Automated Monitoring of UK SDH Quality - a research tool exploring methods and results

Michael Armstrong

Associate staff - University of Dundee
& member of the UK Subtitling Audiences Network

Languages and the Media
5/11/2026

The Audience Strikes Back

Michael Armstrong

TV Viewer and Retired R&D Engineer.

Languages and the Media
5/11/2026

Definitions

This presentation is about broadcast TV subtitles in the UK.

This work measures factors which *degrade* perceived subtitle quality, detecting faults and failures so they can be addressed.

It does **not** attempt to produce an single quality score.

Technical details are not included

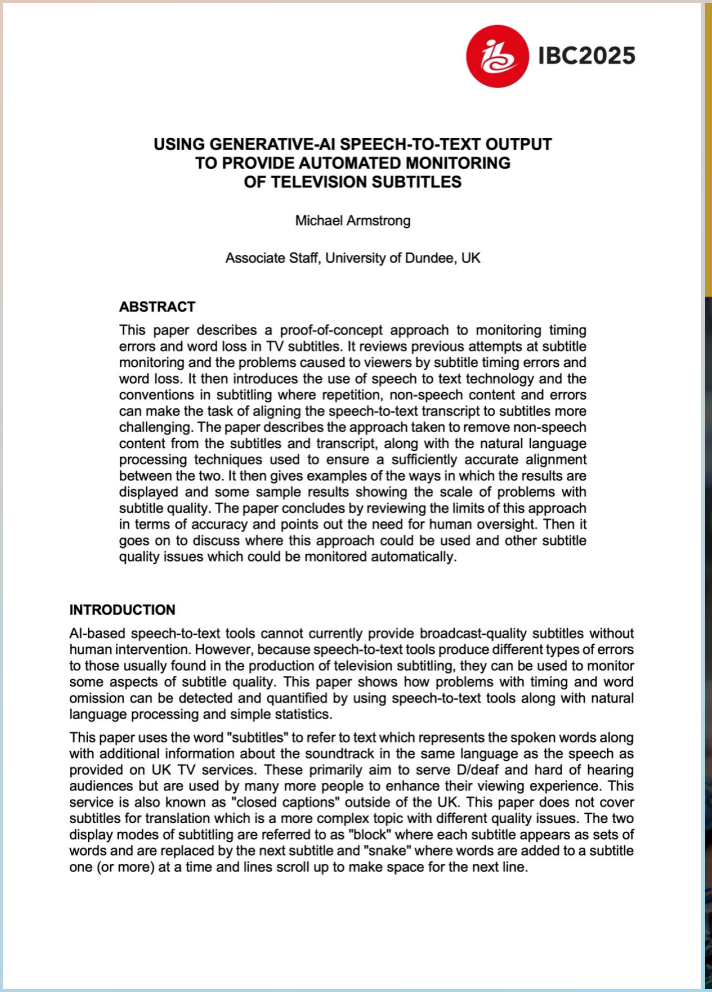
This presentation omits the technical details of this work.

Technical papers are available from

IBC2025 Conference

and

SMPTE Motion Imaging Journal



IBC2025

USING GENERATIVE-AI SPEECH-TO-TEXT OUTPUT TO PROVIDE AUTOMATED MONITORING OF TELEVISION SUBTITLES

Michael Armstrong

Associate Staff, University of Dundee, UK

ABSTRACT

This paper describes a proof-of-concept approach to monitoring timing errors and word loss in TV subtitles. It reviews previous attempts at subtitle monitoring and the problems caused to viewers by subtitle timing errors and word loss. It then introduces the use of speech to text technology and the conventions in subtitling where repetition, non-speech content and errors can make the task of aligning the speech-to-text transcript to subtitles more challenging. The paper describes the approach taken to remove non-speech content from the subtitles and transcript, along with the natural language processing techniques used to ensure a sufficiently accurate alignment between the two. It then gives examples of the ways in which the results are displayed and some sample results showing the scale of problems with subtitle quality. The paper concludes by reviewing the limits of this approach in terms of accuracy and points out the need for human oversight. Then it goes on to discuss where this approach could be used and other subtitle quality issues which could be monitored automatically.

INTRODUCTION

AI-based speech-to-text tools cannot currently provide broadcast-quality subtitles without human intervention. However, because speech-to-text tools produce different types of errors to those usually found in the production of television subtitling, they can be used to monitor some aspects of subtitle quality. This paper shows how problems with timing and word omission can be detected and quantified by using speech-to-text tools along with natural language processing and simple statistics.

This paper uses the word "subtitles" to refer to text which represents the spoken words along with additional information about the soundtrack in the same language as the speech as provided on UK TV services. These primarily aim to serve D/deaf and hard of hearing audiences but are used by many more people to enhance their viewing experience. This service is also known as "closed captions" outside of the UK. This paper does not cover subtitles for translation which is a more complex topic with different quality issues. The two display modes of subtitling are referred to as "block" where each subtitle appears as sets of words and are replaced by the next subtitle and "snake" where words are added to a subtitle one (or more) at a time and lines scroll up to make space for the next line.



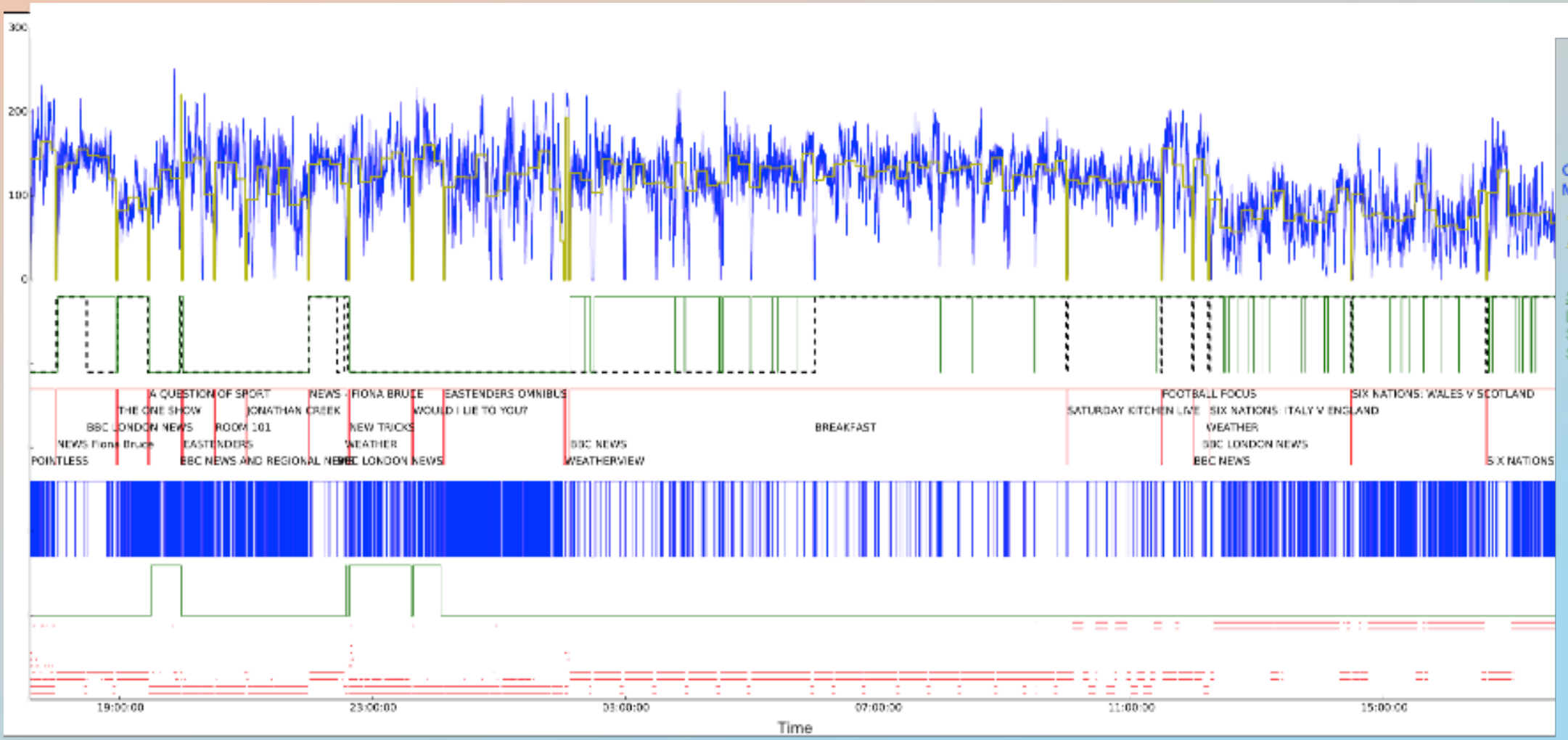
Motivation (this work is unpaid!)

Between 2012 and 2016 I led a small team at BBC R&D which had a positive impact on live subtitle quality.

After retiring in 2024 I became painfully aware that some of the improvements we had made to subtitles were now making matters worse with some delayed by over 30s.

To capture these intermittent problems, I needed to make measurements over long periods of time.

Monitoring Teletext Subtitles- BBC R&D 2014



Adding speech to text

Modern speech to text engines are not 100% accurate and will interpret some speech content differently from a human subtitler.

However, they can give accurate word timings and a reasonably accurate word count.

Therefore, they can be used to measure timing and word loss.

These are the two faults that the audience find most annoying.

Subtitle data

Subtitle data describes how text should be displayed on a screen.

It contains both repetition and non-speech elements.

These need to be removed to leave just the speech content in order to make a reasonable comparison with the transcript.

Alignment

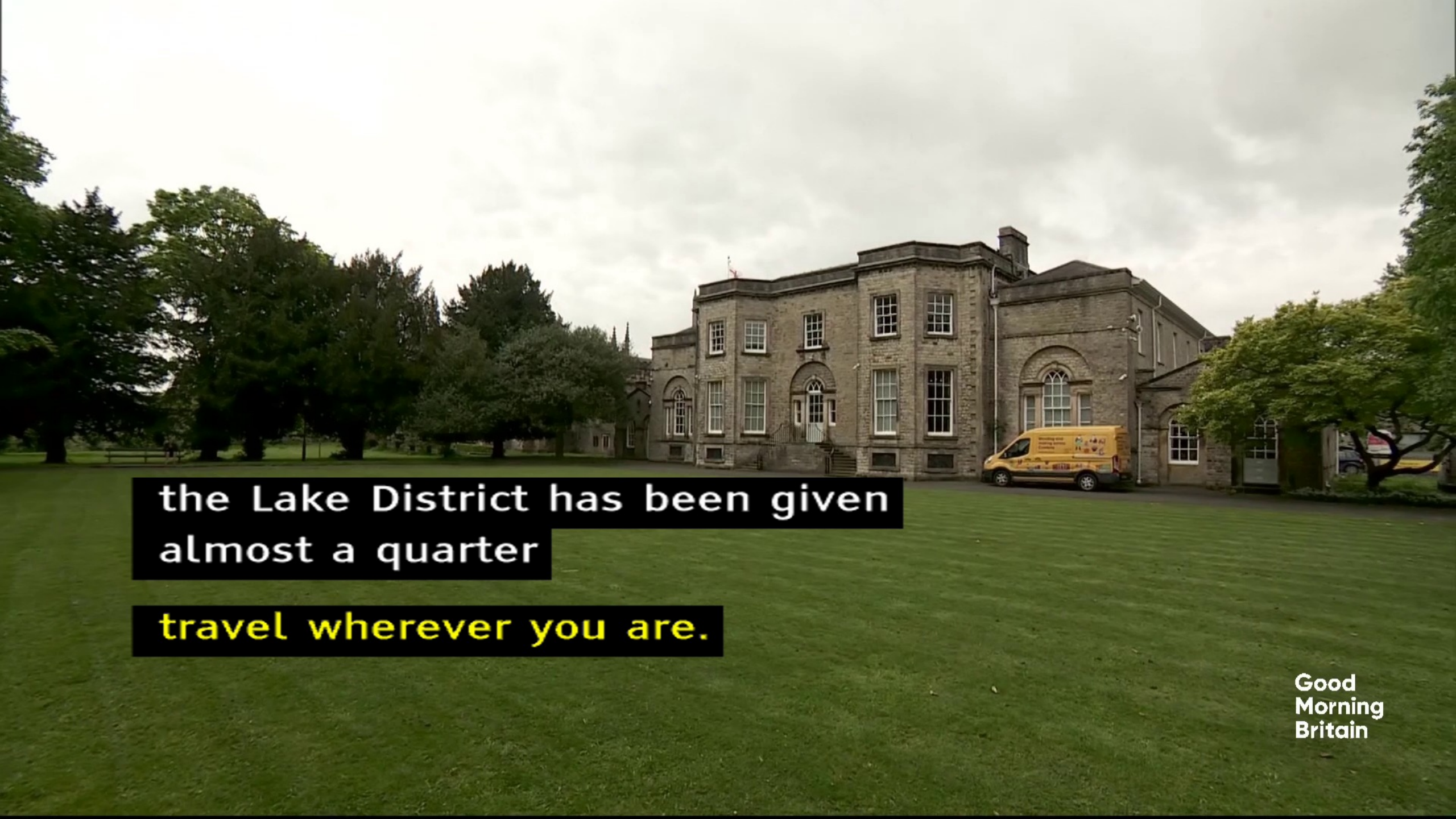
To measure timing, the transcript output by the speech-to-text engine has to be aligned with each of the television subtitles.

- This is straightforward with high quality, pre-prepared subtitles and a clear speech soundtrack.

- However, as the subtitle and audio quality decline, the difficulty of aligning the transcript to the subtitles increases.

Things that make alignment difficult - 1

- The timing in the subtitles and transcript may not match.
- The subtitles may omit many of the spoken words.
- Word errors in both the subtitles and the transcript.
- Spelling differences between the subtitles and the transcript.
- Compound words and contractions vs separate words.
- Numerals vs number words.
- The words in the subtitles can be in the wrong order.
- Long sections of subtitles can be repeated.
- Subtitles can get stuck.


the Lake District has been given
almost a quarter
travel wherever you are.
Good
Morning
Britain

Things that make alignment difficult - 2

- Live subtitles can contain corrections – in varying formats.
- Non-speech utterances, especially in children's TV.
- Speech or singing containing lots of repetition.

Also, the system needs to cope with…

- Programmes with no subtitles.
- Programmes that contain subtitles but no speech.
- Programmes broadcast with the wrong subtitles.

Accuracy

The accuracy of the system was confirmed by measuring known high-quality subtitles, i.e. pre-recorded, factual programmes with a clear speech content.

- The word count differences are less than ±1%.
- The timing measurements are within ±1 second.
- The software manages to align over 97% of the words.

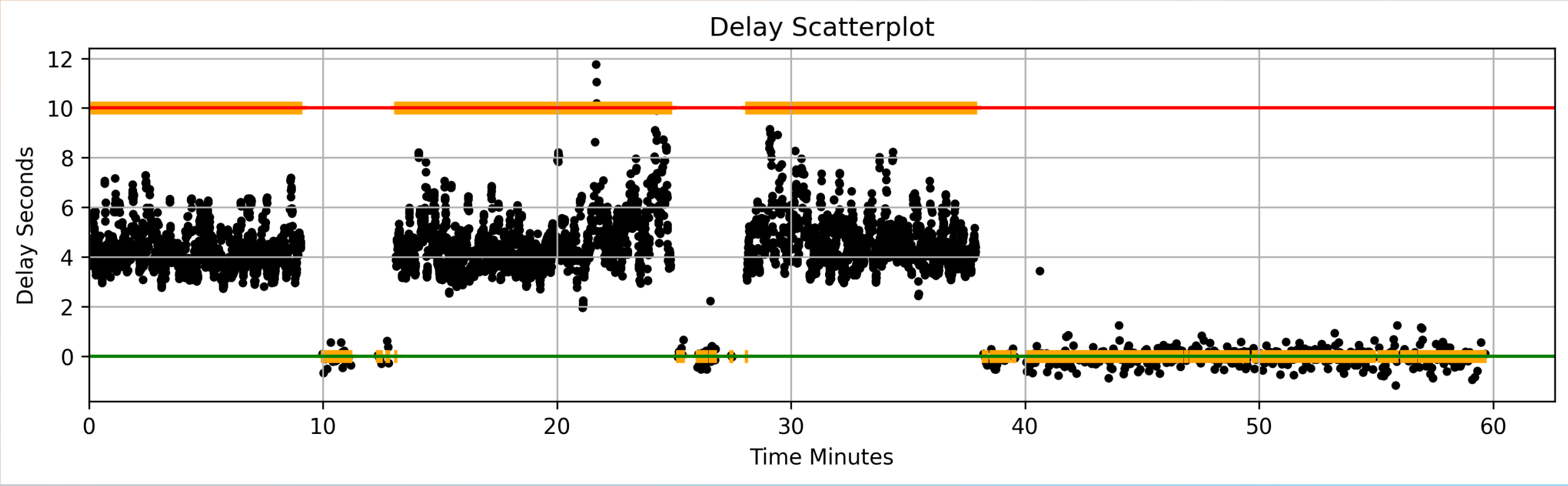
Displaying the results

The results are plotted against a time-line in a series of graphs.

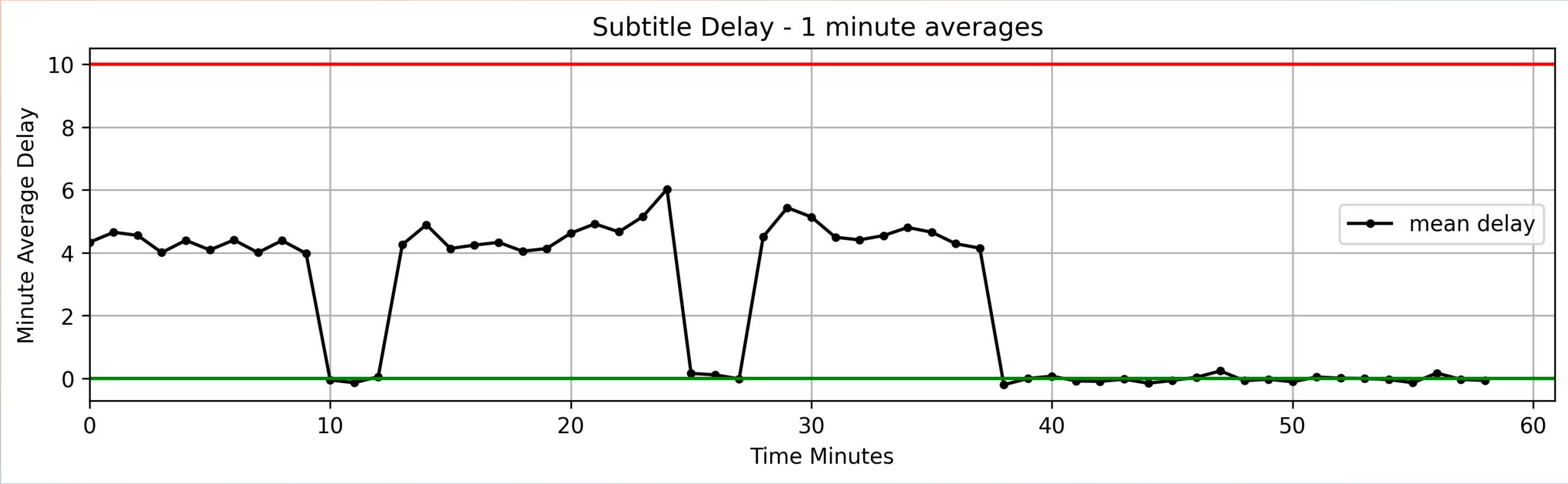
Subtitles over 10 seconds late are effectively useless, indicated by a red line on the plot with a green line for zero delay.

Orange markers on the red line indicate snake/word-at-a-time subtitles and orange markers on the green line indicate block subtitles.

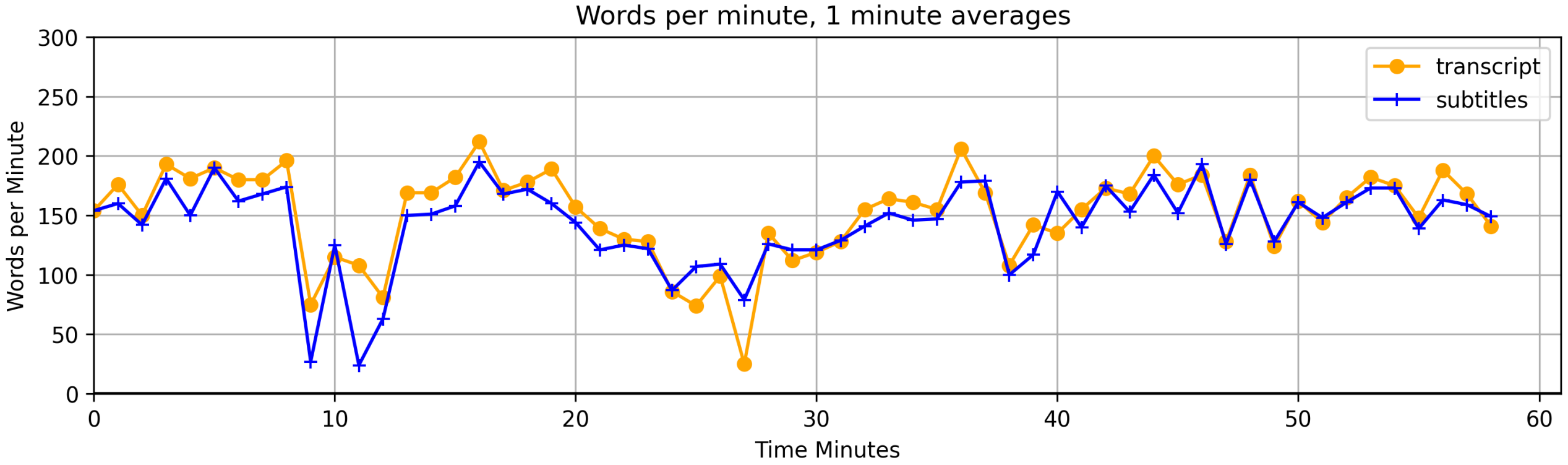
Timing - scatter plot



Timing - one minute-average



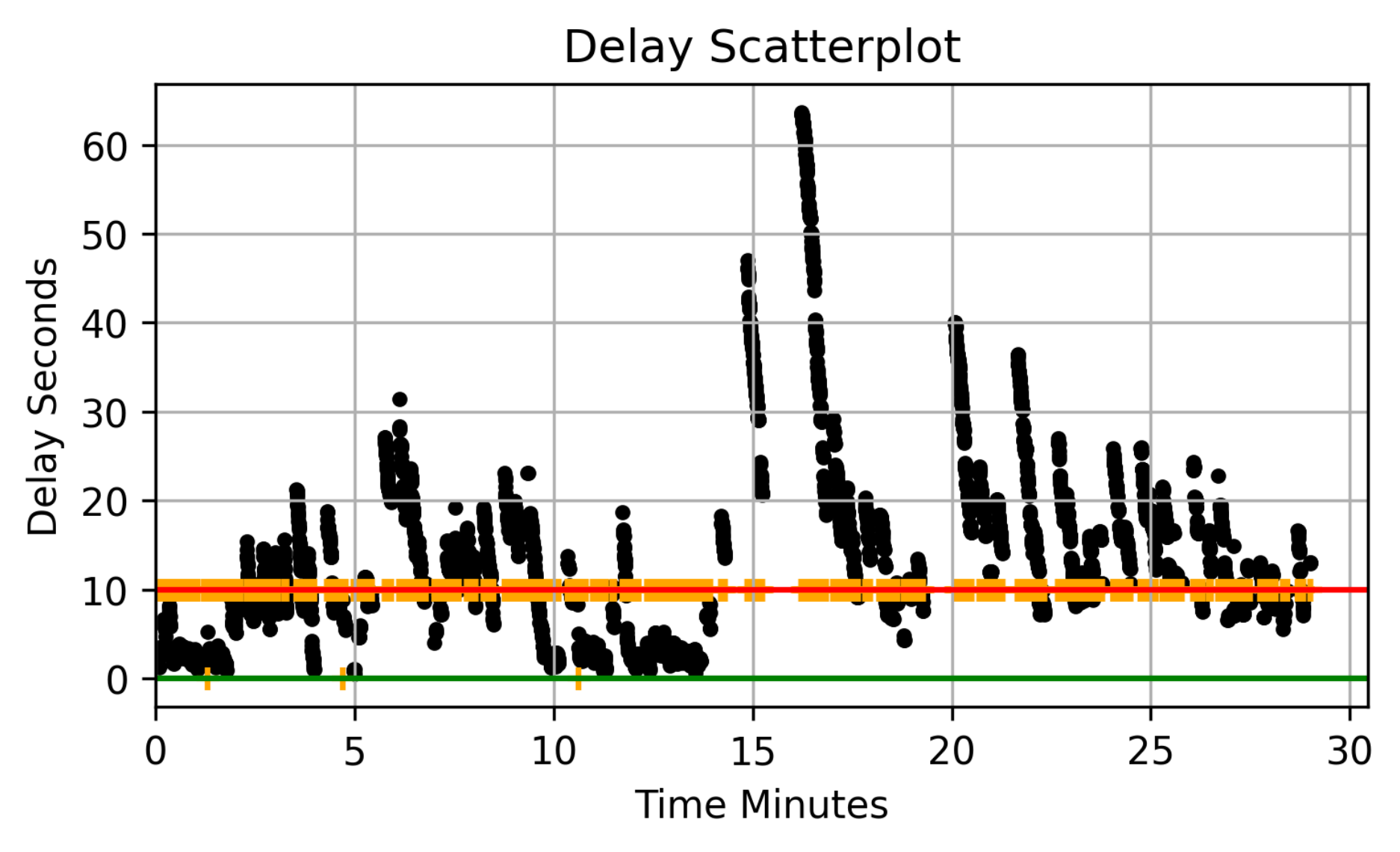
Words in each minute



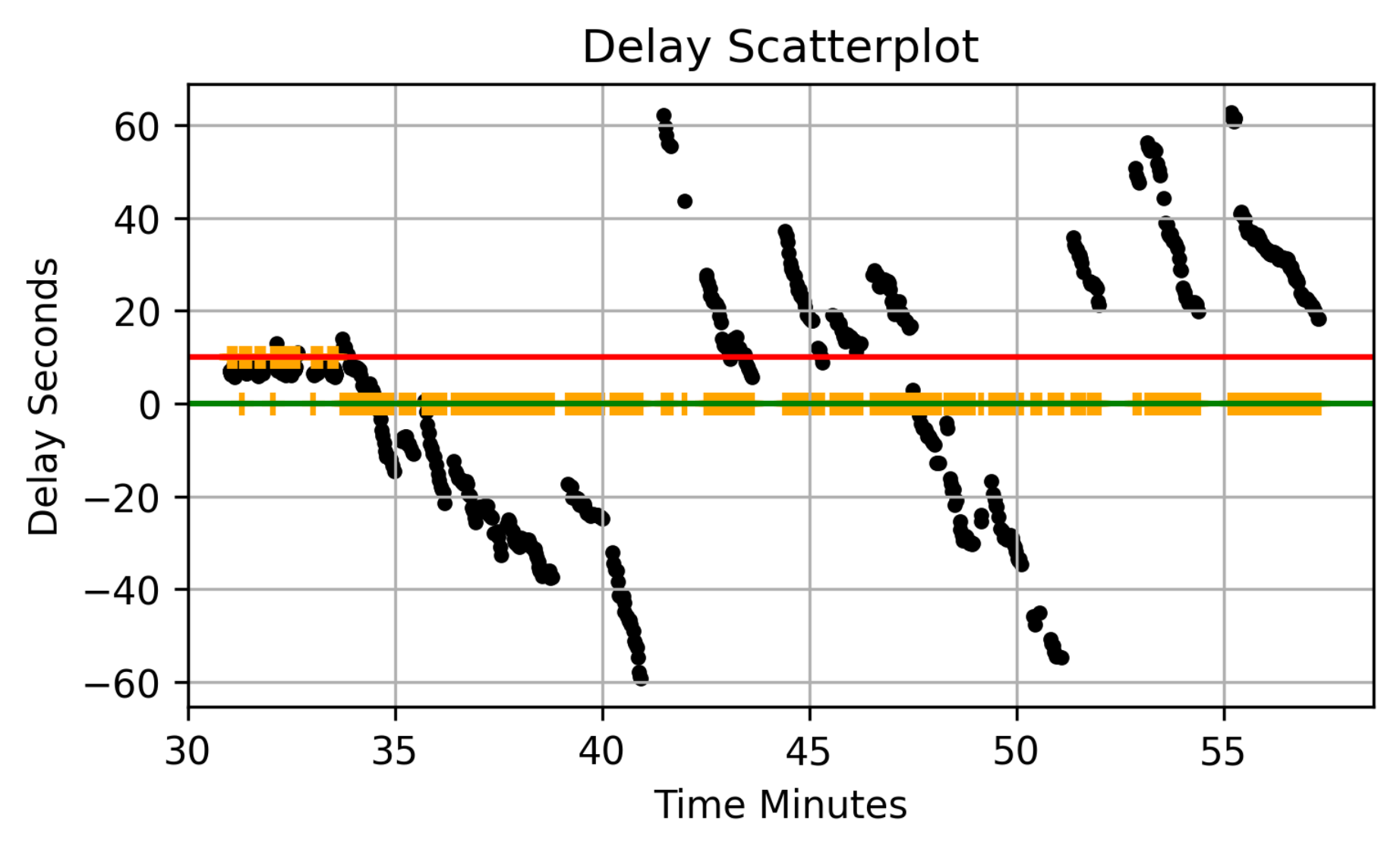
Major quality failings

- Some channels regularly broadcast subtitles over 10 seconds late, sometimes many times that.
- There also are examples of subtitles being broadcast early.
- And examples of subtitles omitting over half the spoken words.
- The following examples are from early 2025.

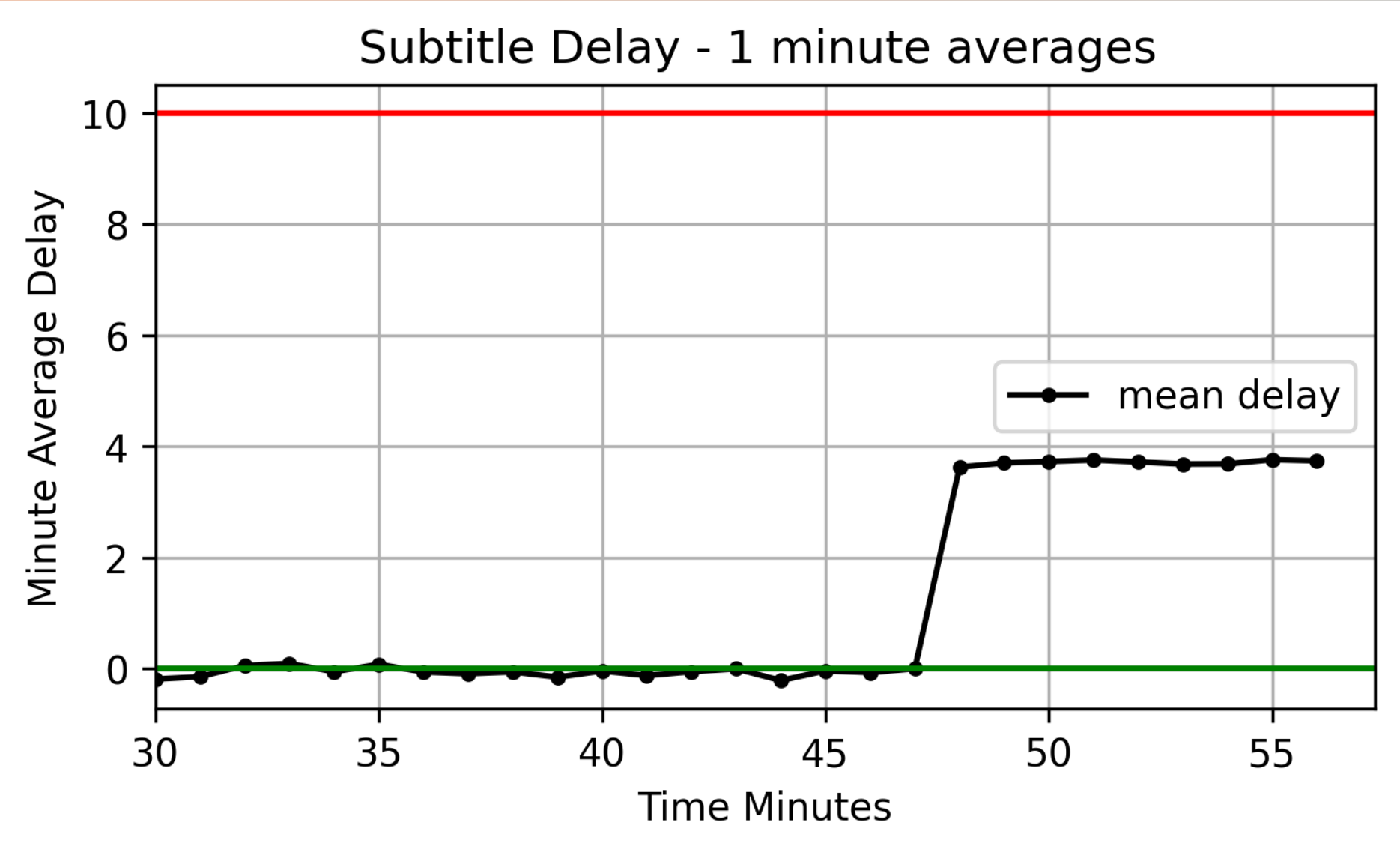
Problems with timing – live (snake)



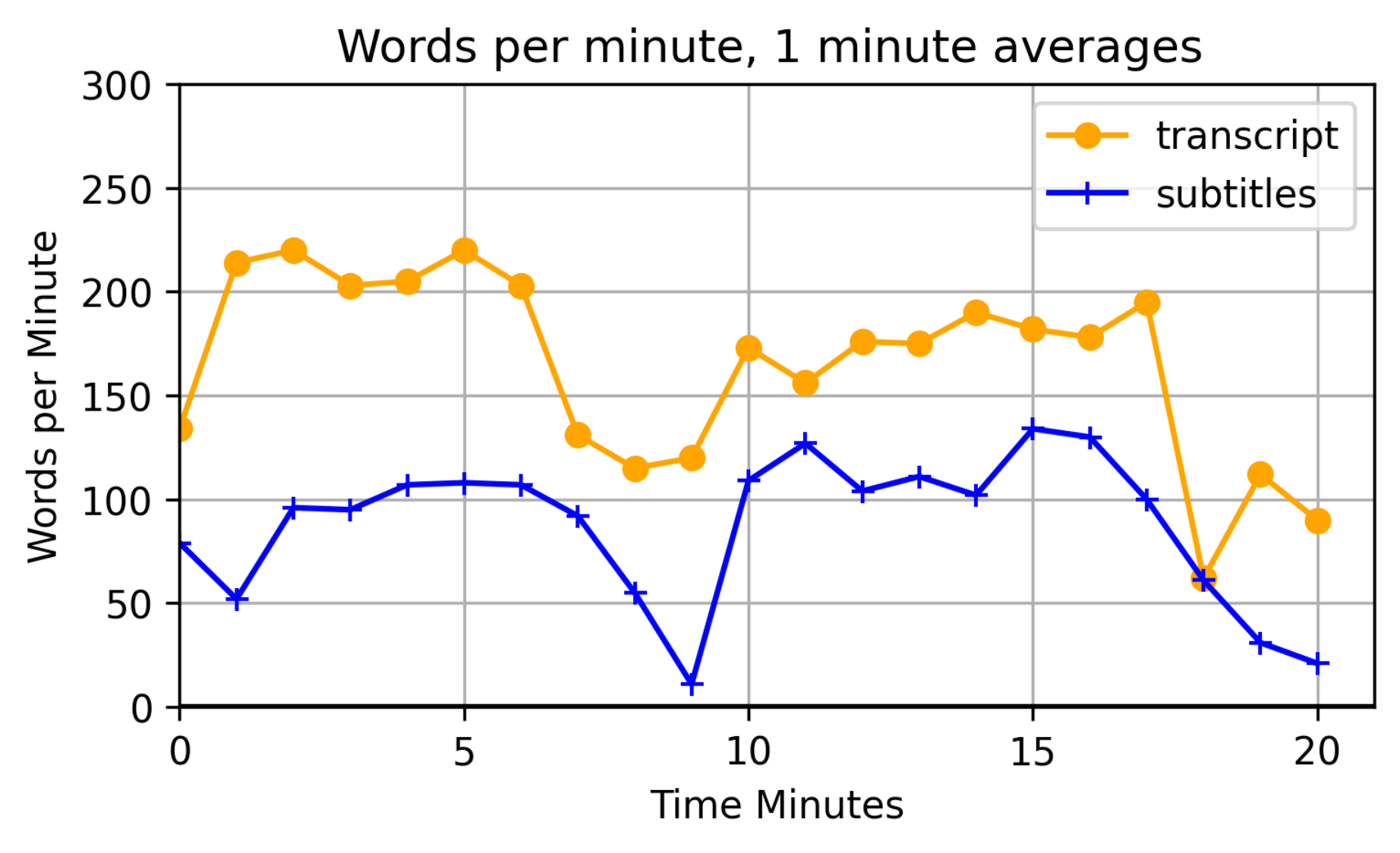
Problems with timing – live (block)



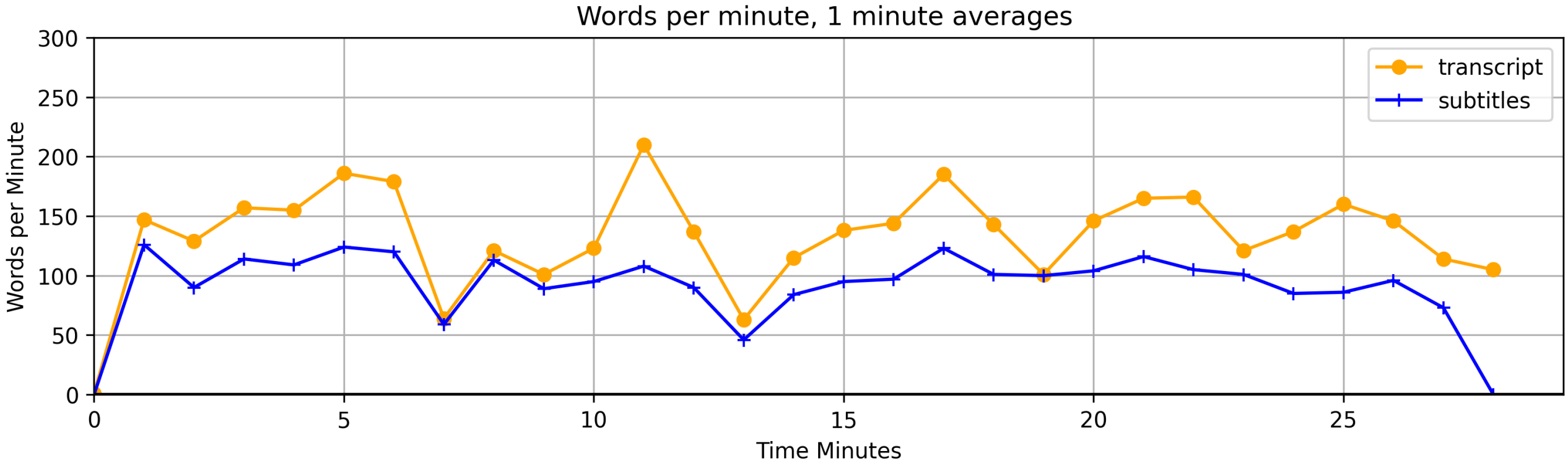
Problems with timing – pre-recorded



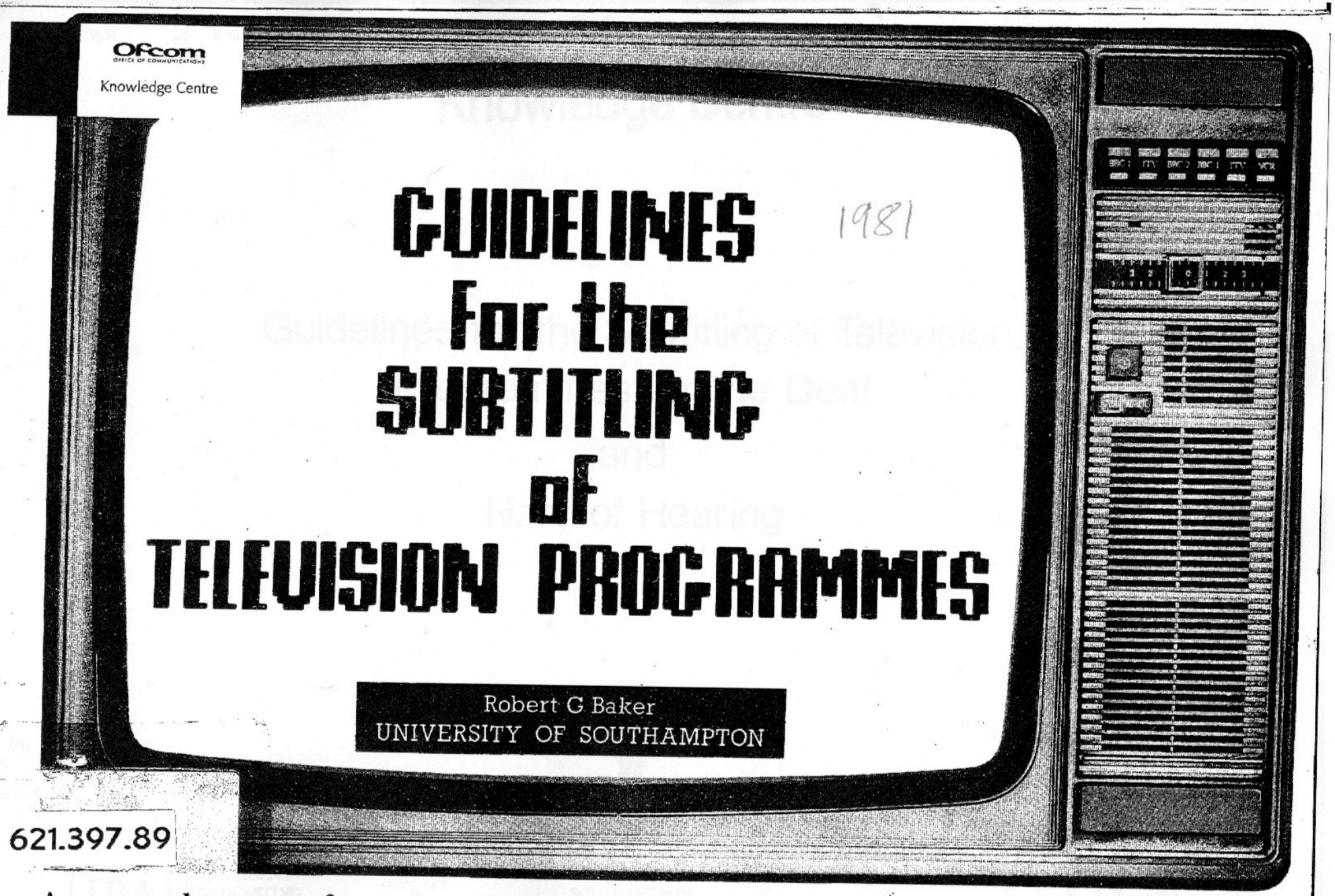
Problems with word loss - live



Problems with word loss – archive (1988)



Ofcom
OFFICE OF COMMUNICATIONS
Knowledge Centre



GUIDELINES For the SUBTITLING of TELEVISION PROGRAMMES

1981

Robert G Baker
UNIVERSITY OF SOUTHAMPTON

621.397.89

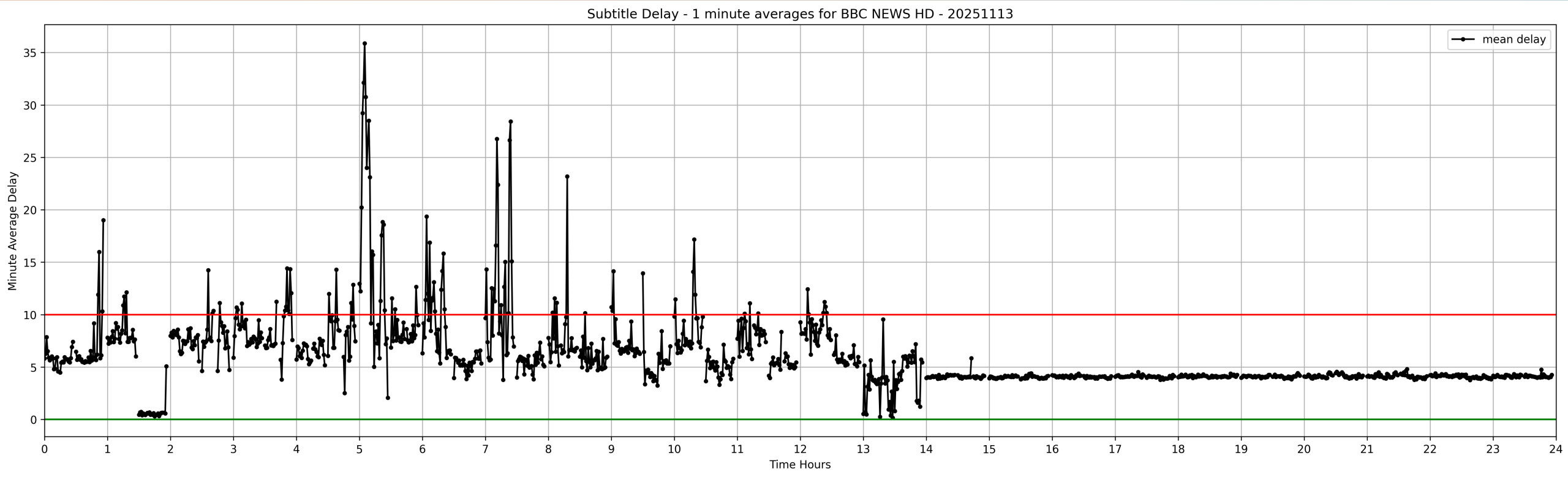
A research report from Southampton University for Independent Television (UK)

24 hour plots

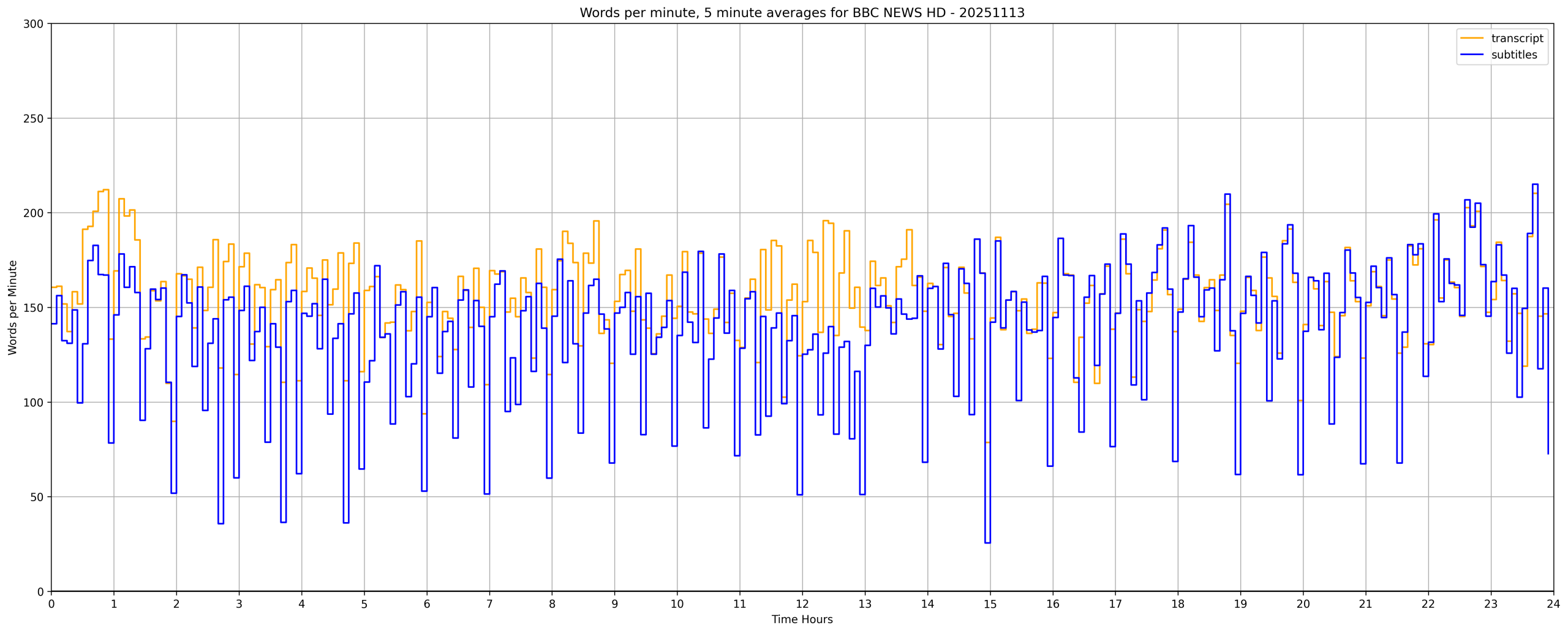
24hour plots were added to reduce the manual effort involved.

EPG data was then added to identify the individual programmes.

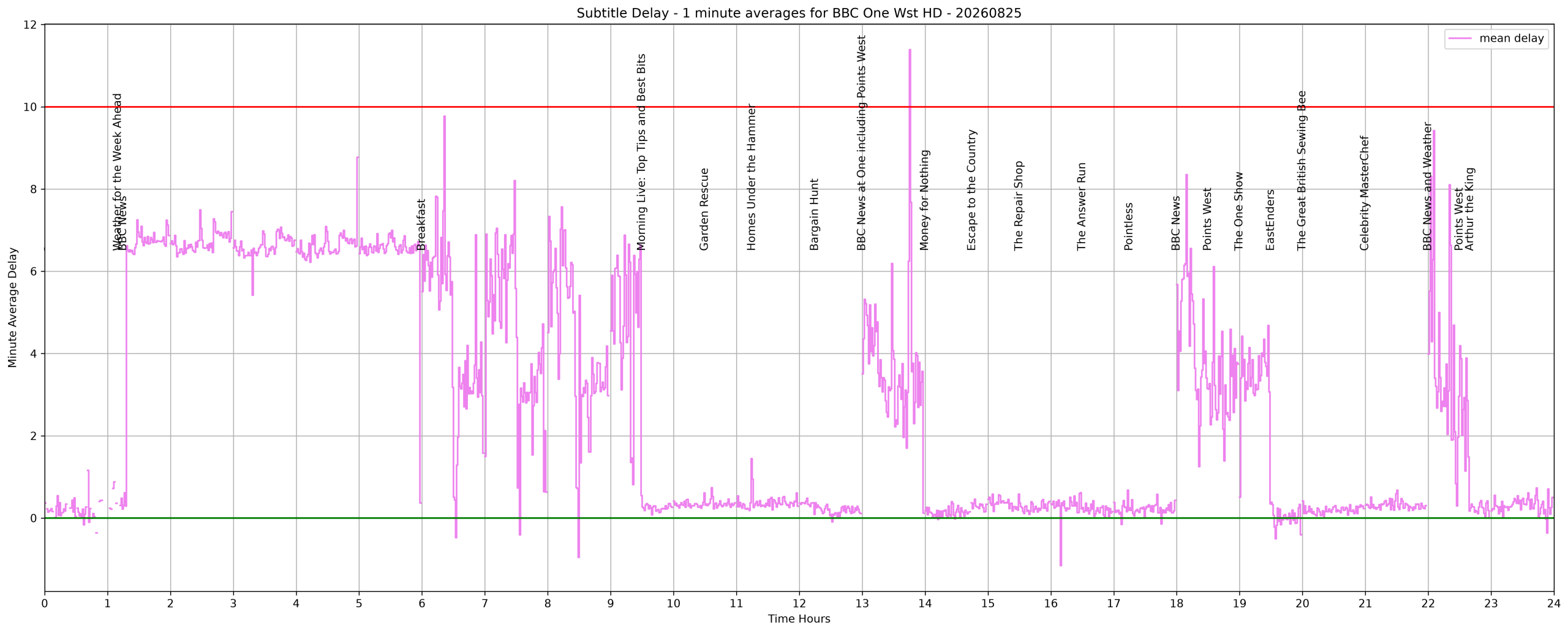
BBC News delay - 13/11/2025 (ASR trial)



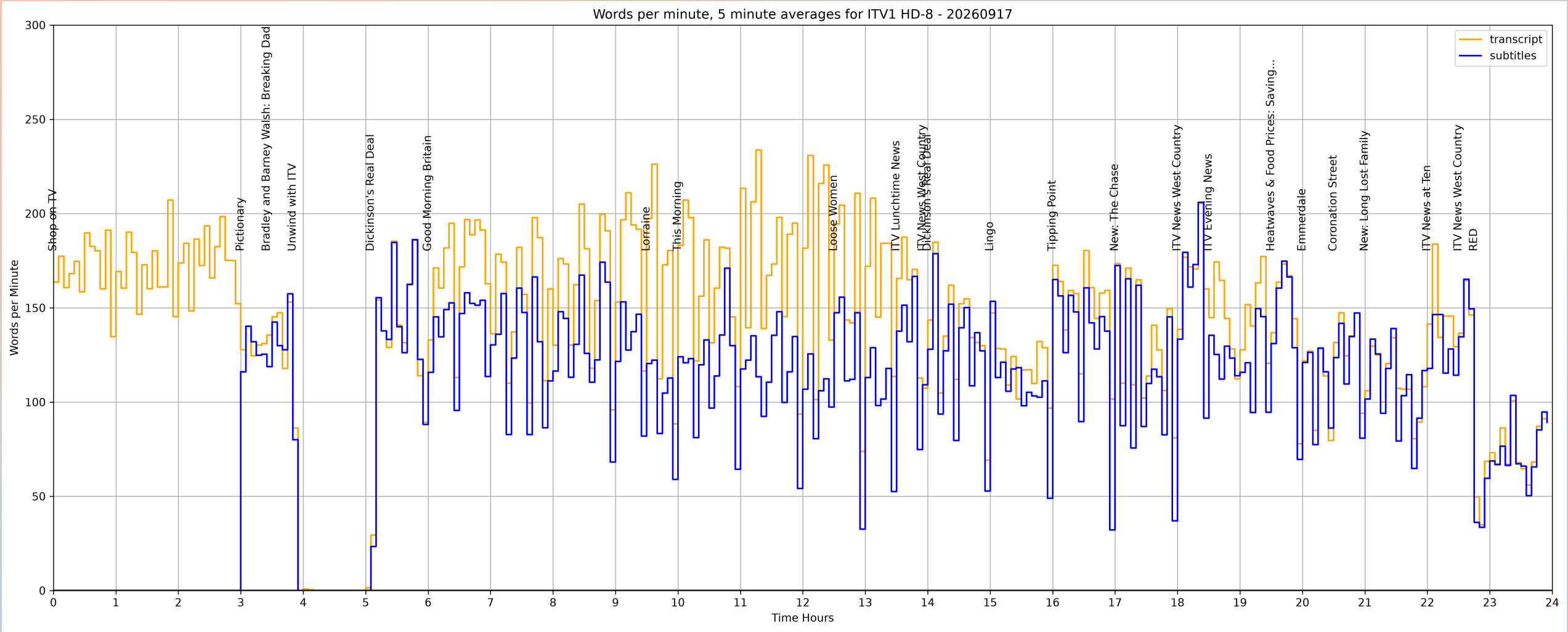
BBC News word rate - 13/11/2025 (ASR)



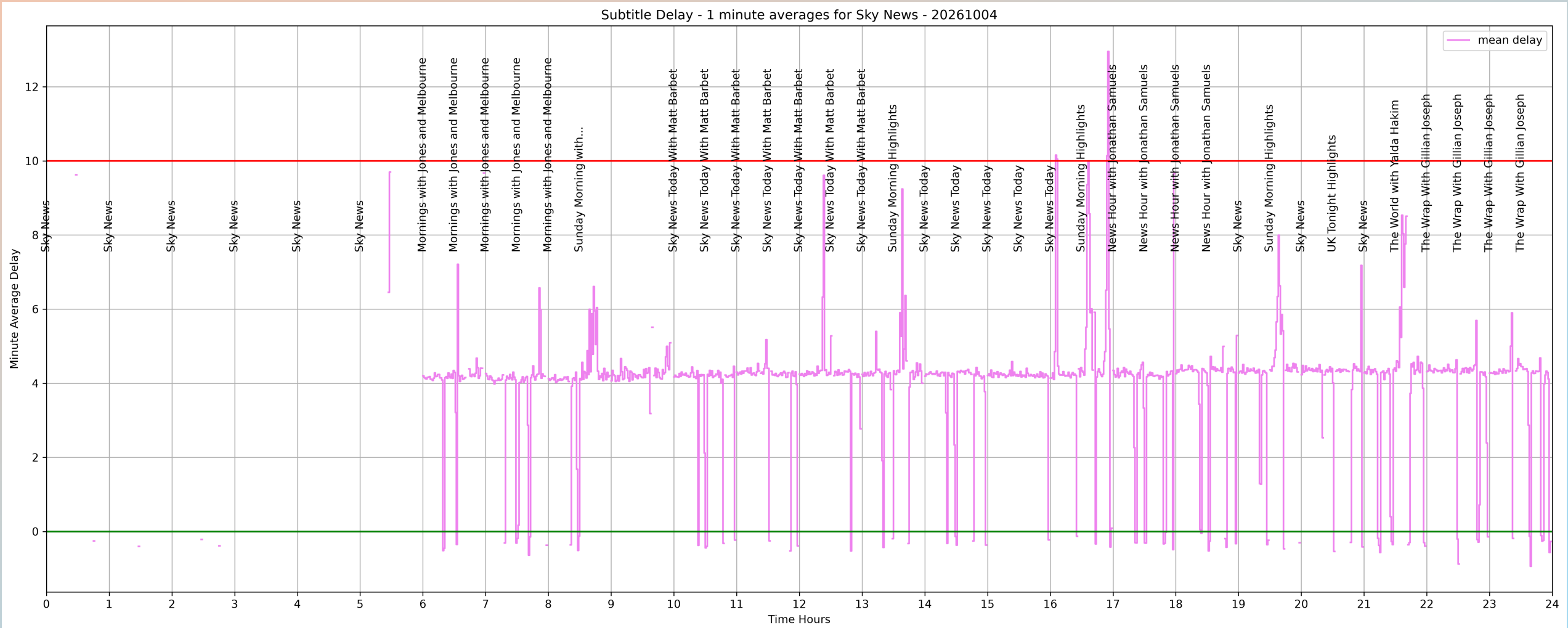
BBC One delay with EPG - 25/08/2026



ITV1 with EPG – wpm 5 min – 17/09/2026



Sky News with EPG – delay 04/10/2026



Conclusions

- This work demonstrates the viability and value of automated subtitle monitoring for delay and word loss.
- Feedback to broadcasters has assisted in some significant improvements in subtitle timing.
- There has also been an increase in the deployment of automatic speech recognition for news channels.
- However, significant problems remain with word loss on daytime magazine programmes, and this represents a major research challenge.

Acknowledgements

I would like to thank

Dr Michael Crabb,

Head of Science at the University of Dundee and

Dr Carol O'Sullivan,

Associate Professor in Translation Studies at the University of Bristol and co-founder of the UK Subtitling Audiences Network

for their support and encouragement.

For further information visit
www.subtitles.org.uk

email:

info@subtitles.org.uk

