Canadian
Association of the Deaf
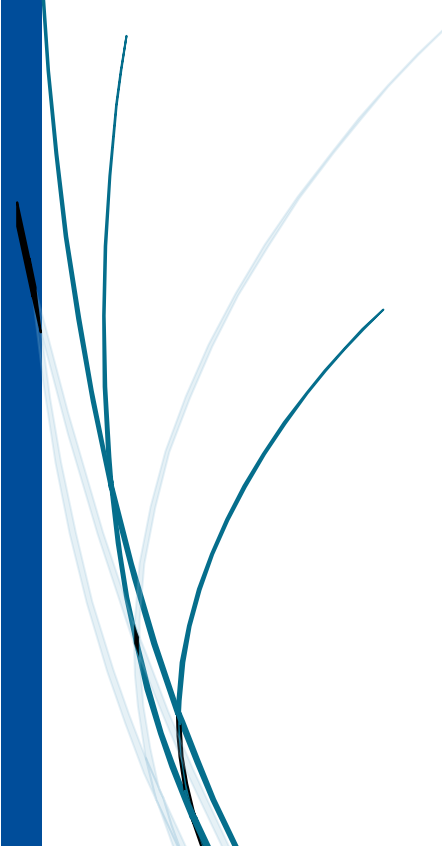
Association
des Sourds du Canada

Key Findings

User Experience of Play-by-Play Captioning in Fast-Paced Sports:

A comparison of two styles of live captioning for
Deaf and Hard of Hearing viewers

Acknowledgements

Richard Belzile, Incoming Project Leader and Chair of the Steering Committee
Canadian Association of the Deaf

Contents

Executive Summary

Live broadcasts for fast-paced sports are an important component to consider when evaluating the user experience of captions because the captions for these types of broadcasts tend to contain many errors, and are the most difficult content to caption due to the fast pace. For this study, we explored the user experience for fast-paced, live sports broadcasts, with a focus on hockey and basketball. Because these broadcasts typically include two off-screen commentators who are commenting on the play-by-play (PBP) action taking place during the game, as well as contributing commentary between the gameplay, there tends to be an extremely high word-per-minute rate associated with captions that are verbatim. In some cases, like basketball, there are easily over 200 words per minute (WPM) being presented as captions, which is effectively too fast to support a comfortable level of reading, not to mention leaving enough time to also enjoy the live-action gameplay. In this project, we explored several aspects of the user experience associated with watching fast-paced live sporting events, including the readability of the captions, the user's ratings for satisfaction and quality of captions, the level of comprehension that is possible either from captions or gameplay areas between the conventional captions style and a novel style where only the captions for the commentary. The results of an eye tracking study designed to monitor eye movement and focus that the users exhibit while watching live captioned sporting events are presented.

Results from this study support previous findings, suggesting that caption users are not fully satisfied with the quality of captioning available today. Fast-paced live-action sporting events in particular introduce additional factors, in that the amount of WPM required for verbatim captioning is often too much to read, and the delays, speaker identification, caption placement, and other factors may further cause viewers to become confused or distracted from what really matters, the gameplay itself.

We also explored a novel strategy to reduce the WPM of captions in sports broadcasting, by displaying only the commentary captions. Results suggested that there were no differences found in viewer preferences between the two styles of captioning. Given the viewers were already familiar with the Play-by-Play style (PBP), our findings suggested that the commentary only style (CO) gave positive impacts on perceived quality satisfaction rating, viewing experience, and supported understanding of the game/flow. A longitudinal study that would allow viewers to become acclimatized to CO captions was identified as necessary to understand the long-term impact and experience of these types of captions.

Further, while the results suggested that the participant's comprehension of the game was more dependent on their experience with the sport over having access to the captions, other valued elements of PBP captions included player names and perhaps a more paraphrased version of the gameplay captions. While other issues were raised by the participants, including placement, text size, synchronization delay, and spelling mistakes, the key takeaway from this research remains, having too many captions does not lead to better comprehension. Too many captions being displayed too fast added difficulties in following the game, tended to be distracting, and disrupted viewers' enjoyment of the game.

The rest of this report will present details and results of our study, including eye-tracking results, comprehension scores, and user feedback relating to the two styles of captions explored in this research.

Overview and Background

Live sports broadcasts for fast-paced sports inherently present many challenges for captioners and viewers, who face an extremely high WPM rate when aiming for/expecting verbatim captions in PBP caption style. For

captioners, keeping up with an announcer's rate of speech is often challenging due to the high-speed rate of dialogue, often provided by multiple commentators, which can result in long delays, numerous spelling mistakes, and other errors that are well-known to have an impact on the quality of the experience for the Deaf[1] and Hard of Hearing (D and HoH) viewers who rely on captions to access the speech content. Even without errors or delays, there can be over 200 WPM occurring throughout the broadcast, which is not only difficult for a viewer to keep up with but also can be a distraction that takes the viewers' attention away from the game and keeps them focused on the captions.

In this research, we explore the user experience of live-action, fast-paced captions that accompany live hockey and basketball broadcasts. To examine the user experience, we employed three methods:

1. Use of an eye-tracking system to monitor the viewer's eye movements while watching video segments of the games.
2. Collect feedback about users' experiences of the captioned segments that they watched using a series of questionnaires.
3. Conducting a conversation-style interview with participants in order to assess comprehension and how well the different styles of captions were able to provide details about the games to the viewers.

Each of the video segments that viewers watched was captioned in two different ways, namely, as broadcast, with PBP captions and alternatively, with CO captions. The CO caption style was created by a professional captioner with instructions to avoid captioning for any gameplay components of captioning, leaving commentary only to be captioned. The process simulated the live captioning for typical sports broadcasting, thus the decision was made by the captioner simultaneously while generating the captions.

The study provided the participant with a choice of watching either basketball or hockey, with each sporting event divided into two segments, so that we could present each participant with an example of both forms of captions for our study.

Project Objectives

The objective of this project is to provide insights into some of the factors that influence the user experience, enjoyment and comprehension of fast-paced live sports broadcasts. A further objective is to explore potential strategies for reducing the number of words presented in the captions, without impacting the viewer's ability to comprehend the game. We describe the details of the study in this report.

---

[1]The word Deaf (with capital D) was used in this document to identify culturally Deaf people who use a signed language for communication and who are actively engaged with the Deaf community.

Project Sponsor and Steering Committee

This project was made possible through the sponsorship of the Broadcasting Accessibility Fund. Project direction was set by a Steering Committee comprising of the following members:

- **Richard Belzile,** Project Leader and Incoming Chair, representing the Canadian Association of the Deaf (CAD)

- **James Hardman**, representing the Canadian Hearing Society (CHS).

- **Beverly Milligan**, representing PAVO Digital

- **Melina Nathanail**, representing 3Play Media

- **Anne Marie Langlois,** representing the Canadian Hard of Hearing Association (CHHA)

- **Gavin Lumsden,** representing Rogers Media

- **Brian McGinley,** representing Bell Media

- **Neil Watson,** representing CBC

Project Description

To help us explore the user experience of fast-paced live sports captioning, and to consider the impact of reducing the number of WPM used for these events, we compared user experiences of two forms of captioning: first, the traditional PBP format of captions, and second, the CO portion of the broadcast.

The goal was to explore the potential impact on quality, satisfaction, and comprehension that the two captioning styles had on the viewers' experiences. Comprehension scores were determined based on the results of a conversation-style interview that we conducted with each participant following the video portions of this experiment.

The eye tracking portion of the study provided us with insights into the user's eye gaze behaviours, focusing on how long they focused on different areas of the screen, and where their eyes tended to focus during the gameplay: either on the areas where the gameplay was taking place or on the captions. Two groups who were most familiar with closed captioning (CC) for live sports, namely, D and HoH people were the target users for this study. We recruited participants who were already familiar with watching live sports with captions, and those who were fans of live hockey or basketball broadcasts.

About the Research

Research Design

The study was designed as a mixed methods experiment, combining precise eye-tracking data for participants taking part in the study with qualitative preference and comprehension data. The study design was approved by the Ryerson (now Toronto Metropolitan University) Ethics Board. Participants watched the video segments for the sports, and subjective ratings of quality, engagement and interest were collected through questionnaires. In addition, a conversation-style interview was conducted with participants to determine their comprehension of the game and non-gameplay information; key events that occurred during the gameplay (e.g., penalties, goals) and comments that were made by the announcers about non-gameplay information were discussed in this interview. The study included two segments of the captioned games (either hockey or basketball), with each segment presenting a different caption condition to the participants (PBP or CO). The entire study took about 1 hour for each participant to complete and was conducted at the Toronto Metropolitan University campus in one of the usability labs. Participants were given an honorarium of $60 plus transportation costs for their participation.

Research Execution

The study lab was set up as follows:

- Tobii eye-tracking equipment that was attached to a PC computer
- A 24" computer monitor was used to present the game clips
- A camera positioned on the computer monitor to record the participant during the session
- A comfortable chair, which was adjustable
- Adjustable dim light pole allowing the participants to see the video clip clearly
- A secondary area set up as a lounge where the discussions were held
- Refreshments including drinks and snack

Study Procedure

Participants were asked to sit in front of the computer monitor containing the eye-tracking hardware and software. Once comfortable, the eye tracker was calibrated for the participant's eyes. A camera was attached to the computer monitor to video record the participants during the sessions. Participants were asked to watch two video clips from a single hockey/basketball game, one with PBP captions and one with CO captions. The caption style was randomly ordered between the two clips. Between each video, participants were asked to complete a mid-study questionnaire to convey their experience with each condition used in the study. Once the second video was watched, participants were asked to complete the post-study questionnaire, invited to the lounge area, and offered a snack and a refreshment before starting the post-study interview. The post-study

interview was constructed in an atmosphere for casual conversation to arise, to simulate how sports fans talk about the game after watching it. Once the post-study interview was completed, participants were paid the honorarium and thanked for their participation in the study.

Comprehension Discussion

To determine if the participants were able to comprehend what was going on in the game, we used a conversation-style interview, that aimed to prompt discussions about specific parts of the game, with the intent of determining if the captions or the gameplay were integral in supporting the user comprehension of the game action.

Specific Considerations for Deaf and Hard of Hearing Participants

Each participant was provided with the option of having CART service, or an American Sign Language (ASL) interpreter present during the study to effectively communicate with the researcher. The researchers who ran the study were hearing and Deaf. Volume on the games was turned down to an inaudible level so that the team could ensure participants were relying on the captions to experience the gameplay actions and commentary.

Demographics

Table 1: Summary of the participant's profile across different groups.

	Hearing condition	Count	Sum	Total number of participants
Basketball	**Deaf**	5 (18%)	13 (48%)	
	Hard of Hearing	8 (30%)		27
Hockey	**Deaf**	11 (41%)	14 (52%)	
	Hard of Hearing	3 (11%)		

Table 2: Breakdown of the sample by different age groups.

Age group	Deaf	Hard of Hearing	Total
18-24	1 (3.7%)	2 (7.4%)	3 (11.1%)
25-34	0	2 (7.4%)	2 (7.4%)
35-44	6 (22.2%)	1 (3.7%)	7 (25.8%)
45-54	4 (14.7%)	3 (11.1%)	7 (25.8%)
55-64	4 (14.7%)	0	4 (14.7%)
65+	1 (3.6%)	3 (11.1%)	4 (14.7%)
Total	16 (59.3%)	11 (40.7%)	27

Table 3: Breakdown of the participant by gender.

Gender group	Deaf	Hard of Hearing	Total
Male	10 (37.0%)	7 (25.9%)	17 (63.0%)
Female	5 (18.5%)	4 (14.8%)	9 (33.3%)
Transgender	1 (3.6%)	0	1 (3.6%)
Total	16 (59.3%)	11 (40.7%)	27

## Explaining the Two Caption Styles

This study introduced a novel form of captioning style as an approach to reducing the number of words presented, without reducing the accessible information. Table 4 provides a breakdown of the differences in technical characteristics of the different clips by caption condition for each sport.

Table 4: Differences in the caption styles that participants experienced and commented on during the study.

|	Basketball		Hockey	
Caption condition	**CO**	**PBP**	**CO**	**PBP**
Average delay	~ 2 seconds	~ 4 seconds	~ 2 seconds	~ 4 seconds
Position	Bottom centre	Top justified	Bottom centre	Bottom justified
Average speed (WPM)	~ 210	~ 165	~ 153	~ 137
Contents	Commentary only	All comments	Commentary only	All comments

To illustrate how the different caption styles appear on the screen, Figure 1 shows that during a play-by-play call, there is no caption present in the CO condition.

Figure 1: On-screen appearance of CO captions and PBP captions during gameplay.

| Hockey CO condition | Hockey PBP condition |
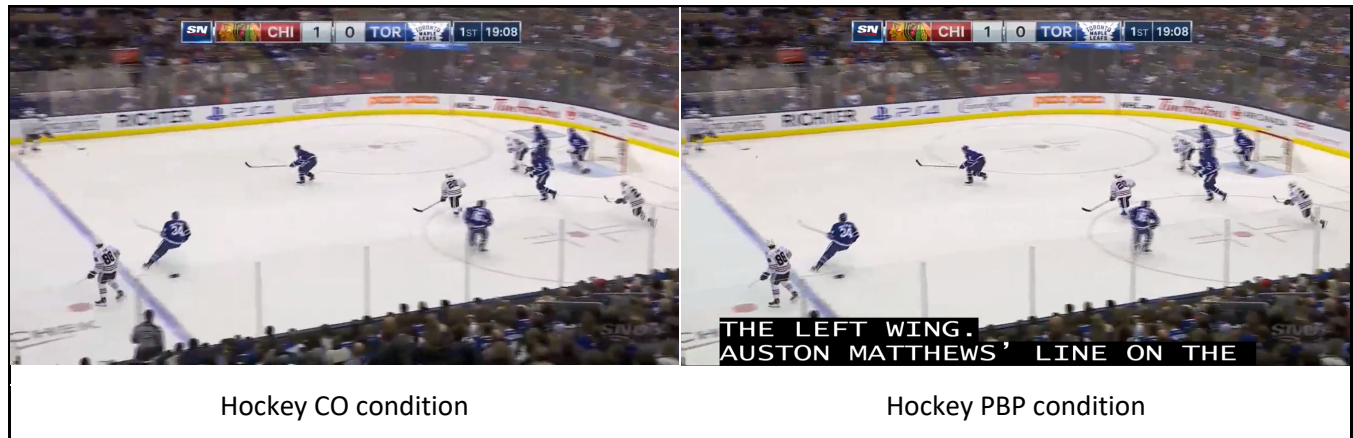
When the gameplay pauses/stops, the CO captions are displayed, as per the commentary taking place during the game. The PBP captions are always displayed whenever one of the commentators is speaking. Note that in the images below, the CO captions are centred on the page, while the PBP captions on the right are spread across the entire top part of the screen.

Figure 2: On-screen appearance of the CO captions and PBP captions during commentary.

| Basketball CO condition | Basketball PBP condition |
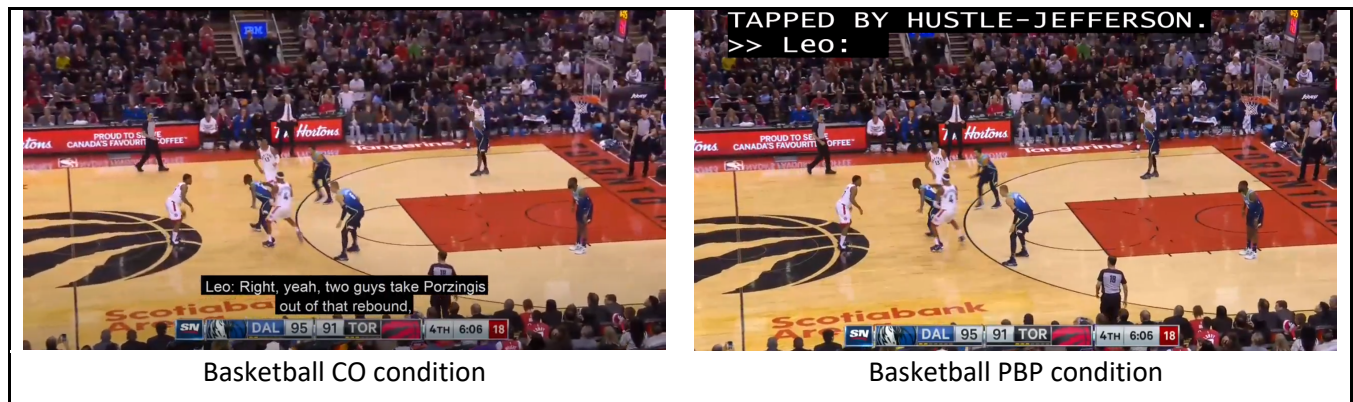
## Comparing CO and PBP caption conditions of Live Captioning

Using the Tobii eye-tracker, the duration and frequency of visits in the Areas of Interest (AOI) was captured, where each visit was defined as the sum of fixations, saccades, and glances in msec. The Areas of Interest were defined as the area containing the captions and the areas not containing captions (e.g., gameplay area, stands, scoreboard, etc.). Based on the eye tracking records, it was possible to perform statistical analysis, which revealed that there were no statistically significant differences in the duration and frequency of visits between CO and PBP captions regardless of the participant's hearing status. Figure 3 shows a scatter diagram of how long, as a percentage of total time, participants' eyes were looking at the two areas of interest between D/HoH viewers for Hockey and Basketball combined.

Figure 3: Scatter diagram of the percent of the time that D/HoH viewers spent looking at the captions and other areas of the screen for both caption conditions.
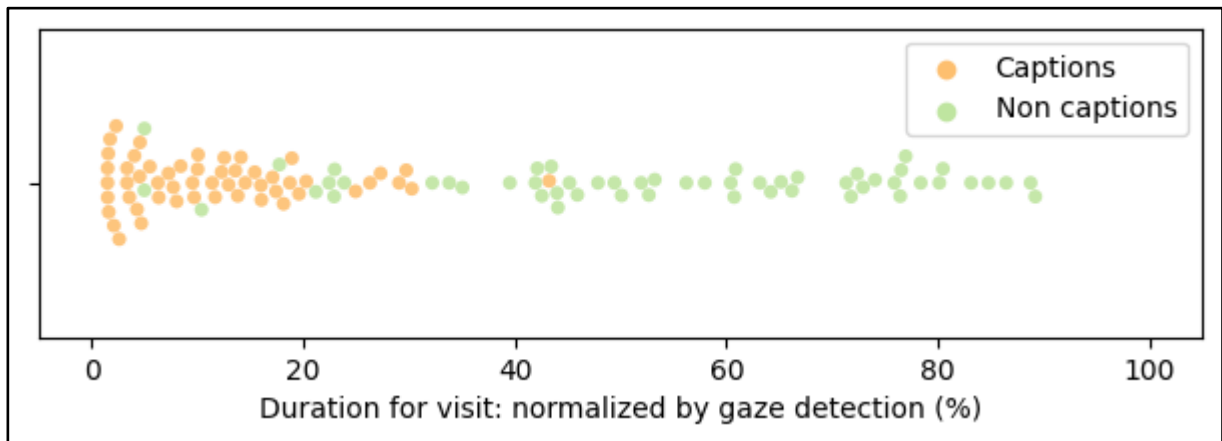

The duration for visits for CO captioning was 11% and 10% for PBP captioning. Although statistical significance was not found, Figure 4 shows the average values for D/HoH views for each caption condition. Participants seem to visit the CO captions more than PBP captions regardless of the sport or hearing status. One way to interpret this would be that PBP captions interfere with watching the gameplay so that participants did not spend much time on or visit the area of the screen at all regardless of whether there were captions for PBP or for the CO. Gameplay captions were not present in the CO condition so participants may have thought that these captions were worth reading. The average number of visits for the CO captions was 124, whereas the average number of visits for the PBP captions was 106 (see Figure 5).

Figure 4: Mean and standard deviation (lines) of the duration for visits on the two areas of interest for CO and PBP conditions
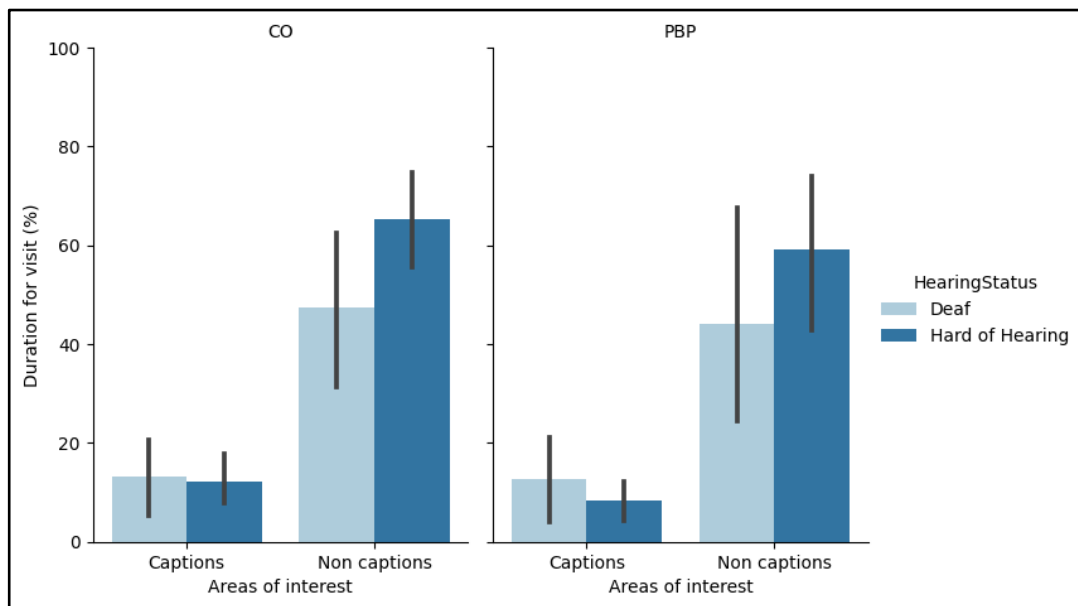
Figure 5: Mean and standard deviation for the frequency of visits on the two areas of interest for CO and PBP conditions
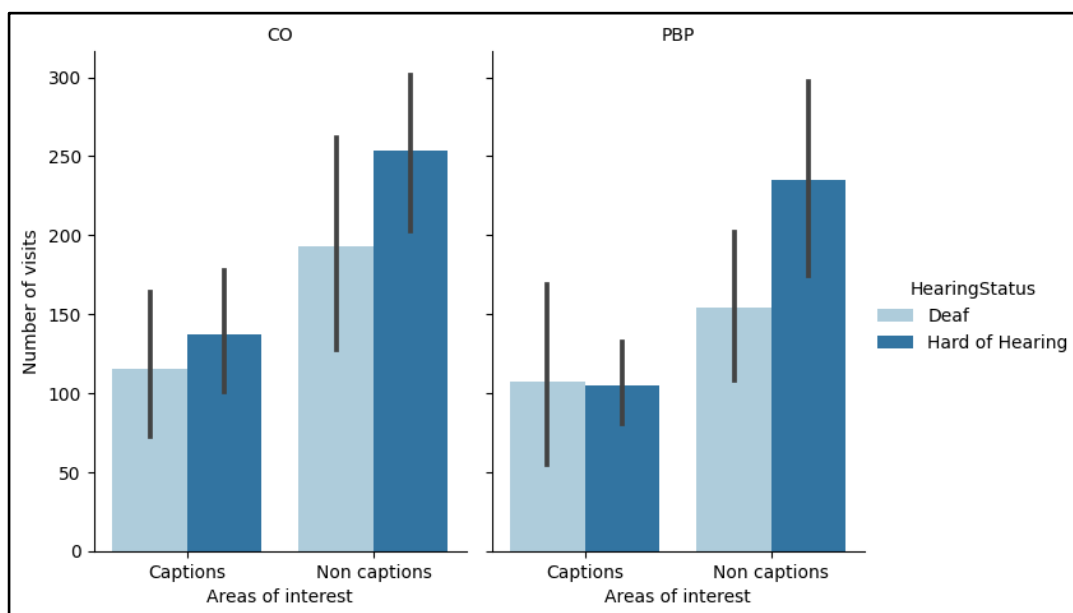
Given that the PBP condition represented conventional captions that participants were used to watching, finding no statistical difference between the two conditions revealed the potential for using CO captioning as the novelty effect likely influenced viewer experience negatively (e.g., "this is different from what I am used to"). This potential appears to be more evident from the positive participant feedback on CO captioning compared with the negative feedback on PBP captioning after viewing the two conditions.

> *"I would prefer the second clip's captioning [referring to CO]. It was good because I understood the game. I do not need the Play-by-Play. When live play is happening the commentators are not even talking about that, so I get distracted easily. I like [CO] because there were no captions during the live play and because I can't hear, I cannot tell if they are talking, and I could catch up on what was going on when live play was dead (replays and timeouts)."*

Another participant described how CO captioning helped follow the game without being worried about missing out.

> *"I could concentrate on the game and not worry that I was missing out on something by trying to keep up with both (...) It was at an eye level that even when the play was starting, I could finish reading and still see what was happening in my peripheral [vision]."*

With respect to PBP captioning, participants mentioned how difficult it was to follow the game and the captioning displayed on the screen. Some of the participants' comments about PBP captioning were:

> *"too distracting" / "stressful, too much" /*

"seriously overwhelming" / "a lot of work for my eyes" / "on-going" /

"a lot of visual back and forth [looking between the captions and game screen]" /

"feel[ing] a little stressed because I do not want to miss the important information before it disappears" /

*"it's like reading a book and when I looked up, I would miss the captions"*

Some participants also added how the conventional captioning [PBP] resulted in them avoiding the captions as a whole and going *"straight to watching the game."* The comments from participants support the aspect of interpreting duration for visits in the non-caption area, where participants want to focus on watching the actual game visuals and play, with fewer distractions if possible.

However, not all comments for CO were positive and not all comments on PBP were negative. One participant preferred PBP captioning because she was used to watching PBP captioning, especially because she *"knew that everything was on it,"*. Other comments referred to how CO captioning missed the player identification/name and numbers as participants used the captioning to recall names.

## Caption Style Preferences, Quality Satisfaction and Viewing Experience

From the questionnaire responses, most (65%) participants selected no preference when asked about which caption condition they preferred. The remaining participants preferred CO captioning (27%) more than PBP captioning (8%). Sixty percent of the D participants and 73% of the HoH participants showed no preference. In the group that stated a preference, about 33% of the D participants and 18% of the HoH participants preferred CO captioning, and 7% of D and 9% of HoH preferred PBP captions. Figure 6 shows the preference distribution among D/HoH participants for both sports combined.

**Figure 6: Number of participants by caption preference.**
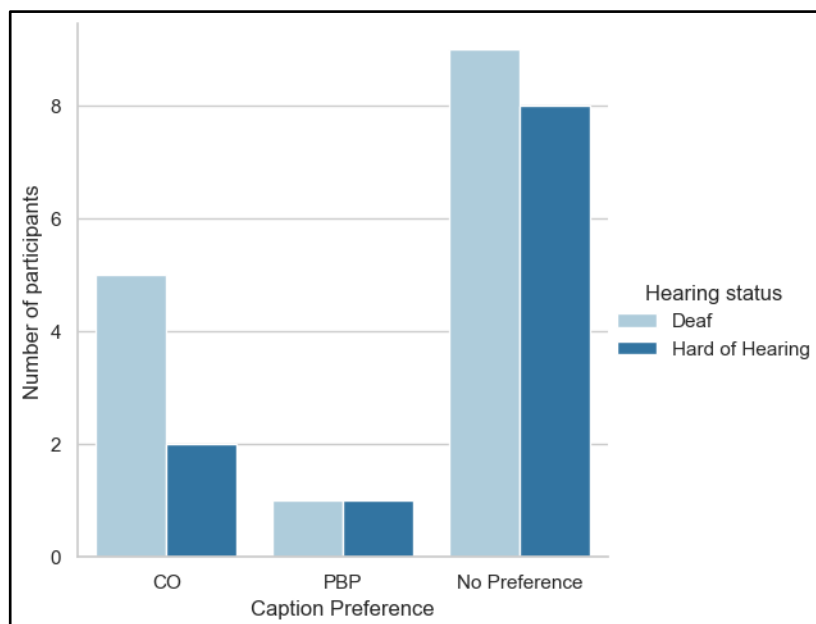
The perceived quality, viewing experience, and caption helpfulness ratings after watching the clips were higher for the CO condition than the PBP condition although these differences were not statistically significant.

Figure 7: Mean and standard deviation of the quality satisfaction rating, viewing experience rating, and caption helpfulness ratings where a rating of 1 is low (e.g, very low quality) and five is high (e.g., very high quality)
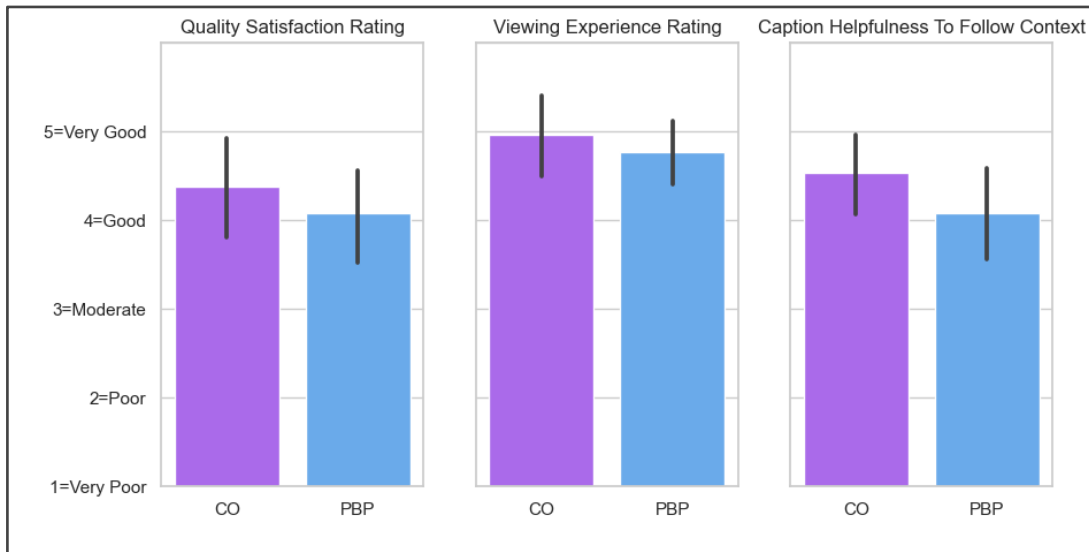
Participants were also asked to rate the quality of four factors (synchronization delay, presentation speed, speaker identification, and caption positioning) along with an overall quality satisfaction rating. There were correlations between the ratings on each of the four quality factors of live CC and the overall satisfaction rating. Overall, the four ratings were correlated with the overall caption quality satisfaction rating, suggesting their importance in judging perceived quality by D/HoH audiences.

Figure 8: Heatmap diagrams of Kendall's tau correlation coefficient from the ratings between CO and PBP conditions.
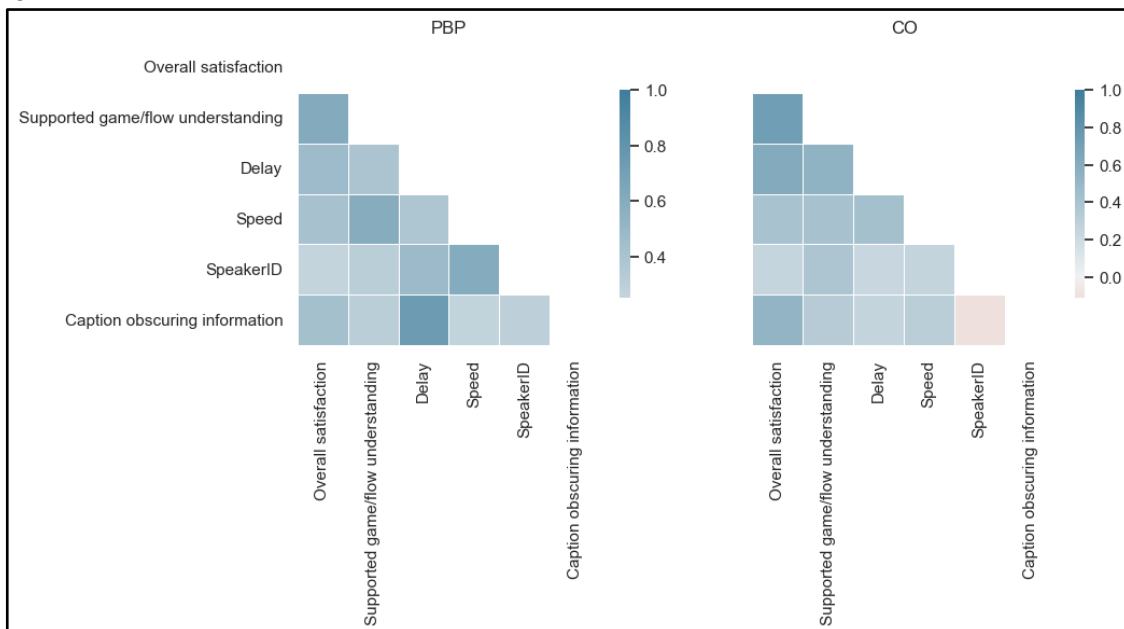
Comprehension

We conducted a post-study interview to assess comprehension of what participants understood from the game. Two types of topics were discussed: gameplay (GP) topics, which were related to the activities that happened during the gameplay (e.g., who scored), and questions related to what the commentators (CM) said during the non-gameplay commentary (e.g., opinion on the quality of the refereeing for the game). The interview was designed to simulate a casual conversation setting similar to how people discuss their ideas after watching a game at a sports bar or watching with friends at home. HoH participants had higher comprehension scores than D participants, especially for CM topics. From statistical analysis, no statistically significant relationships were found between captioning conditions and comprehension scores. However, there seemed to be higher scores when participants discussed topics related to CM compared with the GP topics.

Figure 9: Mean and standard deviations of comprehension scores from GP and CM topic questions.
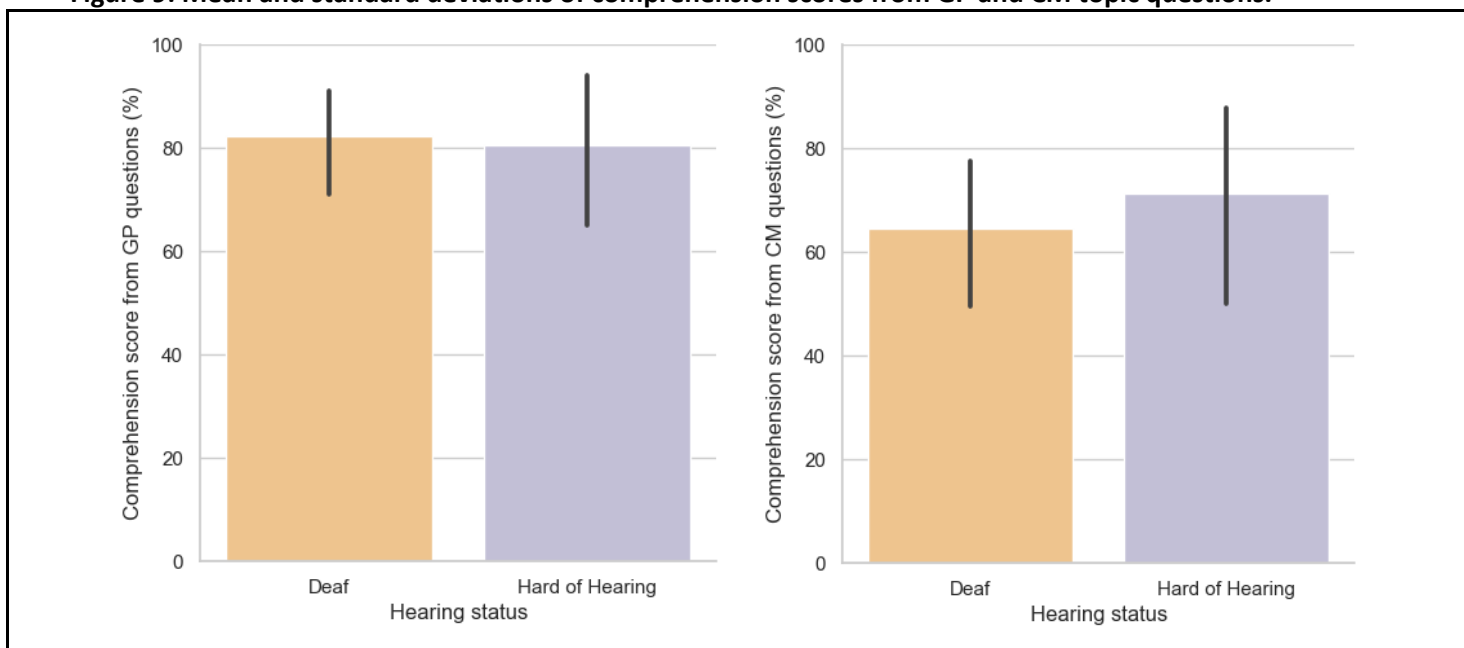
Limitations

While the study provided valuable insights into eye behaviour from watching captions in different conditions, captioning preferences, and essential quality factors of CC from D and HoH participants, there were several limitations that should be acknowledged.

One of the major limitations of this study is the small sample size of participants. Due to the difficulty in recruiting D and HoH individuals who are interested in watching fast sports broadcasting, we had a limited number of participants. This small sample size limits the generalizability of our findings, as the patterns of the study participants may not be representative of the broader population of D/HoH viewers. Future studies should attempt to recruit a larger sample to improve the reliability and generalizability of the findings.

Another limitation of the study is the high variability in the eye-tracking data across participants. We observed a wide range of variances in the fixation patterns between participants, indicating that there may be individual differences in watching captions. The variability in the data may have also been affected by individual differences in cognitive or attentional abilities, which could have affected the way participants processed

different conditions of captioning. It may also be indicative of the limitations of eye-tracking equipment which must be calibrated for each individual. If the person moves or changes their head position, the calibration may be lost, which results in fewer data points.

Furthermore, another limitation includes how captioning styles were created differently. Unlike PBP captioning, which is a standard and familiar captioning practice, CO captioning was a new idea for the captioner. They needed to change their way of captioning in order to produce it. The stimuli provided for this study reflected live captioning practice and were created by an experienced live captioner. However, there were differences in caption positioning and capitalization. The position of PBP captions varied to avoid obscuring information on the screen (e.g., scoreboard), and captions were all capitalized. In contrast, CO captions were always appearing at the same position (bottom centre), and the captions were in the mixed case except for the starting character, which may have affected viewers' opinions of the caption readability and quality, although these relationships have not yet been studied.

Conclusion

This study was another first of its kind in Canada and has revealed some insights into the user experience of watching fast-paced live sports captioning for hockey and basketball. While the industry and viewers know the need to address the general live captioning issues such as synchronization delay, accuracy, speed of captions, speaker identification, caption placement, and readability, a specific aspect of live captioning for fast-paced sports involves the challenges of generating and reading the high number of WPM. In this study, we introduced a new style of captioning for sports broadcasting live-captioning, which only provided captions for commentary with fewer words than the conventional Play-By-Play style.

Results from the eye-tracking portion of the study showed that there were no statistically significant differences between the two captioning styles. However, the descriptive analytics revealed the difference in duration and frequency of visits, where participants spent longer watching and visited more frequently on CO captioning. Furthermore, participants' responses in caption quality satisfaction, viewing experience, and caption support in understanding, showed evidence of preferences for the CO captions. In addition, while many reported no preference for the style of captioning, participant comments allowed a deeper probe into the reasoning behind subjective preferences. It was surprising to find that participants had a considerable amount of positive feedback for CO captioning from this short study despite the study being skewed towards PBP because it was, up until now, the only type of captioning available to viewers.

Finally, the findings from this study provide sufficient evidence to pursue the notion of the commentary only style of captioning in a longitudinal study to determine whether CO captioning is viable and whether viewers are satisfied after overcoming the novelty of this style of captions. In addition, a longitudinal study could indicate whether CO captions sufficiently reduce some of the tradeoffs between speed, accuracy and delay in fast-paced live sports captioning and improve the viewer experience.