked to watch two short films and answer a questionnaire.

It is important that you know that **there are no risks associated** with this experiment. All your answers will be strictly confidential and will be coded in order to maintain your **anonymity**. Your participation depends totally on you, who can feel **free at any time to leave the experiment without any penalty**.

If you would like to know more, please contact doctoral candidate Olivia Gerber (Olivia.Gerber@uab.cat).

***

I agree to take part in this experiment. I am aware that participating totally depends on me and that I can leave it at any time.

NAME AND SURNAME	DATE	SIGNATURE

HBB 4 all
Connected TV Accessibility

UAB
Universitat Autònoma de Barcelona

Comprehension and subtitle segmentation questionnaires (Italian)

Clip 1

Cod	

Qui di seguito ti proporremo alcune domande di carattere generale sul filmato che hai appena visto. Il tuo compito è quello di fare una X sulla casella corrispondente alla singola risposta che vuoi dare per ogni domanda che ti viene proposta.

1. Nella prima scena dello spezzone, il capo cameriere chiede al protagonista della serie, l'agente pubblicitario Don Draper, se il cameriere di colore Sam lo sta disturbando. Draper dice che sta	
parlando con lui	☐
chiedendogli da bere	☐
discutendo	☐

2. L'agente pubblicitario protagonista della serie, Don Draper, vuole sapere perché il cameriere di colore Sam preferisce fumare le sigarette Old Gold. Sam gli spiega che a militare	
i suoi compagni fumavano queste sigarette	☐
ricevevano uno sconto per queste sigarette	☐
davano gratis queste sigarette	☐

3. Nella scena successiva, Don Draper si reca alla casa della pittrice Midge, la sua amante, per passare la notte con lei. Le spiega che la Commissione Federale	
accetta le sue proposte per la nuova campagna pubblicitaria sulle sigarette	☐
rifiuta le sue proposte per la nuova campagna pubblicitaria sulle sigarette	☐
ignora le sue proposte per la nuova campagna pubblicitaria sulle sigarette	☐

4. Midge, l’amante dell’agente pubblicitario Don Draper, lo incoraggia	
ad affrontare la sua paura per la campagna pubblicitaria sulle sigarette	☐
a fare attenzione alle scelte che farà per la campagna pubblicitaria sulle sigarette	☐
a sviluppare nuove idee insieme ad altri colleghi per la campagna pubblicitaria sulle sigarette	☐

5. Quando si svegliano, Don Draper dice alla sua amante Midge che dovrebbero sposarsi. Come reagisce Midge?	
accoglie con interesse la proposta	☐
non vuole prendere impegni né fare progetti	☐
gli chiede del tempo per rifletterci	☐

6. Nella scena successiva, vediamo degli impiegati entrare nell’ufficio del collega Pete Campbell. Lui sta parlando al telefono con la sua fidanzata. Dalla conversazione, possiamo dedurre che la sua fidanzata è	
entusiasta per la festa d’addio al celibato di Pete	☐
tranquilla per la festa d’addio al celibato di Pete	☐
	☐

preoccupata per la festa d'addio al celibato di Pete

7. La segretaria-capo Joan Holloway spiega il funzionamento dell'agenzia pubblicitaria Sterling Cooper a Peggy Olson, la nuova segretaria. Peggy lavorerà nel piano dei dirigenti. Secondo la segretaria-capo Joan, questo piano è	
molto organizzato	☐
molto dinamico	☐
molto caotico	☐

8. Che tipo di consigli dà la segretaria-capo Joan Holloway alla nuova segretaria Peggy Olson?	
Consigli sul lavoro che deve svolgere e sull'ambito personale	☐
Consigli solo sul lavoro che debe svolgere	☐
Consigli solo sull'ambito personale	☐

9. Nell'ultima scena dello spezzone, il proprietario della loro agenzia, il signor Roger Sterling (un signore con i capelli bianchi), parla con Don Draper del loro incontro con i proprietari delle sigarette Lucky Strike. Roger immagina che Draper	
non abbia ancora trovato delle proposte per l'incontro	☐
abbia trovato già delle proposte per l'incontro	☐
abbia bisogno del suo aiuto per trovare delle proposte per l'incontro	☐

10. Il proprietario dell'agenzia pubblicitaria Sterling Cooper, il signore Roger Sterling, chiede a Don Draper se ci sono degli impiegati ebrei nell'agenzia. Perché Roger Sterling vuole che ci sia un impiegato ebreo nella riunione con i clienti ebrei?	
Perché avere la presenza di un impiegato ebreo nella riunione farà che i clienti si sentano più a loro agio	☐
Perché l'impiegato ebreo potrà spiegare più facilmente le strategie ai clienti ebrei	☐
Perché vuole coinvolgere attivamente gli impiegati ebrei dell'agenzia	☐

Qui di seguito ti proporremo alcune domande specifiche sui dialoghi relativi al filmato che hai appena visto. Il tuo compito è quello di fare una X sulla casella corrispondente alla singola risposta che vuoi dare per ogni domanda che ti viene proposta.

1. Nella prima scena vediamo l'agente pubblicitario Don Draper, seduto a un tavolo, che inizia a parlare con Sam, un cameriere di colore. Il capo cameriere interrompe loro e dice a Don Draper che Sam è un po'	
assillante	☐
pettegolo	☐
ciarliero	☐

2. Il cameriere di colore Sam spiega al protagonista Don Draper che sua moglie ha letto sul Reader's Digest che il fumo	
provoca il cancro	☐
uccide	☐
danneggia i polmoni	☐

3. Quando Don Draper si reca a casa della sua amante, la pittrice Midge, per passare la notte, essa gli spiega che sta lavorando per un progetto sulla	
"Celebrazione della Nonna"	☐
"Festa della Nonna"	☐
"Giornata della Nonna"	☐

4. Dopo di che, Midge spiega a Draper che per alcuni mesi si occuperà di disegnare	
cagnolini	☐
cuccioli	☐
lupetti	☐

5. Don Draper spiega alla sua amante Midge che adesso non è più possibile dire che la sigaretta	
calma la tosse	☐
calma la gola	☐
calma la voce	☐

6. L'amante Midge capisce che Don Draper è preoccupato per la faccenda delle sigarette e prova a incoraggiarlo parlandogli della sua genialità. Per farlo, gli dice che con la sua genialità riuscirà a portare le pecore al	
macello	☐
mattatoio	☐
ammazzatoio	☐

7. Nonostante ciò, l'agente pubblicitario Don Draper si mostra insicuro e pessimista. Dice a Midge che presto lo troverà in mezzo a dei giovani dirigenti intenti a	
cavargli gli occhi	☐
spolverargli le ossa	☐
spolparlo vivo	☐

8. Nella scena successiva, vediamo entrare nell'ascensore la nuova segretaria Peggy Olson. Dietro di lei ci sono degli uomini che iniziano a prenderla in giro e uno di loro dice di far salire l'ascensore senza fretta, di prendere	
la via lunga per godersi il panorama	☐
la strada lunga per godersi il panorama	☐
il percorso lungo per godersi il panorama	☐

9. Nell'ultima scena, vediano il protagonista Don Draper che si prepara per un incontro nel suo ufficio. Lo segue Roger Sterling, il proprietario dell'agenzia (l'uomo con i capelli bianchi). Quando entrano nell'ufficio, Roger trova che Draper	
sia stanco morto	☐
sia a pezzi	☐
abbia l'aria distrutta	☐

10. Nella stessa scena, Roger Sterling chiede a Draper se ci sono degli ebrei che lavorano nella loro agenzia. Hanno dei clienti ebrei e vorrebbe che	
si sentissero rassicurati	☐
si sentissero a loro agio	☐
si sentissero accolti	☐

Clip 2

Cod	

Qui di seguito ti proporremo alcune domande di carattere generale sul filmato che hai appena visto. Il tuo compito è quello di fare una X sulla casella corrispondente alla singola risposta che vuoi dare per ogni domanda che ti viene proposta.

1. Nella prima scena dello spezzone, vediamo l'agente pubblicitario Don Draper nel suo ufficio allenandosi quando entra Salvatore Romano, il disegnatore italiano dell'agenzia Sterling Cooper. Salvatore gli mostra il disegno che propone per la campagna pubblicitaria delle sigarette. Don Draper è soddisfatto della proposta?	
sì, gli dice che ha avuto un'idea grandiosa	☐
sì, gli piace, ma deve modificare certi dettagli	☐
no, non gli piace per niente	☐

2. Nella stessa scena, Don Draper dice a Salvatore che non ha l'intenzione di andare alla festa d'addio al celibato dell'impiegato Pete Campbell. Salvatore gli confessa che nemmeno lui vuole andarci perché queste feste sono	
costose	☐
imbarazzanti	☐
noiose	☐

3. All'improvviso entra Greta Gurtman, un'impiegata dell'agenzia Sterling Cooper, per parlare sulla campagna pubblicitaria delle sigarette. Quando lei vede Don Draper e Salvatore, si mostra sorpresa perché entrambi sembrano	
troppo fiduciosi riguardo alla campagna delle sigarette	☐
troppo calmi riguardo alla campagna delle sigarette	☐
troppo indifferenti riguardo alla campagna delle sigarette	☐

4. Qual è il ruolo della signora Greta Gutman all'interno dell'agenzia Sterling Cooper?	
Grafico per le campagne pubblicitarie dell'agenzia	☐
Dirigente creativa per le campagne pubblicitarie dell'agenzia	☐
Ricercatrice per assistere alle campagne pubblicitarie dell'agenzia	☐

5. Come accolgono Don Draper e Salvatore i suggerimenti proposti dalla signora Greta Gutman?	
Non sono d'accordo con lei	☐
Sono soddisfatti	☐
Non capiscono bene le sue proposte	☐

6. La signora Greta Gutman parla del concetto di Freud sul desiderio inconscio di morte che potrebbe usarsi nella campagna pubblicitaria delle sigarette. Don Draper trova che queste riflessioni siano	
irrilevanti per la pubblicità delle sigarette	☐
pratiche per la pubblicità delle sigarette	☐
concrete per la pubblicità delle sigarette	☐

7. In una scena successiva, l'impiegato Pete Campbell entra nell'ufficio dell'agente pubblicitario Don Draper e osserva la nuova segretaria, Peggy Olson. Che tipo di commenti rivolge Pete alla nuova segretaria Peggy?	
commenti positivi e dolci	☐
commenti volgari e maschilisti	☐
commenti critici e costruttivi	☐

8. Successivamente, vediamo l'agente pubblicitario Don Draper e l'impiegato Pete Campbell avviarsi alla riunione con i clienti ebrei. Dalla loro conversazione, cosa si può dedurre sul rapporto tra Don Draper e Pete?	
Don Draper vuole sostenere Pete nella sua carriera	☐
A Don Draper non piace l'atteggiamento di Pete	☐
A Don Draper sta molto simpatico Pete	☐

9. Quando Don Draper e Pete Campbell arrivano alla riunione, Draper si rende conto che il cliente ebreo è una donna, Rachel Menken. Qual è l'atteggiamento di Don Draper nei confronti di Rachel Menken?	
Ha un comportamento maschilista	☐
Si mostra molto diplomatico	☐
Cerca di corteggiarla	☐

10. Perché il cliente ebreo, Rachel Menken, non è d'accordo con la strategia pubblicitaria di Don Draper sui buoni sconto per promuovere il suo negozio?	
Lei crede che perderebbero dei soldi con queste buoni sconto	☐
Lei pensa che i suoi collaboratori ebrei non sarebbero d'accordo su questa proposta	☐
Lei non vuole rivolgersi alle casalinghe	☐

Qui di seguito ti proporremo alcune domande specifiche sui dialoghi relativi al filmato che hai appena visto. Il tuo compito è quello di fare una X sulla casella corrispondente alla singola risposta che vuoi dare per ogni domanda che ti viene proposta.

1. Nella prima scena, il protagonista Don Draper è nel suo ufficio con l'art director Salvatore, il quale gli mostra il disegno che ha fatto per la campagna pubblicitaria delle sigarette. Don Draper gli suggerisce di aggiungere un po' di	
sex appeal disegnando una donna in costume da bagno	☐
erotismo disegnando una donna in costume da bagno	☐
sensualità disegnando una donna in costume da bagno	☐

2. Nella prima scena, il protagonista Don Draper è nel suo ufficio con l'art director Salvatore, quando arriva la signora Greta Guttman per discutere con loro sulla campagna sulle sigarette. Lei afferma che adesso molti considerano le sigarette	
un male	☐
un veleno	☐
una droga	☐

3. La signora Greta Guttman continua il suo discorso con Don Draper e Salvatore, e dice di avere trovato una soluzione. Afferma che il fumo è	
parte integrante della vita americana	☐
parte intrinseca della vita americana	☐
parte sostanziale della vita americana	☐

4. L'agente pubblicitario Don Draper non è convinto dai suggerimenti di Greta Guttman e pensa che la psicologia serve	
ai cocktail	☐
alle feste	☐
agli eventi mondani	☐

5. L'art director Salvatore non è nemmeno convinto di quello che spiega Greta Guttman e cerca di ridicolizzarla dicendo che non è possibile fidarsi di riviste come Reader's Digest, in cui dicono che	
Pinocchio è il libro del secolo	☐
Bambi è il libro del secolo	☐
Dumbo è il libro del secolo	☐

6. In una delle scene successive, Peggy Olson sveglia Don Draper perché l'impiegato Pete Campbell lo aspetta. Draper chiede a Peggy d'intrattenerlo. Lei gli dice che vorrebbe evitare questa situazione nel suo primo giorno, ma non vuole sembrare	
indisponente nei suoi confronti	☐
restia nei suoi confronti	☐
riluttante nei suoi confronti	☐

7. L'impiegato Pete Campbell parla di Peggy Olson al protagonista Don Draper. Gli dice che

se la spasserà lui prima con Peggy	☐
ci proverà lui prima con Peggy	☐
si divertirà lui prima con Peggy	☐

8. Quando il protagonista Don Draper si rende conto delle intenzioni dell'impiegato Pete Campbell con la nuova segretaria Peggy Olso, lo avverte di	
non rovinare la reputazione di una ragazza della segretaria	☐
non infangare la reputazione di una ragazza della segretaria	☐
non macchiare la reputazione di una ragazza della segretaria	☐

9. Nell'ultima scena, Don Draper e i suoi colleghi hanno una riunione con un cliente ebreo, la signorina Rachel Menken. Lei rifiuta duramente le proposte fatte dall'agenzia e questo innervosisce Don Draper. Il proprietario Roger Sterling chiede a Don Draper di	
non alzare un polverone	☐
non alzare i toni	☐
non alzare la cresta	☐

10. L'impiegato Pete Campbell prova a spiegare alla signorina Rachel Menken perché il suo negozio non si può confrontare a Chanel. Secondo lui, Chanel ha il fascino	
della vecchia Europa	☐
dell'Est	☐
di Occidente	☐

Qui di seguito ti proporremo alcune <u>domande relative alla presentazione dei sottotitoli</u> dei due spezzoni che hai appena visto. Il tuo compito è quello di fare una X sulla casella corrispondente al valore della scala che corrisponde meglio alla tua valutazione.

Hai notato delle differenze nella presentazione dei sottotitoli nel primo spezzone rispetto al secondo spezzone?

pochissime	poche	abbastanza poche	nessuna	abbastanza differenze	molte	moltissime

Quale spezzone hai letto in modo più agevole?

il primo	il secondo	non ho notato differenze

Per cortesia, indica con una X la presentazione di sottotitoli che preferisci per le seguenti coppie di immagini per lo schermo di smartphone.

<u>**Coppia A:**</u>

Coppia B:

Per cortesia, rispondi alla seguente domanda sulle tue abitudini di visione di materiale audiovisivo facendo una X sulla risposta che vuoi dare.

Quanto spesso guardi materiale audiovisivo con sottotitoli?

mai	poco	abbastanza poco	né mai né spesso	abbastanza spesso	molto spesso	sempre

Ti chiediamo ora alcune informazioni demografiche (le risposte rimarranno ANONIME):

Sesso:	Maschio			Femmina		
Età:						
Città di residenza:						
Sei di madre lingua italiana?					Sì	No
Se no, da quanti anni parli italiano?						
Conoscevi la lingua che hai ascoltato nel filmato appena visto?					Sì	No
Massima scolarità raggiunta:	Nessun titolo di studio	Licenza Elementare	Licenza Media	Diploma Scuola Secondaria	Laurea	Dottorato
Numero totale di anni di scolarità (somma il numero di anni passati a						

scuola - elementare, media, superiore - e all'università):	

Avevi già visto la puntata della serie che ti è stata presentata	Sì No		
Se sì, quanto tempo fa?			
L'hai vista più di una volta?	Sì No		
Se sì, quante volte?			
Hai già partecipato a un esperimento in cui ti è stato mostrato un filmato?	Sì, questo filmato	Sì, ma un altro filmato	No, mai

Soffri di qualche problema fisico che potrebbe compromettere la visione e la comprensione del filmato?	Sì No
Se sì, quale?	

Translation of comprehension and subtitle segmentation questionnaires

Clip 1

Cod	

Below we will ask you some general questions about the video you have just seen. Your task is to make an X on the box corresponding to the single answer you want to give for each question you are asked.

1.In the first scene of the clip, the chief waiter asks the main character of the series, the advertising agent Don Draper, if the black waiter Sam is disturbing him. Draper says that he is	
talking to him	☐
ordering a drink	☐
discussing	☐

2.The advertising agent in the series, Don Draper, wants to know why the black waiter Sam prefers to smoke Old Gold cigarettes. Sam explains to him that in the army	
his mates smoked these cigarettes.	☐
were getting a discount on these cigarettes.	☐
they were given these cigarettes for free	☐

3. In the next scene, Don Draper goes to the house of the painter Midge, his lover, to spend the night with her. He explains to her that the Federal Commission	
accepts his suggestions for the new cigarette advertising campaign	☐
rejects his suggestions for the new cigarette advertising campaign	☐
ignores his suggestions for the new cigarette advertising campaign	☐

4. Midge, the advertising agent Don Draper's mistress, encourages him	
to face his fear of the cigarette advertising campaign	☐
to pay attention to the choices he will make for the cigarette advertising campaign	☐
to develop new ideas together with other colleagues for the cigarette advertising campaign	☐

5. When they wake up, Don Draper tells his mistress Midge they should get married. How does Midge react?	
welcomes with interest the proposal	☐
does not want to make commitments or make plans	☐
asks him time to think about it	☐

6. In the next scene, we see employees entering the office of their colleague Pete Campbell. He is talking on the phone with his fiancée. From the conversation, we can deduce that his fiancée is	
excited about Pete's stag party	☐
calm for Pete's stag party	☐
worried about Pete's stag party	☐

7. Chief Secretary Joan Holloway explains the operation of the advertising agency Sterling Cooper to Peggy Olson, the new secretary. Peggy will work on the executive floor. According to chief secretary Joan, this floor is	
very organized	☐
very dynamic	☐
very chaotic	☐

8. What kind of advice does head secretary Joan Holloway give to the new secretary Peggy Olson?	
Advice on the work to be carried out and the personal environment	☐
Advice only on the work that needs to be done	☐
Advice on personal matters only	☐

9. In the last scene of the clip, the owner of their agency, Mr. Roger Sterling (a white-haired gentleman), talks to Don Draper about their meeting with the owners of the Lucky Strike cigarettes. Roger imagines that Draper	
hasn't found any meeting suggestions yet	☐
has already found suggestions for the meeting	☐
needs his help finding suggestions for the meeting	☐

10. The owner of the advertising agency Sterling Cooper, Mr. Roger Sterling, asks Don Draper if there are any Jewish employees in the agency. Why does Roger Sterling want a Jewish employee in the meeting with Jewish clients?	
Because having a Jewish employee present at the meeting will make customers feel more comfortable	☐
Because the Jewish employee will be able to explain strategies more easily to Jewish clients	☐
Because he wants to actively involve the Jewish employees of the agency	☐

Below we will ask you some specific questions about the dialogues related to the movie you just saw. Your task is to make an X on the box corresponding to the single answer you want to give for each question you are asked.

1. In the first scene we see advertising agent Don Draper, sitting at a table, starting to talk to Sam, a black waiter. The head waiter interrupts them and tells Don Draper that Sam is a little bit	
annoying	☐
gossipy	☐
chatty	☐

2. The black waiter Sam explains to the protagonist Don Draper that his wife has read in Reader's Digest that smoking	
causes cancer	☐
kills	☐
damages lungs	☐

3. When Don Draper goes to the house of his lover, the painter Midge, to spend the night, she explains to him that she is working on a project on	
"Grandma's Celebration"	☐
"Grandma's Party"	☐
"Grandma's Day"	☐

4. After that, Midge explains to Draper that for a few months she will be busy drawing	
kittens	☐
puppies	☐
wolf cubs	☐

5. Don Draper explains to his lover Midge that now it is no longer possible to say that the cigarette	
calms the cough	☐
calms the throat	☐
calms the voice	☐

6. Lover Midge realizes that Don Draper is worried about the cigarette business and tries to encourage him by telling him about his genius. To do so, she tells him that with his genius he will be able to bring the sheep to the	
slaughter	☐
abattoir	☐
killer	☐

7. Despite this, advertising agent Don Draper is insecure and pessimistic. He tells Midge that soon he will find him in the midst of young executives intent on	
tearing his eyes out	☐
dusting his bones	☐
sprinkling it alive	☐

8. In the next scene, we see the new secretary Peggy Olson entering the elevator. Behind her there are some men who are starting to make fun of her and one of them says to get the elevator up without haste, to take	
the long way to enjoy the view	☐
the long road to enjoy the view	☐
the long route to enjoy the view	☐

9. In the last scene, they see the protagonist Don Draper preparing for a meeting in his office. Roger Sterling, the owner of the agency (the man with the white hair), follows him. When they enter the office, Roger finds that Draper	
is completely tired	☐
is in pieces	☐
is exhauted	☐

10. In the same scene, Roger Sterling asks Draper if there are any Jews working in their agency. They have Jewish clients and would like	
to reassure them	☐
to make them feel comfortable	☐
to make them feel welcomed	☐

Clip 2

Cod	

Below we will ask you some <u>general questions</u> about the video you have just seen. Your task is to make an X on the box corresponding to the single answer you want to give for each question you are asked..

1. In the first scene of the clip, we see the advertising agent Don Draper in his office training when Salvatore Romano, the Italian designer of the agency Sterling Cooper, enters. Salvatore shows him the design he proposes for the cigarette advertising campaign. Is Don Draper satisfied with the idea?	
Yes, you tell him you had a great idea. **Yes, he likes it, but he needs to change certain details** No, he doesn't like it at all	☐ ☐ ☐

2. In the same scene, Don Draper tells Salvatore that he has no intention of going to the employee Pete Campbell's stag party. Salvatore confesses to him that he doesn't want to go either because these parties are	
pricey **embarrassing** boring	☐ ☐ ☐

3. Suddenly Greta Gurtman, an employee of the Sterling Cooper agency, comes in to talk about the cigarette advertising campaign. When she sees Don Draper and Salvatore, she is surprised because they both look like	
too confident about the cigarette campaign **too calm about the cigarette campaign** too indifferent about the cigarette campaign	☐ ☐ ☐

4. What is Mrs Greta Gutman's role in the Sterling Cooper agency?	
Graphic for the advertising campaigns of the agency Creative director for the agency's advertising campaigns **Researcher to assist the advertising campaigns of the agency**	☐ ☐ ☐

5. How do Don Draper and Salvatore welcome the suggestions made by Mrs Greta Gutman?	
I don't agree with her They are satisfied They don't really understand her suggestions	☐ ☐ ☐

6. Mrs. Greta Gutman talks about Freud's concept of the unconscious desire for death that could be used in the cigarette advertising campaign. Don Draper finds these reflections to be	
irrelevant to cigarette advertising handy for the advertising of cigarettes specific for the advertising of cigarettes	☐ ☐ ☐

7. In a later scene, employee Pete Campbell enters the office of advertising agent Don Draper and observes the new secretary, Peggy Olson. What kind of comments does Pete make to the new secretary Peggy?	
positive and kind comments	☐
vulgar and masculine comments	☐
critical and constructive comments	☐

8. Afterwards, we see advertising agent Don Draper and employee Pete Campbell heading to the meeting with Jewish clients. From their conversation, what can be inferred about the relationship between Don Draper and Pete?	
Don Draper wants to support Pete in his career	☐
Don Draper doesn't like Pete's attitude	☐
Don Draper likes Pete very much	☐

9. When Don Draper and Pete Campbell arrive at the meeting, Draper realizes that the Jewish client is a woman, Rachel Menken. What is Don Draper's attitude towards Rachel Menken?	
He has a macho attitude	☐
He shows himself very diplomatic	☐
He's trying to flirt with her	☐

10. Why doesn't the Jewish customer, Rachel Menken, agree with Don Draper's advertising strategy on discount vouchers to promote her store?	
She thinks they're going to lose money with these discount vouchers	☐
She thinks her Jewish associates would disagree with this proposal	☐
She doesn't want to address housewives	☐

Below we will ask you some specific questions about the dialogues related to the movie you just saw. Your task is to make an X on the box corresponding to the single answer you want to give for each question you are asked..

1. In the first scene, the protagonist Don Draper is in his office with art director Salvatore, who shows him the drawing he did for the cigarette advertising campaign. Don Draper suggests that he adds a bit of	
sex appeal by drawing a woman in a swimsuit	☐
eroticism by drawing a woman in a swimsuit	☐
sensuality by drawing a woman in a swimsuit	☐

2. In the first scene, the protagonist Don Draper is in his office with art director Salvatore, when Mrs. Greta Guttman arrives to discuss with them about the cigarette campaign. She says that now many consider cigarettes	
harm	☐
poison	☐
drugs	☐

3. Mrs. Greta Guttman continues her speech with Don Draper and Salvatore, and says she has found a solution. She says that smoking is	
an essential part of American life	☐
an intrinsic part of American life	☐
a substantial part of American life	☐

4. Advertising agent Don Draper is not convinced by Greta Guttman's suggestions and thinks that psychology is useful at	
cocktails	☐
parties	☐
social events	☐

5. Art director Salvatore is not even convinced of what Greta Guttman explains and tries to ridicule her by saying that it is not possible to trust magazines like Reader's Digest, in which they say that	
Pinocchio is the book of the century	☐
Bambi is the book of the century	☐
Dumbo is the book of the century	☐

6. In one of the following scenes, Peggy Olson wakes up Don Draper because employee Pete Campbell is waiting for him. Draper asks Peggy to keep him entertained. She tells him that she would like to avoid this situation on her first day, but she doesn't want to seem	
unsympathetic to him	☐
reluctant towards him	☐
reluctant towards him	☐

7. Employee Pete Campbell tells Don Draper about Peggy Olson. He tells him that

he's going have a good time with Peggy first	☐
he'll try it first with Peggy	☐
he'll have fun with Peggy first	☐

8. When the protagonist Don Draper realizes the intentions of the employee Pete Campbell with the new secretary Peggy Olso, he warns him not to	
ruin the reputation of a girl in the secretary's office	☐
tarnish the reputation of a girl in the secretary's office	☐
stain the reputation of a girl in the secretary's office	☐

9. In the last scene, Don Draper and his colleagues have a meeting with a Jewish client, Miss Rachel Menken. She strongly rejects the offers made by the agency and this makes Don Draper nervous. The owner Roger Sterling asks Don Draper not to	
raise a fuss	☐
go up a notch	☐
raise the crest	☐

10. Employee Pete Campbell tries to explain to Miss Rachel Menken why her shop cannot be compared to Chanel. According to him, Chanel has the charm	
of old Europe	☐
Eastern	☐
Western	☐

Below we will ask you some questions about the presentation of the subtitles of the two clips you have just seen. Your task is to make an X on the box corresponding to the value of the scale that best corresponds to your assessment.

Did you notice any differences in the presentation of the subtitles in the first segment compared to the second one?

almost none	a few	some	none	quite a few	many	a lot

Which fragment did you read most easily?

the first one	the second one	I didn't notice any difference

Please use an X to indicate your preferred subtitle presentation for the following pairs of smartphone screen images

Pair A:

Pair B:

Please answer the following question about your viewing habits of audiovisual material by marking an X on the answer you want to give.

How often do you watch audiovisual material with subtitles?

never	almost never	just sometimes	neither never nor always	quite often	very often	always

We now ask you for some demographic information (the answers will remain ANONYMOUS):

Gender:	Man			Woman		
Years:						
Hometown						
Are you Italian mother tongue?					Yes	No
If not, how many years have you spoken Italian?						
Did you know the language you heard in the video you just saw?					Yes	No
Maximum education achieved:	No qualifications	Elemtary school	Middle school	High school	University degree	PhD

Total number of years of education (sum the number of years spent at school - primary, middle, high - and university):						

Have you seen the episode of the series that you were presented with?	Yes No		
If so, how long ago?			
Have you seen it more than once?	Yes No		
If so, how many times?			
Have you already participated in an experiment in which you were shown a video?	Yes, this fragment	Yes, another film	No, never

Do you have any physical problems that could compromise your viewing and comprehension of the film?	Yes No
If so, which one?	