The International Journal for Translation & Interpreting Research
trans-int.org

Busting "ghost subtitles" on streaming services

Jan-Louis Kruger
Macquarie University, Australia
North-West University, South Africa
Janlouis.kruger@mq.edu.au

Sixin Liao
Macquarie University, Australia
Sixin.liao@mq.edu.au

DOI: 10.12807/ti.117202.2025.a03

Abstract: In this paper we focus on a phenomenon that all subtitle users experience: "ghost subtitles". "Ghost subtitles" are subtitles we notice in our peripheral vision, only to find them gone by the time our eyes have moved down to start reading or disappearing while we are still reading. "Ghost subtitles" often meet the minimum duration and maximum speed requirements set by platforms or broadcasters but disregard the time it takes to move gaze from the image to the subtitle (i.e., processing latency). The one-speed-fits-all approach means that, many subtitles are not on screen long enough to allow viewers to finish reading them, which could result in frustrating viewing experiences. To determine how prevalent fast subtitles are on streaming platforms, this paper presents an analysis of the distribution of subtitle speeds based on a corpus of subtitles from one of the major streaming platforms. We further investigated the impact of subtitle speed and audio language on processing latency based on eye-movement data from a total of 109 participants in two separate experiments. We found that almost 15% of subtitles in our corpus were faster than 20 cps, and almost 8% of subtitles were shorter than one second. We also found processing latencies of around 400 ms for fast subtitles to around 700 ms at speeds of 12 cps, and between 580 ms and 760 ms in different audio conditions. This points to the importance of setting subtitle speed and duration in a way that allows viewers enough time to process both the image and the subtitle properly.

Keywords: ghost subtitles, processing latency, streaming services, subtitle speed

1. Introduction

We have all been frustrated by subtitles at some point. These frustrations depend on user expectations and needs to some extent (i.e. whether the viewer understands or can hear the dialogue in the soundtrack), but typically occur when there are errors or other issues such as subtitle speed and legibility.

Subtitle speed is a complex and heavily debated issue related to both deaf and hearing audiences (c.f. Burnham et al., 2008; Kruger et al., 2022; Liao et al, 2021; Romero-Fresco, 2009; Sandford, 2015; Szarkowska, 2016; Szarkowska & Bogucka, 2019; Szarkowska & Gerber-Morón, 2018; Tyler et al., 2009). In most guidelines there are therefore references to minimum and maximum duration of subtitles as well as average or maximum speeds. This is premised on the assumption that viewers require a certain amount of time to be able to read the subtitles and look at the image.

However, two main interest groups are either reluctant or actively opposed to setting speed limits, or they push for higher speeds. The first is broadcasters

and streaming platforms where the reason can be safely assumed to be largely financial. Bringing down subtitle speed is not something that can be automated reliably, unlike producing verbatim subtitles. This inevitably means that slower speeds are more expensive to produce as human subtitlers have to be engaged to ensure that the reduced text does not result in shifts in meaning. Although there are obvious financial incentives to move to faster speeds, this interest group cites research selectively and relies on anecdotal information from users rather than considering the real cost to the viewer experience (see, for example, Ofcom, 2024). A BBC Research White Paper by Sandford (2015) illustrates the extent to which even the findings of in-house studies are disregarded. After reporting that the participants rated subtitles above approximately 180 words per minute (wpm) to be fast or too fast, and rated enjoyment lower at speeds above 180 wpm, the conclusion is still that "subtitles should match the speech in timing and wording".

The second interest group consists mainly of users such as deaf and hard-of-hearing viewers as well as hearing viewers who use same-language subtitles (i.e., subtitles in the same language as the audio, also known as intralingual subtitles) for whatever reason. These viewers are often strongly in favour of verbatim subtitles, citing reasons such as equitable access and perceived censorship, or consider any mismatch between dialogue and subtitles to be an error regardless of whether the meaning has remained intact. This applies mainly to intralingual or same-language subtitles for the deaf and hard-of-hearing (SDH, also known as closed captioning or CC) as mentioned above, although real or perceived inconsistencies between spoken dialogue and subtitles are also an issue for users of interlingual subtitles (i.e., subtitles in a different language than that of the audio) who rely on subtitles to understand a film in a foreign language.

The evidence base for the reception of non-verbatim subtitles is, with very few exceptions, based on anecdotal accounts, in-house reports by broadcasters, or complaints from users, as well as a handful of surveys (cf. Romero-Fresco, 2009; Sandford, 2015; Szarkowska, 2016). Although the reception of fast subtitles has been investigated (Liao et al., 2021; Szarkowska & Gerber-Morón, 2018), very few studies have engaged with users who are not proficient readers due to lower literacy, cognitive decline, or reading disorders, nor have these studies been done using full-length films. The purpose of this article is not to debate the issue of speed and whether subtitles should be verbatim or adhere to a particular speed, but rather to provide two key findings: one regarding the current subtitle speeds on a major streaming platform based on a corpus study, and the other on the latency between subtitle onset and reading based on empirical evidence from eye movements, which provides some context to the way viewers process subtitles.

The guidelines for English subtitles provided by Netflix (Netflix, 2024) state that the minimum duration of subtitles should be 5/6 (five sixths) of a second. In terms of timing, the guidelines state:

> Subtitles should be timed to the audio or, if necessary, within 3 frames of the audio. If more time is required for better reading speed, the out-time can be extended up to 12 frames past the timecode at which the audio ends.

The guideline on speed aligns with the recommendation of Ofcom (2024), the official broadcasting regulator in the UK, that "in general, verbatim subtitles synchronised with the audio offer the most equitable access to content".

What this change signals is that subtitles no longer have to adhere to speed limits and that verbatim subtitles synchronised with the audio should be the goal. It is likely that other guidelines will follow suit, conceivably also major

streaming platforms like Netflix, although the latter currently still has maximum recommended speeds. According to Netflix's English Timed Text Style Guide (Netflix, 2024), the maximum speed for English subtitles is 20 characters per second (cps). In this article we investigate the distribution of subtitle speed and specifically what percentage of Netflix English subtitles are faster than 20 cps and shorter than 1 second.

2. Subtitle speed: what happens when subtitles are fast?

Subtitle speed (or subtitle presentation rate) is typically expressed either in characters per second (cps) or words per minute (wpm). The conversion between the two measures is not straightforward and is influenced by whether spaces are included in the cps calculation. At 20 cps, a subtitle of 74 characters (two full lines of 37 character each including spaces and punctuation) would need to be displayed for 3.7 seconds. A speed of 20 cps translates to 1200 characters per minute (20*60 seconds). The average word length in English is 5 characters excluding spaces. To convert 20 cps to wpm would therefore mean dividing 1200 characters by 5+1 = 6 (assuming that every word is followed by either a space or a punctuation mark) which comes to 200 wpm. Most calculations, however, use 5 characters per word for this calculation which means that a speed of 20 cps would be 240 wpm (1200 characters per minute divided by 5). Subtitling software can calculate speed in wpm using the actual number of words in a subtitle, but due to varying word lengths, this makes wpm a much less consistent measure of speed (cf. Fresno & Sepielak, 2022). Recommended speeds vary significantly across countries and guidelines and will not be discussed here as the focus is on Netflix, where there are also different recommendations in terms of maximum speed for English and for other languages as well as for children.

In an experiment investigating the reading of subtitles at high speeds using eye tracking[1], Szarkowska and Gerber-Morón (2018) found that viewers could keep up with fast subtitles (20cps), as indicated by the fact that they managed to look at the subtitles and image while maintaining a good comprehension of the video. In the first experiment they recruited 74 participants (27 English, 26 Spanish and 21 Polish) who did not understand Hungarian to watch 3 clips (4-6 minutes) with Hungarian audio and subtitles in their native language. Subtitles were presented at 12, 16 or 20 cps, with each clip at a different speed. The speed and the video clip were counterbalanced, and the presentation order was randomised. Different speeds were created by editing the text at times rather than by just changing the duration, which means that the text in the subtitles was not controlled and could have created a confound. They found no impact of speed on comprehension or scene recognition, but higher speeds resulted in higher self-reported cognitive load (difficulty and effort but not frustration). In terms of eye movements, they found that slower subtitles resulted in more,

[1] Eye tracking is an increasingly popular methodology in studying subtitled film and involves recording the eye movements of viewers while watching, for example, a subtitled film. By analysing fixations (periods when the eyes are relatively still in a specified area and during which visual information is taken in) and saccades (the rapid movements between fixations), researchers can investigate many questions related to the cognitive processes underlying information processing, such as attention distribution among different sources and how written text such as subtitles is processed. The assumption is that where the eyes fixate is strongly related to what is being processed cognitively (known as the eye-mind hypothesis) (see Just & Carpenter, 1980).

longer fixations as well as longer total reading times. Unsurprisingly, they found that the proportional reading time (the time spent looking at the subtitles as a percentage of the time the subtitles were on screen) was highest at the fast rate when the subtitles were on screen for the shortest amount of time. They interpret the higher proportional reading time as indicating more efficient processing, although this is perhaps an overstatement, since it is inevitable that when a certain amount of text is on screen for a shorter period, the proportion of that display time that a viewer will look at the subtitle will be higher than when the text is on screen for longer. It is also somewhat problematic to see a higher proportional reading time as evidence of more efficient processing as this rests on an untested hypothesis. Participants reported being able to cope with all three speeds equally well.

In their second experiment, Szarkowska and Gerber-Morón (2018) tested only two speeds (12 and 20 cps) with Spanish and Polish native language participants with high English proficiency watching two clips in English with native language subtitles (i.e. in this case they could understand the audio). As in the first experiment they found no effect of speed on comprehension. They found no effect on cognitive load (difficulty or effort), but fast subtitles resulted in lower frustration. For English and Spanish participants, enjoyment was higher at the higher speed, and for Polish participants there was no difference. In terms of eye movements, they found the same patterns as in the first experiment with more, longer fixations and higher total time at slower speed and higher proportional time for the higher speed. Again, there was no effect of subtitle speed on the experience of participants in terms of their comfort with the speed. The overall conclusion of the study is that viewers can cope with faster speeds (with faster being 20 cps here). Their finding that faster subtitles are read more efficiently based on higher proportional reading time in this experiment could also be questioned, since overall slower subtitles were read for longer and with more and longer fixations and more revisits which could conceivably also be interpreted as more in-depth and not necessarily more effortful processing.

In another study on subtitle speed (reported in Liao et al., 2021 and Kruger et al., 2022), eye-movement data were collected in a two (video: present, absent) by three (subtitle speed: 12, 20, 28 cps) within-participant design. In this study, 21 English speakers were recruited to watch six video clips (each between 9 and 10 minutes) from a documentary series without sound. Three subtitle speeds were created for each clip by manipulating only the duration of the subtitles. In terms of global measures (i.e. calculated across all subtitles rather than on individual words), Liao et al. (2021) found that higher speeds resulted in significantly fewer and shorter fixations and longer saccades than lower speeds, which suggests that faster speeds result in more skimming and superficial processing. In terms of local or word-level measures, they found that higher speeds resulted in an attenuated word frequency effect[2], which suggests that viewers process words more superficially when subtitle speed increases. In other analyses on the same dataset, Kruger et al. (2022) found that higher speeds resulted in more words being skipped (also linked to longer saccades), more subtitles not being read to the end before disappearing, and fewer words and subtitles revisited. They interpret this as further evidence of superficial processing of subtitles and less integration between subtitles and the background image.

[2] The word frequency effect is a well-documented effect where less common (low frequency) words take longer to process than more common words (see Brysbaert et al., 2018; Rayner et al., 2004).

The evidence from this study, although reported with different emphases in the two articles, point to a general conclusion that faster subtitles are read fleetingly and more superficially, and that fast subtitles are less likely to be read to the end before they disappear. According to Kruger et al. (2022), at 20 cps, 19% of the text at the end of subtitles is not fixated and at 28 cps this goes up to 24%, meaning that almost a fifth to a quarter of the text at the end of each subtitle cannot be processed at the two fastest rates. At these speeds viewers are also unable to reread parts of the subtitles they may not have understood on the first reading (short of rewinding, which is possible but would impact the viewing experience significantly and is unlikely to occur if subtitles are persistently fast). This is specifically relevant in the context of the current article, where we want to determine the distribution of subtitle speed and the incidence of fast subtitles in the corpus.

However, none of these studies looked at the reading of short subtitles in particular. For example, in the experiment conducted by Liao et al. (2021), the subtitle length was between 27 and 55 characters, which was chosen to eliminate very short subtitles and to avoid viewers having to make return sweeps to the beginning of the second line. As will be discussed below, subtitle length takes on a particular relevance when shorter subtitles are considered, and generic maximum speeds will be ineffective in preventing processing problems with shorter subtitles.

3. Methods

3.1 Corpus

To provide some evidence on the actual speed of subtitles that viewers are exposed to, we analysed a large corpus of subtitles from Netflix as the leading streaming platform (Nickinson, 2025). An overview will be provided of the distribution of subtitle speeds in the case of English interlingual subtitles of content from a range of languages in films and series produced by Netflix (i.e. also excluding sound identifiers). As the title of this article suggests, the analyses focus on "ghost subtitles", or subtitles that are on screen for such a short period of time that viewers would be unlikely to be able to read them at all or read the subtitle to completion before it disappears. This is likely to occur with subtitles that are on screen for very short times (e.g., less than 1 second) in relation to the amount of text that needs to be read even though the overall speed in characters per second may be below the maximum recommended speed set by Netflix of 20 cps. A key to understanding this is the fact that there is a latency between the time a subtitle appears on screen and the time the viewer moves their eyes from the background image to the subtitle to start reading. Latency will be discussed in relation to data collected in two eye-tracking experiments on subtitle reading (Liao et al., 2021, 2022).

Kruger et al. (2022) reports on a small sample of subtitled films from Netflix (11 films with a total of 23,356 English subtitles). They found a wide range of speeds (0.8 to 50.4 cps) with an average speed of 12.6 cps and with 15.2% of subtitles faster than 20 cps. To provide a more comprehensive overview of subtitle speed, we built a much larger corpus of English subtitles for the current study. The corpus constitutes all the English subtitles for films in languages other than English (including all languages where there were more than 15 Netflix productions), that were available on Netflix Australia in December 2023. In total we extracted the subtitles for 1449 videos (films and only the first episode of series – in order not to skew the data towards series) from 17 languages. The total duration of the videos was 46,825 minutes.

3.2 Latency

For this article, latency is defined as the time it takes a reader to start reading the subtitle after it has appeared on screen. This latency means that viewers have less time to read subtitles than the time the subtitle is on screen. To get an idea of the average latency, we analysed data from two eye-tracking experiments reported by Liao et al. (see Liao et al., 2021 and 2022 for detailed description of the experiments and other results). Our analyses in this paper only consider eye-movement data in the video-present condition, although the experiments also included a condition where the subtitles were displayed without video.

In the speed experiment, English first language participants were used, and English subtitles were presented at three speeds without audio: 12, 20 and 28 cps. This experiment thus provides evidence on latency for English viewers reading first-language (L1) subtitles, particularly when the absence of audio creates a strong need for the subtitles. Different display times at different speeds also create varying degrees of urgency to read the subtitles before they disappear. The hypothesis is that as the urgency for reading the subtitles fast increases (i.e. as the subtitle speed increases), the latency should decrease, with viewers moving their eyes to the subtitles more quickly to maximise their chances of reading them before they disappear.

In the audio experiment, Chinese first language participants who used English as their second language were recruited to watch videos with English (L2) subtitles. The audio was manipulated into three conditions: Chinese (L1) audio, English (L2) audio, and no audio. Unlike the speed experiment in which the speed was consistent throughout a video, in the audio experiment the speed varied from 6 to 18 cps, with an average of 12 cps and a standard deviation of 1.5 cps). This experiment therefore provides evidence on second language speakers of English reading subtitles in a more authentic context when different audio conditions create different needs for the subtitles (e.g., there is a low propensity to read subtitles when the L1 audio provides easier access to the L2 spoken dialogue). The hypothesis is that the latency would decrease from Chinese to English to no audio as the need for the subtitles increases.

Table 1: Basic information about the two eye-tracking experiments.

Experiment	Number of Participants	Experimental Conditions	Number of Subtitles
Speed Exp. (Liao et al., 2021)	75 English native speakers (aged 18-56, average: 20; 14 males)	Within-subject design of two manipulated variables: video presence (with or without video) and subtitle speed (12, 20 and 28 cps).	490
Audio Exp. (Liao et al., 2022)	34 Chinese native speakers who were advanced speakers of English (aged 20 to 38, average: 26; 8 males)	Within-subject design of two manipulated variables: video presence (with or without video) and audio condition (without audio, English audio, Chinese audio. Subtitles had an average speed of 12.35 cps (SD = 1.54; range: 6–18 cps).	503

Note: Eye movements from nine participants from the speed experiment and three participants from the audio experiment were excluded in the analyses due to calibration issues.

3.2.2 Apparatus

Eye movement data was recorded using an Eyelink 1000+ eye tracker (SR Research Ltd., Canada) with a sampling rate of 2000 Hz and a screen resolution of 1920 x 1080 pixels. Videos were presented at a resolution of 1280 x 720 pixels at 30 frames per second in the centre of the screen with subtitles presented below the screen in 30-point Courier New font. Participants were positioned 95 cm from the screen and their heads were stabilized using a chin rest to minimize head movement. Only the right eye was tracked. A 9-point calibration and validation were performed before each video with a maximum calibration error of 0.5°.

3.2.3 Analyses

To investigate latency, the time between the onset of a subtitle and the first fixation in the subtitle that was followed by a forward saccade was calculated. This was done to determine the exact point at which reading began by excluding fixations because of incorrect landing positions (i.e. where the eyes landed at a non-optimal viewing position and the participant had to make a regression to an optimal subtitle reading position to start reading).

Data were analysed using linear mixed models (LMMs), specifically the lme4 package (version 1.1-23) in R (Version 3.6.3); p values were computed via the *lmerTest* package (Version 3.1-2, Kuznetsova et al., 2017). For the subtitle speed experiment, speed was entered as a fixed effect, with participant and subtitle item being entered as random effects. For the audio experiment, audio condition was treated as a fixed effect, with participant and subtitle item being random effects. For the analyses for both experiments, sliding difference contrasts were used via the *contr.sdif* function, which compares consecutive factor levels of each variable. The *emmeans* package (version 1.47) was used to compute and extract the estimated means between different levels and the simple effects. When fitting the model, we started with a maximal structure (i.e., with both random intercepts and random slopes in the random effects). Insignificant random effects were removed to simply the model following the Parsimonious Mixed Model approach (Bates et al., 2015). Details of the final models are provided in Appendix 1 and 2.

4. Results

4.1 Corpus results

The total number of subtitles was 1,291,741. The distribution across the languages can be seen in Table 2 below. Across the corpus, the subtitle speed ranges from 0.19 cps to 75.25 cps, with an average speed of 14.79 cps (SD: 5.28) (see Table 3 and Figure 1 below).

Table 2: A summary of the corpus

Source Language	Number of Videos	Number of Subtitles	Percentage of total subtitles (%)	Average video duration (mins)	Total video duration (mins)
Danish	19	16676	1.29	62.07	1179.29
Dutch	17	13386	1.04	81.06	1378.04
French	124	122364	9.47	72.53	8993.70
German	70	57812	4.48	68.99	4828.97
Hindi	93	91912	7.12	82.04	7629.80
Indonesian	22	23708	1.84	100.20	2204.32
Italian	69	63163	4.89	74.19	5118.94
Japanese	164	101761	7.88	49.52	8122.02
Korean	172	156633	12.13	69.76	11998.47
Mandarin	42	40785	3.16	76.81	3226.18
Norwegian	19	16122	1.25	73.64	1399.23
Polish	57	52030	4.03	80.29	4576.51
Portuguese	91	84129	6.51	61.55	5600.76
Spanish	391	359526	27.83	68.22	26674.51
Swedish	25	20984	1.62	65.01	1625.23
Thai	22	17683	1.37	76.55	1684.07
Turkish	52	53067	4.11	79.01	4108.66

Table 3: Mean and standard deviation (SD) of speeds (cps) across languages.

Language	Mean	SD
Danish	14.06	5.10
Dutch	13.75	5.37
French	15.28	4.92
German	14.41	4.95
Hindi	14.82	5.63
Indonesian	13.61	5.17
Italian	14.61	5.04
Japanese	13.17	5.08
Korean	15.24	5.71
Mandarin	15.48	6.15
Norwegian	13.87	5.56
Polish	14.69	5.31
Portuguese	14.94	4.81
Spanish	15.06	5.13
Swedish	13.84	5.11
Thai	13.68	5.18
Turkish	15.57	5.28

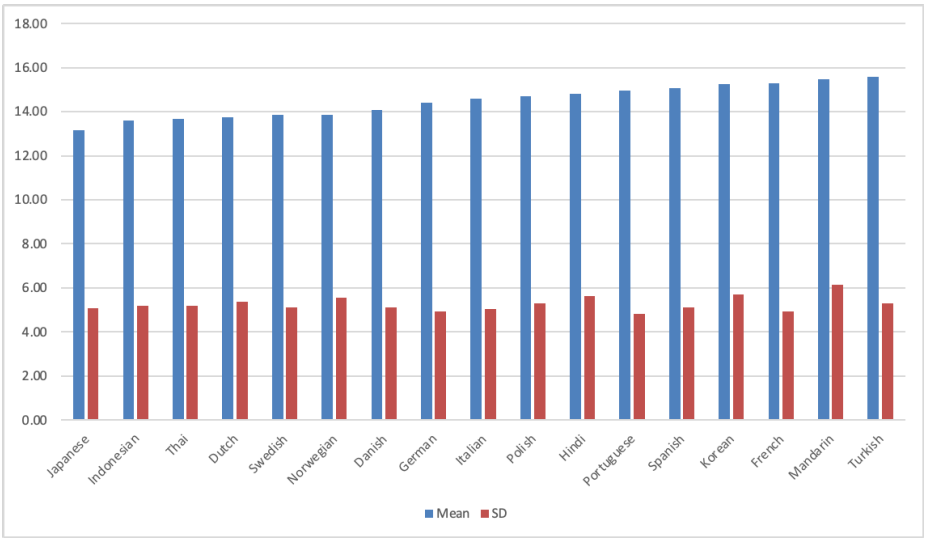

Figure 1: Average speed across different languages (SD: standard deviation; Y axis represents subtitle speed in cps, while x axis represents language).

In terms of the distribution of subtitle speeds, 14.85% of the subtitles were faster than 20 cps (see Figure 2). There were 7.14% of subtitles shorter than 1 second and 0.06% shorter than 0.83 second — the minimum duration in Netflix guidelines. Among subtitles shorter than 0.83 second, the average subtitle length was 8.8 characters (SD: 5.09), with a wide range from 3 to 43 characters; the average speed was 11.6 cps (SD: 6.76), with a range from 3.75 to 57.5 cps.

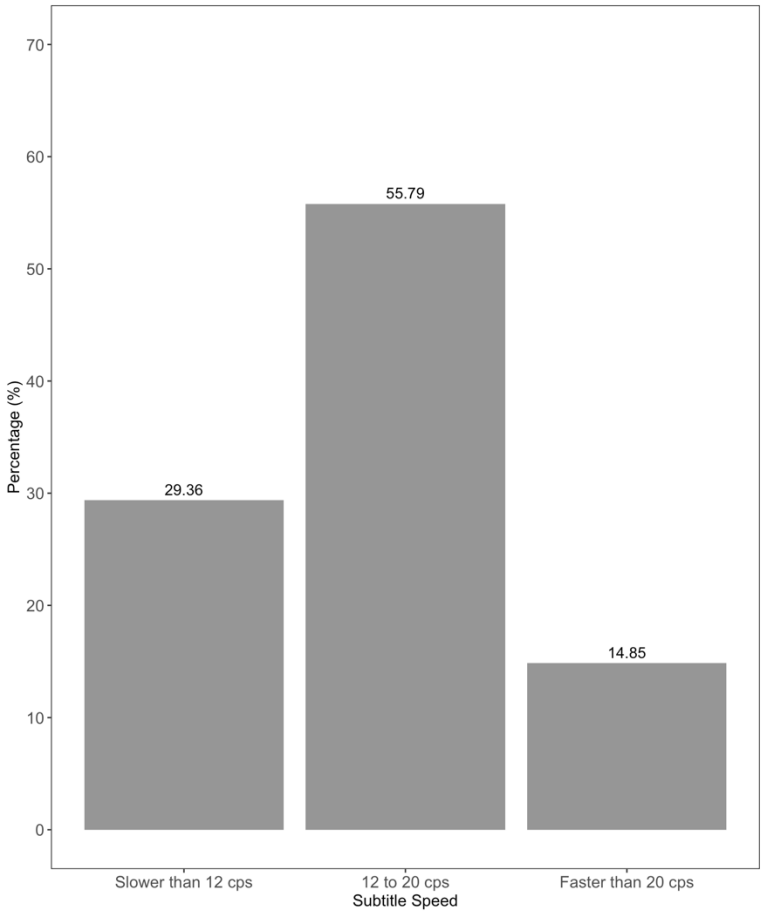

Figure 2: Distribution of subtitle speeds at different ranges.

Comparing languages, we note a significant difference in the percentage of fast subtitles (see Table 4 and Figure 3). This might reflect a difference in conventions in different languages or a difference in speech rate although this is mere speculation and would need to be verified in further studies. Results of the linear mixed effect model show that Mandarin has faster speed compared to most other languages (Danish, Dutch, German, Indonesian, Italian, Japanese, Norwegian, Polish, Portuguese, Swedish and Thai) (all *p*s < 0.05).

Table 4: Percentage of subtitle speeds faster than 20 cps across languages (subtitles faster than 20 cps include those less than 1 second).

Language	Over 20 cps (%)	Less than 1 second (%)
Danish	11.66	7.22
Dutch	10.94	8.45
French	15.04	7.54
German	10.59	8.58
Hindi	16.54	8.55
Indonesian	9.81	8.40
Italian	13.04	6.35
Japanese	8.08	7.26
Korean	19.98	6.50
Mandarin	21.84	9.10
Norwegian	12.32	9.13
Polish	13.91	8.22
Portuguese	13.21	5.63
Spanish	15.35	6.43
Swedish	10.29	7.30
Thai	9.82	10.43
Turkish	19.47	6.66

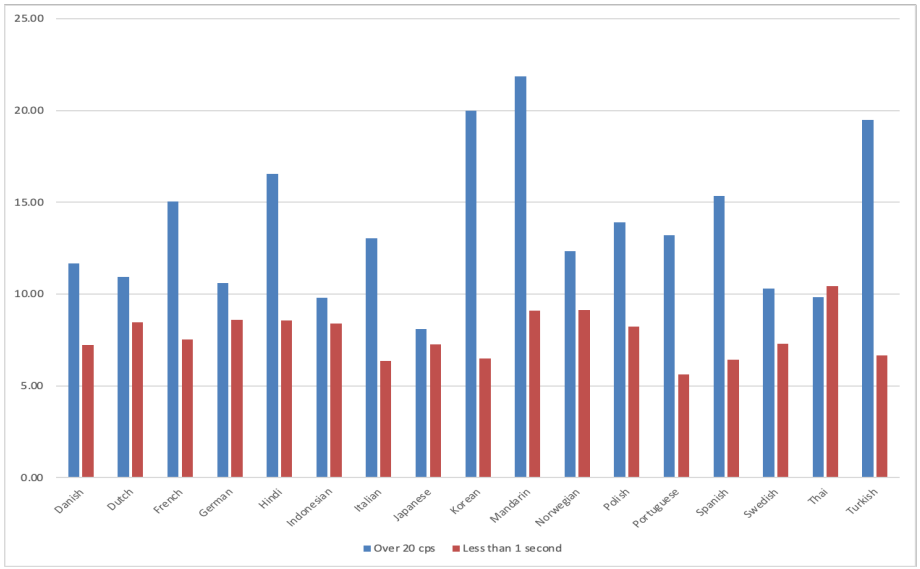

Figure 3: Percentage of subtitle speeds faster than 20 cps across different languages. Subtitles faster than 20 cps include those less than 1 second.

4.2 Subtitle processing latency: Evidence from eye movements

4.2.4 Results

As shown in Table 5 and Figure 4, subtitle speed had a significant impact on latency. Participants in the 20 cps and 28 cps conditions had shorter processing latency than in the 12 cps condition (*p*s < 0.001). There is no significant difference between 20 cps and 28 cps ($p = 0.29$).

Audio also had a significant impact on latency: participants with Chinese (L1) audio had longer processing latency compared to those with English (L2) audio or without audio (*p*s < 0.05). There is no significant difference between English-audio and no-audio conditions ($p = 0.97$) (see Table 5 and Figure 5).

Table 5: Mean and standard deviations of latency (ms) in different conditions.

Speed Experiment		Audio Experiment	
Condition	Mean (SD)	Condition	Mean (SD)
12 cps	716.44 (337.95)	No Audio	618.88 (278.55)
20 cps	396.84 (90.66)	Chinese Audio	764.07 (206.86)
28 cps	374.97 (60.67)	English Audio	579.70 (182.73)

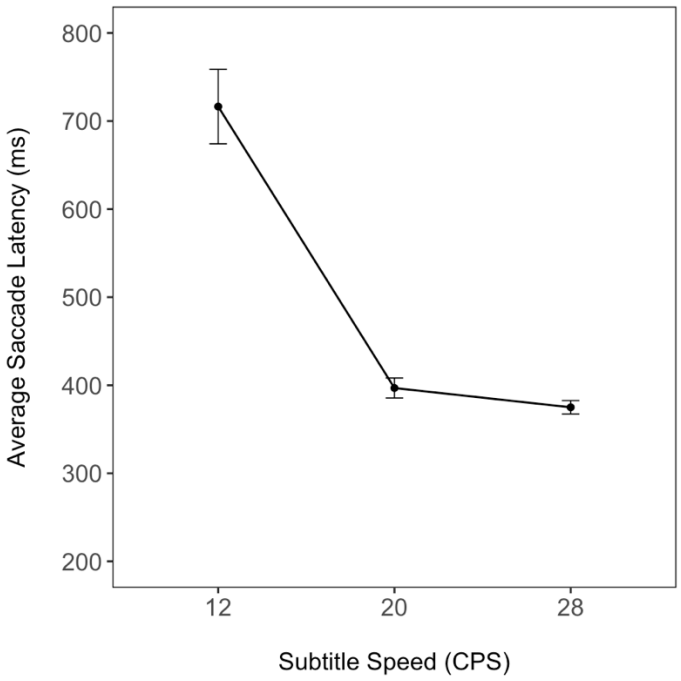

Figure 4: Average latency as a function of subtitle speed when audio was absent.

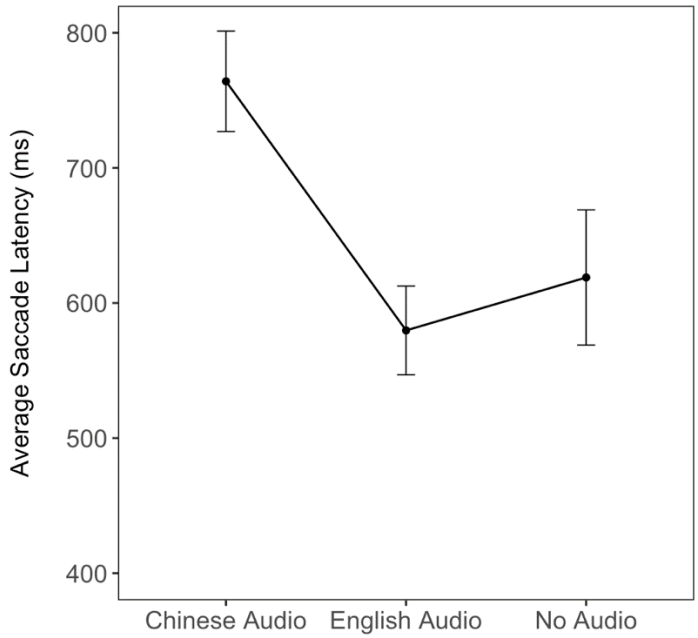

Figure 5: Average latency in different audio conditions with varying speeds.

5. Discussion

The corpus analysis of English subtitles on content in languages other than English clearly shows that, although all the languages have a similar average speed of around 15 cps, an average of 14.85% of the subtitles are faster than 20 cps. In other words, the maximum speed is exceeded in a large number of subtitles across all languages. There is also a noticeable variability across the different languages: Korean, Mandarin and Turkish (constituting around 20% of the corpus) have around 20% of subtitles faster than 20 cps, with the lowest percentage of subtitles faster than 20 cps being in the range of 8 to 12 % (in Dutch, Danish, German, Indonesian, Japanese, Norwegian, Swedish and Thai).

When looking at the percentage of fast subtitles, with 7.14% of subtitles shorter than 1 second across the corpus and with this percentage being higher (between 8 and 10.5%) in 8 of the languages (Dutch, German, Hindi, Indonesian, Mandarin, Norwegian, Polish and Thai), the percentage of subtitles that could potentially fall in the "ghost subtitles" category of subtitles that could disappear before viewers had the time to start reading them is therefore high. Although fewer than 1% of subtitles were shorter than 0.83 seconds, the average subtitle length for those subtitles was 8.8 characters (ranging from 3 to 43 characters), meaning that these were often longer than one word. Considering the latencies in our two experiments (particularly the average of more than half a second in the second experiment), viewers would therefore have less than half a second to read subtitles shorter than one second. In the case of subtitles shorter than 0.83 seconds, they would either arrive at the location of the subtitle just as it disappears, or they would have about a quarter of a second to read the subtitle. The average fixation duration in silent reading is between 225 and 250 ms (Rayner, 2009), and in subtitle reading similar durations have been found in some studies (see Bisson et al., 2014), although shorter durations have also been reported in subtitle reading of between approximately 180 and 200 ms (see Liao et al., 2021; Szarkowska & Gerber-Morón, 2018). In other words, any subtitle longer than a single word on screen for less than a second would be unlikely to

allow for more than one or two fixations, meaning that any of these brief subtitles with more than one word would be difficult to process fully.

Given the reality of subtitle speeds in our corpus, the results from the analyses of latency take on particular significance. In the case of the speed experiment, it is evident that, at consistently high speeds, viewers adopt a strategy of moving their eyes to the subtitles significantly faster with a latency of only 397 ms and 375 ms at speeds of 20 and 28 cps respectively compared to 716 ms when the speed was consistently 12 cps. This supports our hypothesis that higher speed would result in shorter latencies. The results show that viewers tend to adapt their eye-movement routines to prioritise text processing when there is an increased demand on reading with reduced time. This means that the viewers have to compromise their processing time for the visual content, which may impair their viewing experience. This is also frustrating news to filmmakers considering the vast amount of time that they devote to building the visual narratives. Given that the speed did not vary in this experiment, participants could indeed adopt this strategy in response to the demands. If we consider that around 7% of subtitles in our corpus were shorter than one second, this still means that even at consistently high speeds, participants would have just more than half a second to read 7% of the subtitles.

Since most viewers would watch translated subtitles with sound, the results of the analyses of latency in the second experiment are perhaps more telling. In the two conditions where participants would have a higher reliance on subtitles (without audio, and with English audio), the latency was 619 ms and 579 ms respectively (with no significant difference between these two conditions). With Chinese audio, where the reliance on subtitles would be lowest, the latency was significantly higher at 764 ms, which is comparable to the 716 ms latency at a speed of 12 cps in the first experiment. These findings also support our hypothesis that higher reliance on subtitles would result in shorter latencies. What this means is that regardless of their reliance on the subtitles, participants in this experiment would have between a third and a quarter of a second to process subtitles shorter than a second.

6. Conclusion

What our results from the corpus analyses reveal is that a large percentage of English subtitles on Netflix in a range of languages is even faster than the maximum recommended speed of 20 cps. Likewise, a large number of subtitles are also on screen for less than one second. Even if this guideline on maximum speed and minimum duration were to stay in place (although indications from the Ofcom discussion paper already signals a move towards the removal of maximum speeds), many subtitles would therefore not be possible to process fully or at all.

The results of the latency analyses underscore this conclusion. For very brief subtitles, latencies in excess of half a second would mean that viewers are likely to miss a significant number of "ghost subtitles". A significant percentage of faster subtitles also means that viewers would be unlikely to be able to fully process longer "ghost subtitles" when taking latencies of more than half a second into account.

We have focused on two of the extremes, namely fast subtitles and very brief subtitles. What these analyses do not consider is the number of subtitles that are longer than one second and slower than 20 cps, but that would also become very challenging to process fully due to the latency. One concrete example would be a subtitle of 24 characters (4 words of 5 characters each with spaces and punctuation) at a speed of 18 cps which would be on screen for 1.3

seconds. If we factor in a latency of 653 ms (the average of the latencies in the sound experiment), this will mean that a viewer would have 677 ms to read the subtitle. Even one unfamiliar word in such a subtitle would slow down the reading (due to the well-established word-frequency effect), meaning that the subtitle is likely to disappear before it could be read in full. Such incomplete processing is likely to impact comprehension and enjoyment, although this still needs to be confirmed empirically (but see Kruger et al., 2022 for evidence of incomplete reading of fast subtitles). One way to counter this would be to introduce diversified speeds based on the volume of text taking latency into account. This could be done by simply subtracting a latency (of at least 500 ms) from the duration of each subtitle before calculating the speed. With recent developments in AI, this would be simple to implement. Word frequency could also be factored into the calculation of the required speed for a particular subtitle.

Most eye tracking studies on the reading of subtitles have been performed on skilled readers (mostly university students). It is fair to assume that readers with lower reading proficiency (including viewers with reading disorders like dyslexia, beginner readers, readers in a foreign language, older readers with cognitive decline, and some deaf viewers) would have even more difficulty processing faster subtitles which are also likely to result in higher fatigue and reduced comprehension and enjoyment. These are of course empirical questions that need to be investigated in future research.

Our study also did not consider other factors that might impact latency, such as the volume of text and visual complexity. It is conceivable that viewers would be able to judge the volume of text in a subtitle in their peripheral vision and move their eyes to the subtitle faster if there is more text to process. It also stands to reason that when engaged in processing parts of a video with higher visual complexity (due to editing, screen composition, contrast, movement, etc.), latency would increase, and viewers would be more likely to miss brief subtitles. These, too, are empirical questions that need to be investigated in future research.

A limitation of our study is that we only analysed English subtitles on videos in a selection of languages in Australia. Future research would need to investigate whether similar patterns can be observed in other languages across the globe. This study is also limited in that the one experiment reported did not include sound, and the second experiment only included Chinese participants reading English subtitles. More research is required to investigate whether these latencies also occur in different populations reading subtitles in different languages.

Based on our results, we can conclude that subtitle speed on at least this major streaming platform requires careful consideration, and there is no reason to believe that the situation would be significantly different on other platforms or broadcasting networks. Our findings on latency provides evidence to inform these considerations. Viewing video with subtitles is a cognitively demanding activity in which various cognitive processes must be coordinated, making it imperative for providers who value access and customer satisfaction to consider more than just the financial impact of subtitling solutions. Allowing viewers time to enjoy all aspects of AV content will benefit viewers but also allow content to reach its potential.

References

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. arXiv preprint. https://doi.org/10.48550/arXiv.1506.04967

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science, 27*(1), 45-50. https://doi.org/10.1177/0963721417727521

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

Fresno, N., & Sepielak, K. (2022). Subtitling speed in Media Accessibility research: some methodological considerations. *Perspectives, 30*(3), 415-431. https://doi.org/10.1080/0907676X.2020.1761841

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

Kruger, J.-L., Wisniewska, N., & Liao, S. (2022). Why subtitle speed matters: Evidence from word skipping and rereading. *Applied Psycholinguistics, 43*(1), 211–236. https://doi.org/10.1017/S0142716421000503

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software, 82*, 1-26.

Liao, S., Yu, L., Kruger, J.-L., & Reichle, E. D. (2022). The impact of audio on the reading of intralingual versus interlingual subtitles: Evidence from eye movements. *Applied Psycholinguistics, 43*(1), 237–269. https://doi.org/10.1017/S0142716421000527

Liao, S., Yu, L., Reichle, E. D., & Kruger, J.-L. (2021). Using Eye Movements to Study the Reading of Subtitles in Video. *Scientific Studies of Reading, 25*(5), 417–435. https://doi.org/10.1080/10888438.2020.1823986

Netflix. (2024). Timed text style guide: General requirements. https://netflixtechspark.zendesk.com/hc/en-us/articles/360001200447-Timed-Text-Style-Guide-General-Requirements

Nickinson, P. (2025, February 1). *The 10 most popular streaming services, ranked by subscriber count.* Digital Trends. https://www.digitaltrends.com/home-theater/most-popular-streaming-services-by-subscribers/#:~:text=The%2010%20most%20popular%20streaming%20services%2C%20ranked%20by,8%208.%20ESPN%2B%3A%2026%20million%20...%20More%20items

Ofcom. (2024, 15 April). *Ensuring the quality of TV and on-demand access services: Statement on changes to the TV Access Services Code and updates to Ofcom's best practice guidelines on access services.* https://www.ofcom.org.uk/siteassets/resources/documents/consultations/category-1-10-weeks/264291-ensuring-the-quality-of-tv-and-on-demand-access-services/associated-documents/statement-on-ensuring-the-quality-of-tv-and-on-demand-services.pdf?v=356711.

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. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*(8), 1457–1506.

Romero-Fresco, P. (2009). More haste less speed: Edited versus verbatim respoken subtitles. *Vigo International Journal of Applied Linguistics*, (6), 109-133.

Sandford, J. (2015). *The impact of subtitle display rate on enjoyment under normal television viewing conditions.* (BBC Research White Paper WHP306). British Broadcasting Corporation. https://digital-library.theiet.org/doi/10.1049/ibc.2015.0018

Szarkowska, A. (2016). *Report on the results of an online survey on subtitle presentation times and line breaks in interlingual subtitling. Part 1: Subtitlers* (Report). University College London.

Szarkowska, A. (2018). *Report on the results of the Sure Project study on subtitle speeds and segmentation*. University College London.

Szarkowska, A., & Bogucka, L. (2019). Six-second rule revisited: An eye-tracking study on the impact of speech rate and language proficiency on subtitle reading. *Translation, Cognition & Behavior, 2*(1), 101-124. https://doi.org/10.1075/tcb.00022.sza?locatt=mode:legacy

Szarkowska, A., & Gerber-Morón, O. (2018). Viewers can keep up with fast subtitles: Evidence from eye movements. *PloS one, 13*(6) https://doi.org/10.1371/journal.pone.0199331

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

Appendix 1. Model details in the speed experiment

Fixed Effects				
	Est/Beta	SE	t	p
Intercept	5.97	0.02	269.67	<0.001
Speed (20-12)	-0.23	0.02	-10.82	<0.001
Speed (28-20)	-0.03	0.02	-1.50	0.13
Speed (12-28)	0.26	0.02	12.31	<0.001
Random Effects				
	Variance	S.D.		
Item (Intercept)	0.02	0.15		
Subject (Intercept)	0.02	0.14		
Model fit				
R^2	Marginal	Conditional		
	0.04	0.18		
Final model: Speed + (1 | Subject) + (1 | Subtitle Item) Number of observations: 3824				

Appendix 2. Model details in the audio experiment

Fixed Effects				
	Est/Beta	SE	t	p
Intercept	6.18	0.03	203.98	<0.001
Audio (EA-CA)	-0.23	0.05	-4.69	<0.001
Audio (NA-EA)	0.02	0.07	0.25	0.800
Audio (CA-NA)	0.22	0.06	3.93	0.002
Random Effects				
	Variance	S.D.	Correlation	
Item (Intercept)	0.06	0.25		
Subject (Intercept)	0.02	0.13		
Subject (slope) (EA-CA)	0.04	0.19	0.27	
Subject (slope) (NA-EA)	0.09	0.30	-0.32	-0.76
Model fit				
R^2	Marginal	Conditional		
	0.022	0.211		
Final model: Audio + (Audio | Subject) + (1 | Subtitle Item) Number of observations: 2109				