THE IMPACT OF SUBTITLE DISPLAY RATE ON ENJOYMENT UNDER NORMAL TELEVISION VIEWING CONDITIONS

J. Sandford

The British Broadcasting Corporation, UK

ABSTRACT

One of the properties often identified as having an impact on the television viewing experience for subtitle users is the rate of subtitling (measured in words per minute) (1,2). Previous studies on the subject have often restricted participants from using residual hearing or lip-reading as they naturally would when viewing television (3,4,5,6). Additionally, some studies were carried out with potentially biased participants (5,6). No research has been done to date at a large scale on the rate of scrolling subtitles as are often used in live subtitling (5,6).

This paper presents the results of a study examining the impact of subtitle display rate on enjoyment for a representative sample of subtitle users. Specially created and off-air material was used with both block and scrolling subtitles. Sound was available and lip-reading was possible. The results challenge previous assumptions.

INTRODUCTION

The rate of subtitles is often highlighted in subtitling guidelines as an important factor in viewer understanding and enjoyment (2), but no scientific justification is provided. When the few papers currently available on the subject of subtitle rate are examined, it becomes apparent that the quality of previous research is poor and findings vary wildly as a result. Furthermore, research in the field repeatedly fails to use un-biased regular subtitle users (i.e. people who use subtitles once a day or more) as participants and fails to use normal television viewing conditions (3,4,5,6).

This paper presents the findings of a new study that improves on previous work and answers some questions while questioning the validity of others.

- What is the ideal rate of subtitles for subtitle users?
- At what subtitle rates is enjoyment diminished for subtitle users?
- How do these rates compare to the enjoyment of speech at various rates for hearing viewers?
- Does the rate of subtitles even have an impact on enjoyment?

BACKGROUND

Subtitle Rate Measurement

Subtitle rate (also known as the speed of subtitles) is most often measured in Words Per Minute (WPM). This may be calculated in a number of ways. The most common method

used is to take an average over a period of time by dividing the number of words in a clip or programme by its length in minutes. This method is used in much of the available academic literature. While this is a simple method to implement, it may provide low readings for clips with long periods without speech. This may be accounted for by excluding long periods of silence from calculations. Studies generally choose their clips carefully to avoid this problem.

The measurement of rate in this study used this method. Clips in part 1 of the study had no periods without subtitles. Periods without subtitles in the clips in part 2 were excluded from calculations.

Subtitle Rate in Guidelines

Guidelines often quote optimal and maximum rates for subtitles. Figures of approximately 140WPM as the optimum subtitle rate and around 180-200WPM as the maximum rate are common. The guidelines examined fail to cite research supporting these figures but justify them by stating that above these rates, subtitles will be difficult to follow (2).

Prior Research

The small amount of published research on subtitle rate varies wildly in quality. Participants are sometimes selected from biased or non-representative groups. These include people who work in subtitling, people who do not use subtitles and people from interest groups who may subconsciously aim to represent the standard views of their group (5,6,7). Many studies also purposefully aim to reduce experimental influences to the subtitles alone by using footage without those speaking in shot or by using clips without audio (3,4,5,6). This has the un-desirable side effect of creating an un-natural viewing experience. Viewers normally use visual cues such as lip-reading or facial expression to support subtitles. Most subtitle users also have some level of hearing and thus use subtitles in conjunction with audio. Viewers' experience is a combination of these sources of information.

Previous research has shown no drop in comprehension at rates of at least 230WPM (3,4,7), far higher than the maximum rates in current guidelines. One study which aimed to find the most enjoyable rate of subtitles identified a speed of 145WPM, which is approximately the average speed of American subtitles found in a study conducted by the same researcher (3,8). However, the materials in this study used footage without people speaking in shot and without audio.

Requirements for This Study

In order to identify maximum, minimum and optimum subtitle rates, this study built upon the method of Jensema 1998 (3). Participants were presented with clips of a range of speeds and asked to rate the speed and their enjoyment of the subtitles. In addition, a control group of hearing viewers were asked to rate the speed and their enjoyment of the speech on the same clips but without subtitles. Where Jensema used un-natural clips, this study replicates normal television viewing conditions with the speakers in shot and audio available. The two main display methods of subtitles, block and scrolling (also known as word-at-a-time), were tested. To identify if these results held for real-world content, a range of off-air (broadcasted) clips identified as being far faster than current guidelines were also tested.

STUDY – PART 1

For part 1 of this study, clips were created at a range of rates with both block and scrolling subtitles. Subtitle users were asked to rate these in terms of speed and enjoyment. This would allow the optimum rate of subtitles and rates at which subtitles become too fast or too slow to be identified. The same rates would also be identified for speech for hearing viewers.

Materials

24 news clips were created for this study. A local news team, studio and presenter were used to keep the style consistent, familiar and realistic. Both recorded audio and video were used in the clips with no other audio or visual content added.

All clips were 30 second long to eliminate differences in fatigue between clips. The number of words was therefore changed to alter the rate (e.g. a 170WPM clip would be scripted to 85 words). Speech and subtitles were correctly aligned in all clips to eliminate the confusing effects introduced by differing speech and text for those who used lip-reading or audio in conjunction with the subtitles. Rates of 90, 110, 130, 150, 170, 190, 210 and 230WPM were used which approximate those used in Jensema 1998 (3). Three sets of 8 clips were created with one set shown with block subtitles, one with scrolling and one used for introductory materials for each participant. Having 3 clips at each rate would also reduce any effects of individual scripts. Realistic scripts were created by identifying local news stories consisting of approximately the number of words required and rewording sections to make the length correct. The stories were purposely chosen from regions other than the one the study was conducted in to reduce the possibility of familiarity with the stories. A 3 second still of a black background with a dark logo was displayed before and after each clip to allow the participant to comfortably switch between rating and viewing clips.

Where subtitles were used, their style and layout was matched to that of BBC News. Splitting of lines was based on the maximum number of characters allowed in a subtitle and not on grammatical boundaries to avoid effects of artistic choices and matched news subtitling styles. All clips were subtitled verbatim and had versions with both block and scrolling subtitles produced. In the case of scrolling subtitles, each word was introduced on its first spoken utterance. In the case of block subtitles, each subtitle was introduced on the first spoken utterance of the first word. The final subtitle of each clip was removed at the point the final utterance finished. There were no breaks in the subtitles. The speed of each clip used in analysis was the measured rate, not the target rate. Where clips were shown without subtitles, the subtitled rate and not the spoken rate (which may be higher for numbers etc.) was used in analysis to allow direct comparisons to be made with the subtitled clips.

Methodology

The study was conducted with the participant seated at a distance of approximately 5H from the television (where H is the height of the television) (9). A table was provided with a mouse to allow the participant to interact with the user interface for questions presented on the television. The television remote was also provided for setting the volume.

For each participant the three sets of videos were assigned as introductory material, block subtitles and scrolling subtitles. Additionally, half of the participants were shown their block

subtitles set first and half their scrolling subtitles set first. Participants were first shown the 150WPM clip in their introductory set with block subtitles and asked to set the volume of the television as they would normally have it when watching television with subtitles. For each main set, they were then shown 3 introductory videos of slow (90 or 110WPM), medium (170 or 190WPM) and fast (210 or 230WPM) rates in that order with the subtitling format (scrolling or block) of that set. The lower speed shown with the first set. Participants were asked to rate each clip in terms of speed ("Too slow" to "Too fast") and enjoyment ("Not enjoyable" to "Very enjoyable") of the subtitles on continuous scales with labels at each end only. Each question was displayed separately to reduce cross-rating interference. The clips in the main set were then shown ordered according to a Latin-square. Participants were asked to rate these clips as before.

A control group of hearing participants were asked to set the volume and shown introductory content as with the subtitles group but were only asked to rate one other set of clips, not two. Any wording on screen that referred to subtitles in the main group referred to speech in the control.

Participants

25 frequent subtitle users were recruited through an external agency. A split of male and female participants were recruited along with a spread of ages, hearing impairments and social grades. All were regular users of subtitles as an access service and were familiar with televised BBC News content. No participants who use sign language as a first language were recruited. First language BSL users, who make up around 8% of hearing impaired people in the UK (10), may be seen as second language users and will require a specific detailed study. A convenience sample of 16 hearing participants not involved in production quality was recruited from BBC North at MediaCityUK for the control group.

Results

Figures 1 & 2 show the ratings for speed and enjoyment for both groups of participants. Table 1 shows exact values where mean rate is on a scale where 1 is "Too slow" and 5 is "Too fast". "Slow", "Good" and "Fast" rates are taken from a linear regression at ratings of 2, 3 and 4 respectively. Mean enjoyment is on a scale where 1 is "Not enjoyable" and 5 is "Very enjoyable". Peak enjoyment was found with a quadratic regression.
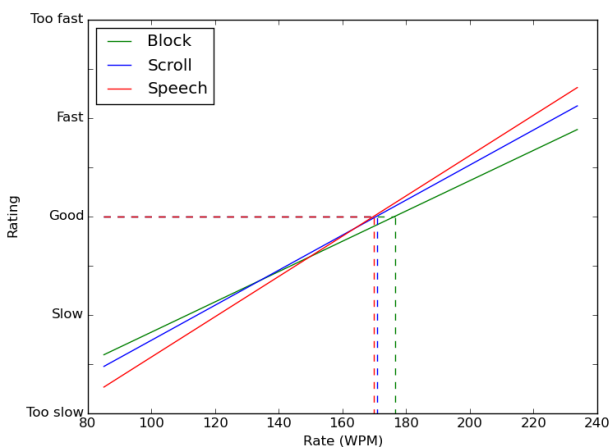

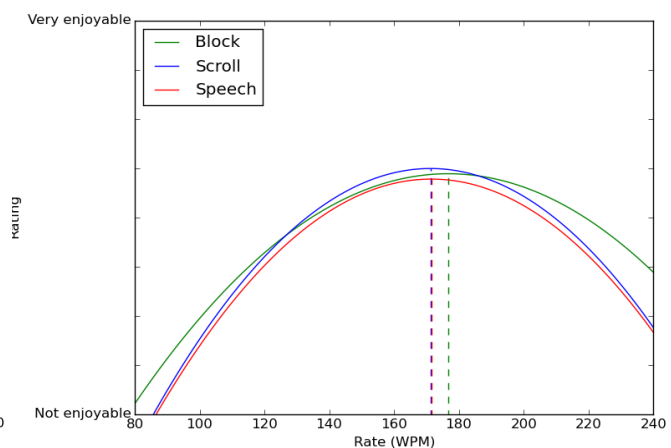Figure 1 – Perceived rate of subtitles and speech against measured rate

Figure 2 – Perceived enjoyment of subtitles and speech against measured rate

	Rate R^2	Rate mean (1-5)	Rate "Slow" (WPM)	Rate "Good" (WPM)	Rate "Fast" (WPM)	Enjoy R^2	Enjoy mean (1-5)	Enjoy peak (WPM)
Block	0.64	2.77	112	177	242	0.35	2.86	177
Scrolling	0.68	2.84	115	171	227	0.42	2.76	171
Speech	0.77	2.83	121	170	219	0.39	2.67	171

Table 1 – Study part 1 results for perceived rate and enjoyment

The optimum ("Good") rate was found to be highest for block subtitles and was approximately the same for scrolling subtitles and speech. The range of rates between "Fast" and "Slow" was widest for block subtitles and narrowest for speech. However, the overall similarity between all of these results demonstrates that the rate of subtitles is not an issue under the conditions tested. When the rate of speech is perceived to be bad, the rate of subtitles is also perceived to be bad. Also, when the rate of speech is perceived to be good, the rate of subtitles is perceived to be good.

STUDY – PART 2

Part 1 of the study has shown that, under the conditions tested, the rate of subtitles is not an issue. Part 2 aims to explore if this remains true for broadcast content. A range of clips from broadcast content identified as having subtitle rates above 200WPM were selected to see how their ratings compare to the material in part 1.

Materials

8 clips identified using a monitoring system were selected to cover a range of programming styles. These contained differing numbers of people talking and varying shots such as close-ups, long shots and shots of people/content other than the person talking. The clip lengths, styles, and mean instantaneous subtitle rates for each clip are shown in Table 2.

The subtitles for these clips were shown as they were presented when broadcast. All timing, styling and positioning was maintained including any inaccuracies. All clips had pre-prepared block subtitles. No live scrolling subtitles were used due to the nature of how live subtitles are created. They are often created by a single subtitler speaking the subtitle content into speech recognition software and carrying out minor formatting with a keyboard. This method means live subtitles rarely reach the highest subtitling rates.

Methodology

These clips were shown to the regular subtitle users immediately after part 1. The same rating system was used and the clips were shuffled using a Latin-square.

Results

Figures 3 & 4 show the ratings of part 2 against the regression lines calculated in part 1. Exact values are shown in Table 2. All mean perceived rates are closer to "Good" than the prediction from part 1 and well under the "fast" mark. The mean perceived rate of 3 clips subtitled at approximately 230WPM fall within 1% of a perceived perfect rate. Mean enjoyment is also well above the prediction from part 1 for all clips.
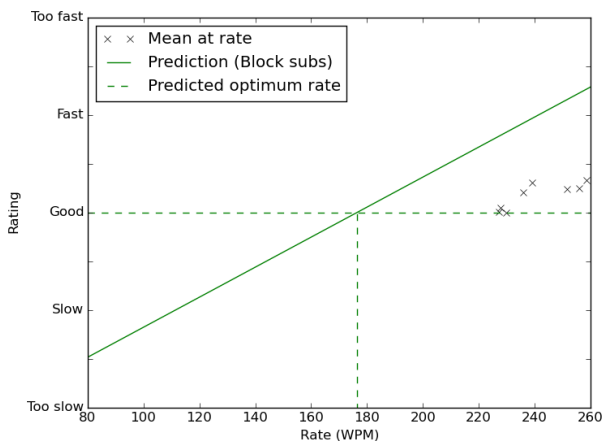
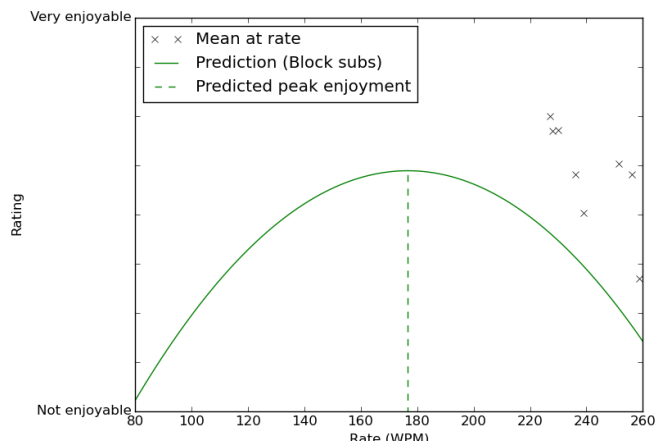Figure 3 – Perceived rate of off-air content against measured rate

Figure 4 – Perceived enjoyment of off-air subtitles against measured rate

Title	Style	Clip length	Rate (WPM)	Rate mean (1-5)	Rate IQR (1-5)	Enjoy mean (1-5)	Enjoy IQR (1-5)
rockford	Drama	3m06s	252	3.24	0.28	3.51	1.41
homes	Factual	0m56s	227	3.01	0.05	4.00	1.42
topgear	Talk show	0m56s	256	3.24	0.49	3.40	1.54
weather	News	0m59s	239	3.31	0.38	3.01	0.61
wogan	Talk show	2m04s	230	2.99	0.06	3.85	1.32
escape	Factual	2m11s	228	3.05	0.05	3.84	1.63
kitchen	Cookery	1m04s	259	3.33	0.63	2.34	1.17
perfection	Quiz	0m56s	236	3.21	0.23	3.40	0.78

Table 2 – Study part 2 clip info, and means and IQRs for perceived rate and enjoyment

These results show that the perceived ideal rate found in part 1 does not apply to all content. Different content felt right at different speeds. Any difference in perceived rate must therefore be related to other issues. Table 2 also shows the spread of the data for each clip in the form of the inter-quartile range (IQR). Some clips have far larger ranges than others. This may indicate that personal preference or personal resilience to other issues within subtitles has a large effect on perceived rate and enjoyment.

It should be noted that a technical error resulted in only 21 of the 25 participants viewing the weather clip.

DISCUSSION

Part 1 of this study found the optimum rate of subtitles to be 171WPM for scrolling subtitles and 177WPM for block subtitles. These rates are approximately the maximum currently allowed under some guidelines. The point at which most subtitle users would find the

subtitles too fast was found to be 227WPM for scrolling and 242WPM for block subtitles. These values are far above the 200WPM some guidelines warn would be difficult for many viewers to follow (2). Furthermore, part 1 of this study went on to find that ratings for subtitles at various rates were close to or better than those for hearing viewers rating the speed of speech on clips without subtitles. From this we infer that when there is an issue with the rate of subtitles, the same issue can be expected in speech and vice-versa. Issues of rate may then be expected to be noticed by content producers in the speech before the subtitles are created. Further to this, participants within this study repeatedly commented that "if there is a mismatch [between the subtitles and speech] then that's a problem". They were confused by our request for them to rate the speed of subtitles as this is dictated by the speed of speech and cannot be changed. Many see it as important that the subtitles are as close as possible to the speech in both timing and wording to make the understanding of speech/lip-reading and the subtitles combined as easy as possible. Part 1 of this study showed that not only is the rate of subtitles not an issue when the subtitles are verbatim and correctly timed, but that the way that the issue of rate is interpreted by previous academic literature and guidelines does not match the perceptions of users.

Part 2 of this study aimed to discover if the findings in part 1 held true across a range of real-world content. Clips as high as 230WPM in this section were tightly rated as perfect in speed. The three best rated clips contained discussions between multiple people, a situation where the overall rate of speech is naturally higher. Furthermore, some of these clips occasionally had the speaker out of shot or in wide shots. The consistently good ratings suggest that people's following of the content was not impaired greatly by the inability to lip read for short periods. That said, some participants did express a preference for the speaker's face to be in shot. This not only enables lip-reading but also the interpretation of emotions absent from the subtitles. This section of the study clearly demonstrated that there is no single optimum or maximum permissible rate for subtitles. These rates are highly dependent on the type of content and what feels natural.

Broadcasters receive complaints about the rate of subtitles and part 1 of this study shows people have a consistent perception of what content is too slow and too fast. If perceived rate is not caused by actual rate, as shown in part 2, then what is it caused by? In part 1 of this study, it was likely caused by a sense of what is a natural rate. The fast/slow speech in these clips felt oddly fast or slow. The fast speech of a malfunctioning robot in a movie is intentionally odd. But the fast speech of a frightened character feels right. Secondly, people may identify hard to follow content as too fast/slow. In the case of subtitles with delay or errors, it becomes harder to match audio/lip-reading with the subtitles necessitating higher concentration. This may be exacerbated by high information density at high rates. Conversely, a sentence that crosses the boundary between two slow subtitles will also require effort to hold the first part of the sentence long enough to combine it with the second and make sense of it as whole. It is also possible previous studies have failed to make the meaning of their question clear and clarify the answers of participants. This study identified multiple cases of participants using the term "too slow" to describe increased latency. Complaints of live subtitles being "too fast" may be explained by a combination of latency, errors and necessitated editing all requiring high concentration as well as subtitling systems causing subtitles to "bunch up" and be played out at inconsistent rates. It should be noted that the use of these terms was clarified with participants in this study.

Previous studies and guidelines have insisted relatively low rates are needed to enable viewers to follow the content - even if they request otherwise. This study has shown that low rates are not required for viewers to feel that they are following the content sufficiently.

CONCLUSION

This study has shown that the perceived rate of subtitles for frequent users tends to align with those of speech for the hearing. It has shown that different content feels right at different speeds. Furthermore, the perceived rate of subtitles is not representative of the actual speed but is a symptom of technical issues and the overall natural feel of the programme. To avoid perceived issues with rate, subtitles should match the speech in timing and wording. We found no problems associated with the rate of subtitles when they matched natural speech, regardless of the rate in words per minute.

REFERENCES

1. Ofcom. 2013. The quality of live subtitling - Improving the viewer experience.

2. Mikul, C. 2014. Caption quality: International approaches to standards and measurement.

3. Jensema, C. 1998. Viewer reaction to different television captioning speeds. American Annals of the Deaf, Vol. 143. pp. 318-324.

4. Burnham, D. et al. 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, Vol. 13. pp. 391-404.

5. Romero-Fresco, P. 2010. Standing on quicksand: Viewers' comprehension and reading patterns of respoken subtitles for the news. Approaches to Translation Studies, Vol. 32.

6. Romero-Fresco, P. 2012. Quality of live subtitling: The reception of respoken subtitles in the UK. Approaches to Translation Studies, Vol. 36.

7. Steinfeld, A. 1999. The benefit to the deaf of real-time captions in a mainstream classroom environment.

8. Jensema, C., McCann, R. and Ramsey, S. 1996. Closed-caption television presentation speed and vocabulary. American Annals of the Deaf, Vol. 141. pp. 284-292.

9. Noland, K. C. and Truong, L. H. 2015. A survey of UK television viewing conditions. BBC R&D - WHP 287.

10. Action on Hearing Loss. Accessed: 28 April 2015. Statistics. Action on Hearing Loss. http://www.actiononhearingloss.org.uk/your-hearing/about-deafness-and-hearing-loss/statistics.aspx.