ACADEMIC PROJECT

# Evaluating the UX of Captioning Methods within a Cinema Environment


## Jenny Cheng
## MSc UX Engineering

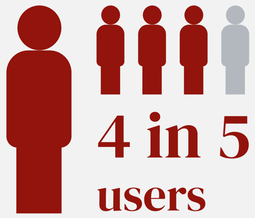
Goldsmiths
UNIVERSITY OF LONDON



## Summary

Closed captioning devices (CCDs) are vital for D/deaf and hard of hearing individuals, offering synchronised text captions to access content. Studies have found that around 60%-80% of users under the age of 35 also prefer subtitles. This presents an opportunity for CCDs in cinemas to enhance film comprehension, even for general audiences and foreign films without affecting others. Instead of using CCDs, UK cinemas currently depend on open subtitling or on-screen captioning. There are drawbacks to this strategy since it requires all users to employ captions; also, screenings are often restricted in quantity and available at unsociable hours of the day. To contrast, the U.S. has federal law requiring such CCDs to be provided at venues. Despite widespread use in North America, CCD usability in cinemas remains under-explored.

This study employs eye-gaze tracking glasses to investigate cognitive load effects of two common cinema CCDs: closed captioning glasses (CC Glasses) and closed captioning stands (CC Stands). Eye movement including fixations (when eyes stop moving), saccades (twitches & jerks), and saliency maps were used to evaluate cognitive load. Findings show CC Glasses mimic familiar gaze patterns, though ergonomic considerations are important. CC Stands aid comprehension but their smaller display prolongs fixations, increasing cognitive load. Analysis reveals captioning characteristics also impact cognitive load.


between 18-25 in the UK use subtitles some or all of the time [1]

## Methodology

Participants (18) were invited to the campus cinema and instructed to watched a clip from Japanese film *Tampopo* (1985) either using the simulated CC Glasses or CC Stand method, their eye movements were recorded throughout the task using the Tobii Pro Glasses 3. Post-task, a questionnaire gauged comprehension and CCD usability. Eye-tracking generated saliency map (heat-map) visualisations, and area of interest (AOI) metrics which formed the bulk of the data for analysis. Two AOIs were mapped for all footage: Captions AOI, and Content AOI (Fig 6).

## Simulated CCD Builds

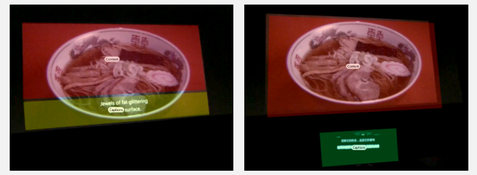
Due to challenges in obtaining professional cinema CCDs, both captioning methods were simulated (Fig 1,2,5). On-screen subtitles with eye-gaze tracking glasses simulated CC Glasses. While the CC Stand was built using a gooseneck phone stand, mounting a smartphone as the subtitle display.

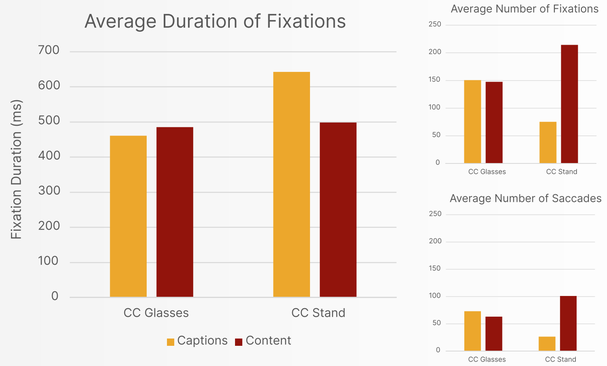
## Results

All participants found that using captions aided with comprehension regardless of captioning method. Overall, both CCD methods perform similarly, with no conclusive difference in inferred cognitive load measures. Though some users indicated discomfort using a second screen, and eye-tracking glasses. Saliency map visualisations demonstrated the convergence of fixation points around the Caption AOI (Fig 3,4). Through AOI analysis, CC Stand users were found to be fixating less on the captioning display, but focusing on the captions for a longer duration.

Fig 5. Simulated CC Stand apparatus with adjustable arm mounted to desk

Fig 6. AOI analysis required manual frame-by-frame mapping to account for head movements






### Closed Captioning Glasses (CC Glasses)

CC Glasses, like the *Sony Entertainment Access Glasses*, integrate real-time captions on the lenses, synced wirelessly with cinema captioning and audio descriptions. Using augmented reality (AR), they aim to display text without requiring a separate screen.

### Closed Captioning Stands (CC Stand)

CC Stands like *CaptiView* have separate displays attached to an adjustable arm, showing real-time captions synchronised with film audio. Wireless connectivity permits users to adjust screen position, brightness, and font size for personalised visibility and comfort.

Fig 1. Simulated user view point using CC glasses


Fig 3. CC Glasses cumulative saliency map

Fig 2. Simulated user view point using CC


Fig 4. CC Stand cumulative saliency map

## Key Findings + Recommendations

AOI proximity and size play an important role when considering cognitive load. CC Stand's smaller display reduces fixations in the Caption AOI, causing fewer but longer fixations compared to CC Glasses. This aligns with literature on second screen impacts on cognitive load. Additionally, familiarity of stimuli has been found to be associated with shorter fixations [2][3][4]

The study emphasises the demand for adaptable cinema subtitling solutions, requiring human-centered design innovations. Considerations include ergonomics of wearable AR glasses, dynamic subtitles adapting to gaze patterns, and predictive AI systems. Yet, cost and policy enforcement for Cinema CCD access remain vital.


[1] Stagetext, Sapio Research (2021) 'Collecting opinions about subtitles for Stagetext', Sapio Research. Available at: https://sapioresearch.com/success-stories/stagetext/ (Accessed: 7 August 2023).

[2] Chan, J.L. and DeSouza, J.F.X. (2013) 'The effects of attentional load on saccadic task switching', Experimental Brain Research, 227(3), pp. 301–309. Available at: https://doi.org/10.1007/s00221-013-3452-1.

[3] Lee, S., Baek, J. and Han, G. (2018) 'Effects of using a second-screen application on attention, learning, and user experience in an educational content', Interactive Learning Environments, 26(3), pp. 289–307. Available at: https://doi.org/10.1080/10494820.2017.1324494.

[4] Van Cauwenberge, A., Schaap, G. and Van Roy, R. (2014) "'TV no longer commands our full attention": Effects of second-screen viewing and task relevance on cognitive load and learning from news', Computers in Human Behavior, 38, pp. 100–109. Available at: https://doi.org/10.1016/j.chb.2014.05.021.