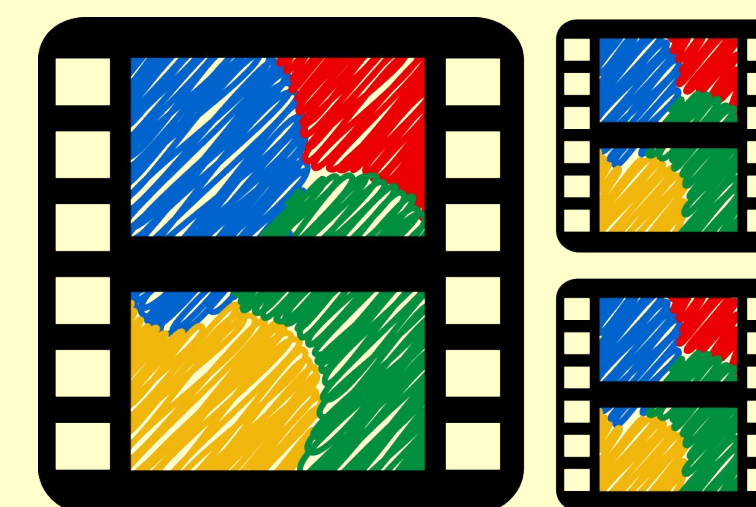# Impact of Punctuation Quality on User Experience in Closed Captioning

REU Accessible Information and Communication Technologies Site, Gallaudet University

Promiti Datta, University of Toronto
Amit Balchandani, Gallaudet University
Pablo Jakubowicz, Gallaudet University

Abraham Glasser, Rochester Institute of Technology
Dr. Raja Kushalnagar, Gallaudet University

1. Background and Objectives

- ASR Technology is creating many new options for captions production
  - Some of them systems included punctuation while other do not
  - We show how punctuation is related to the "readability" experience
  - We hope this encourage ASR developers to include punctuations in their systems
- Captions are widely judged by the WER standard

$$WER = \frac{S + D + I}{N}$$

  - **Word Error Rate** = (Sub + Del + Ins) / N where N is the total number of words in the reference transcript
  - WER calculations does not include punctuation usage
- **Punctuated WER** (PWER) - Account for punctuation inconsistencies by turning periods, commas, and question marks into words, and recalculating WER
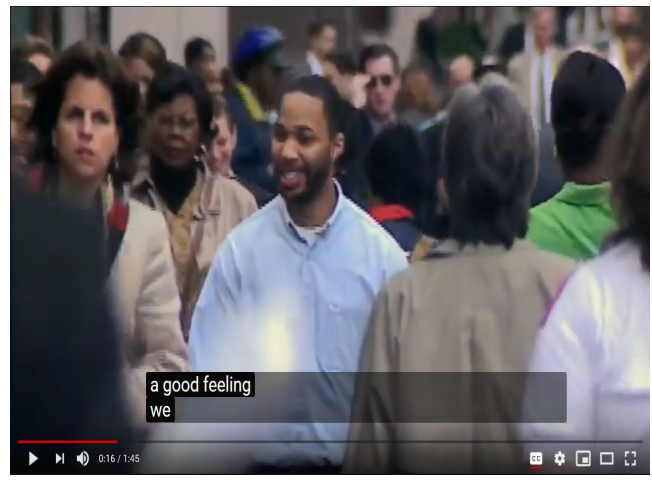
ex) Good morning, welcome back. → Good morning COMMA welcome back PERIOD

RQ: How does punctuation in captions affect the readability and experience of Deaf / Hard of Hearing users watching live television and news broadcasts?

Hypothesis: Punctuation errors do detriment the experience of reading subtitles for Deaf / Hard of Hearing people.

2. Methodology

Types of Captions:

Automatic Speech Recognition

NO Punctuation — YouTube (auto-generated) "a good feeling we"

Punctuation — Live Transcribe (Google's app) "a good feeling."

Human

Live Human - No Punctuation "GOOD FEELING >>WE WERE ALL JUST"

Live Human "GOOD FEELING."

- Collected live TV samples and curated the following types of caption for each video (shown in chart above):
  - Live Human - LH
  - Live Human with Punctuation Removed - LH0
  - Google Live Transcribe - LT
  - YouTube Auto Captions - YT
- Showed each participant 4 videos with 4 different caption types .
- Participants completed a short survey after each video asking about the caption readability, understandability, etc.
- Rotated videos and caption types for each participant in a Latin Square ordering
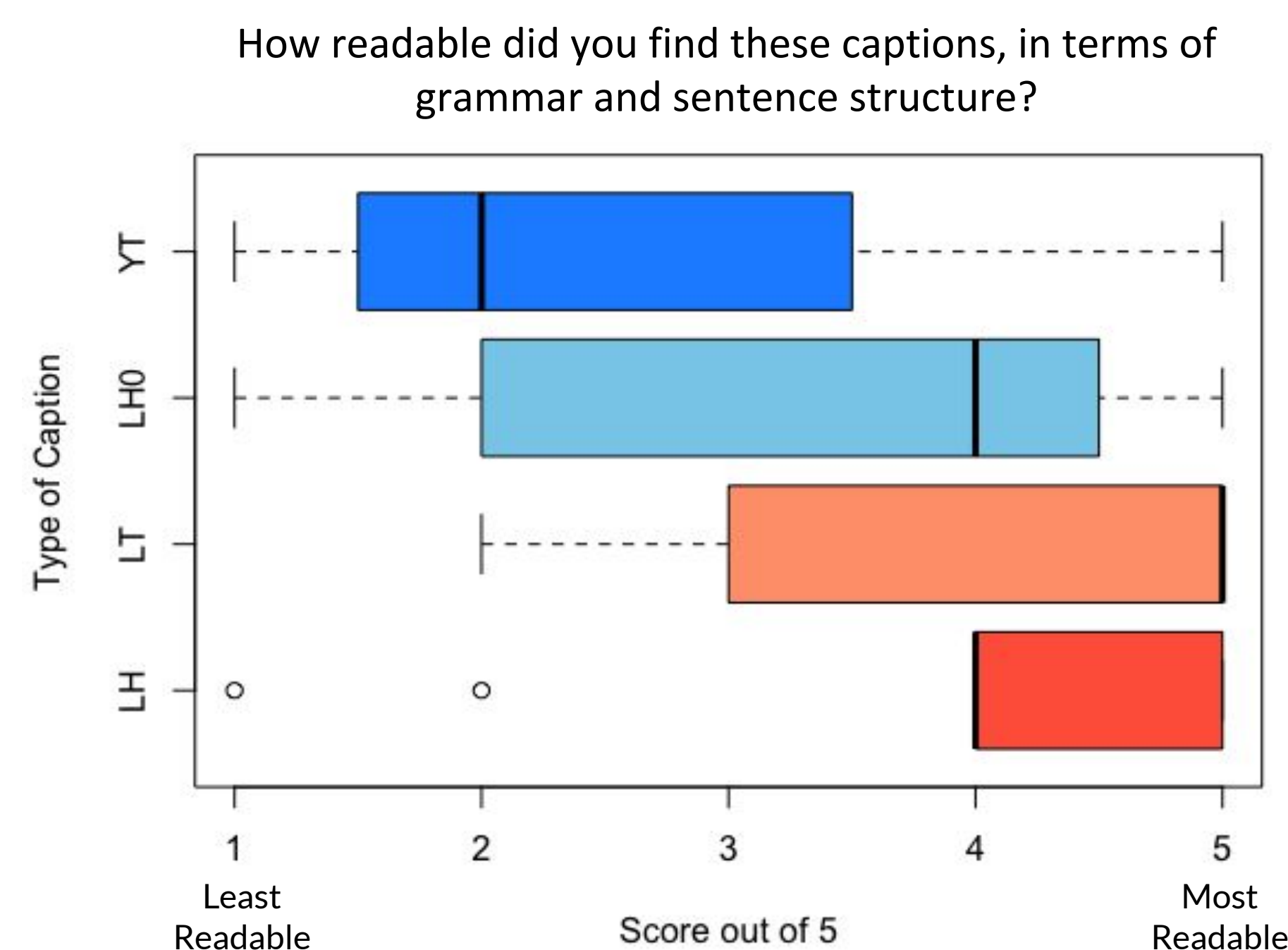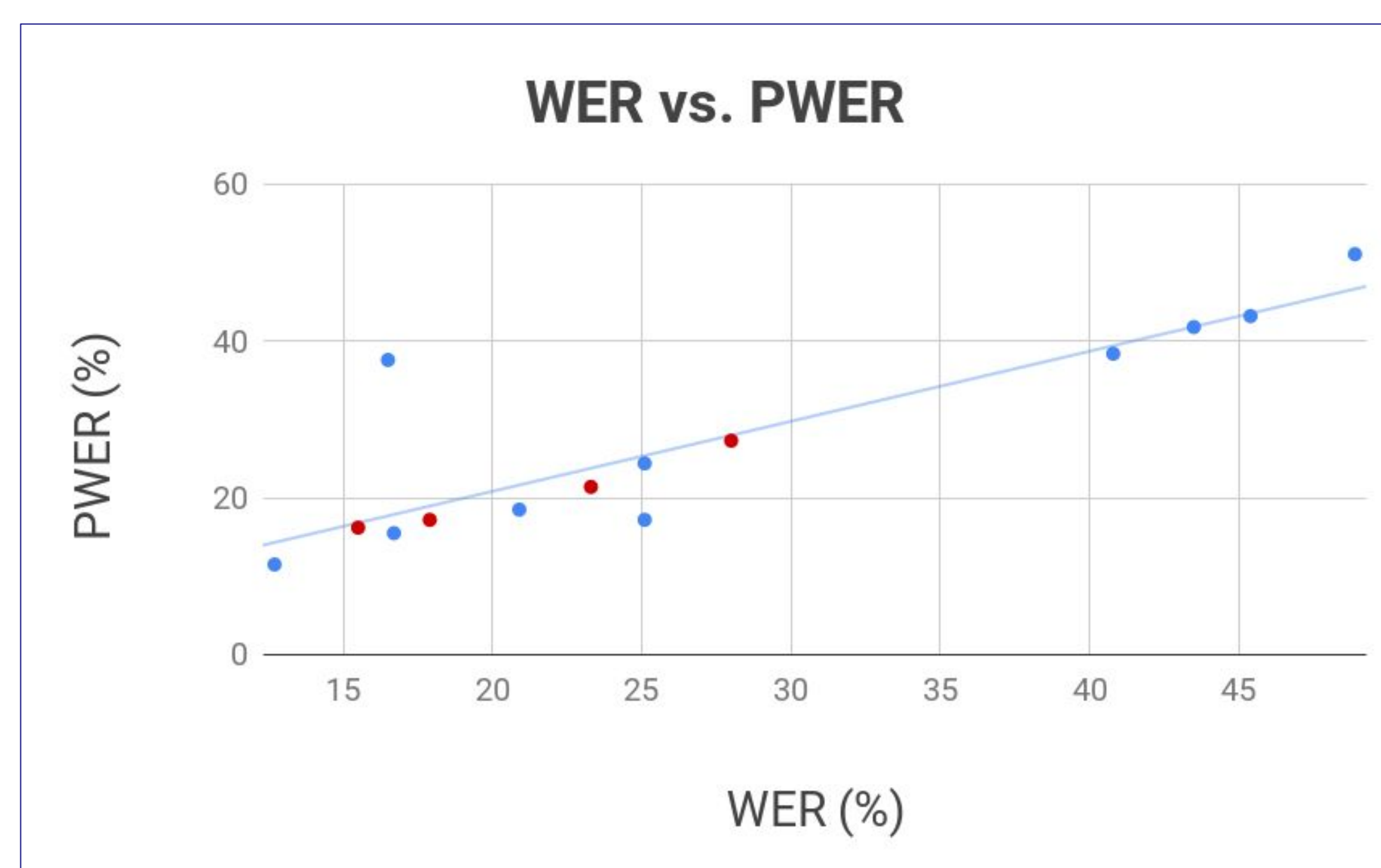
Figure 1

How readable did you find these captions, in terms of grammar and sentence structure?

Figure 3

WER vs. PWER

3. Results

- **Live Human** vs. **Live Human w/ No Punctuation**:
  - **Not enough significance** of punctuation or not within **human generated captions** (p-value = 0.054)
- **Live Transcribe** vs. **YouTube**:
  - **Significant difference** in readability for punctuated vs. non-punctuated **ASR captions** (p-value = 0.001)
- **Live Transcribe** vs. **Live Human w/ No Punctuation**:
  - **Significant difference** means that people rate ASR captions with punctuation over human captions without punctuation (p-value = 0.048)
- Overall **Punctuated** vs. **Non-Punctuated**:
  - Punctuated captions are better rated than non-punctuated caption (p-value = 0.026)

4. Qualitative Feedback

Punctuated Captions Improve Understandability:
- "When I read the non-punctuated captions, they look like run-on sentences and I have a **hard time trying to figure out when they stop talking**"
- "[Punctuated Captions] also help me **separate concepts, sentences, paragraphs** and so on. It makes everything much more transparent"

Less Cognitive Burden:
- "For non-punctuated captions, it made me **lose motivation** to understand everything because I lost track."
- "[Non-Punctuated Captions] **wear my eyes out** when I keep reading and notice there is no period. It affects my writing and reading skills."

Figure 2
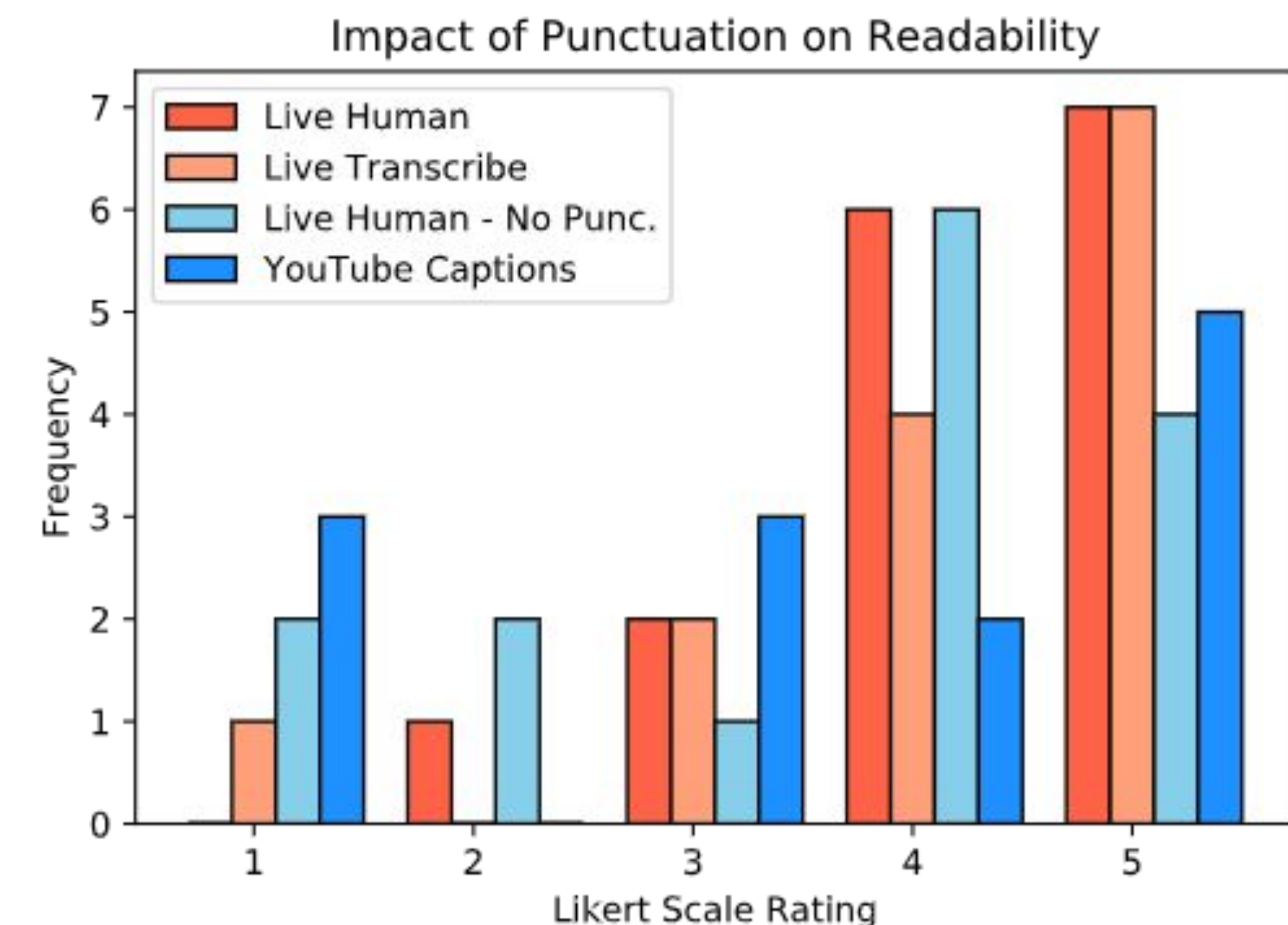

Impact of Punctuation on Readability

Figure 1: Participants rated the readability of the captions after each video they watched on a scale of 1 - 5. The summary for each caption type is shown.

Figure 2: Participants rated the impact of the punctuation or lack of punctuation on the caption readability. Low scores meant that the punctuation quality did not have much impact on readability, whereas high scores meant punctuation quality had a strong positive or negative impact, depending on whether there was punctuation present or not.

Figure 3: Relationship between WER and PWER for punctuated Live Transcribe captions coming from television clips with similar content. Data points in red represent the videos used in the user experiment.

5. Conclusions

Effect of Punctuation on Reading Experience:
- Punctuation errors do detriment the captioning experience for Deaf / Hard of Hearing people.
- People prefer "good quality" ASR captions with automatic punctuation over human generated captions with no punctuation.

Relationship Between WER vs. PWER and Other Possible Metrics for Punctuation Quality:
- *Figure 3* shows that PWER and WER have a strong linear correlation (r = 0.867)
- In the future, we would like to analyze the effect of transcript length and number of punctuation marks on these metrics
- We also want to incorporate context into our evaluation of punctuation

6. References

[1] Abraham Glasser, Kesavan Kushalnagar, and Raja Kushalnagar. 2017. Deaf, Hard of Hearing, and Hearing Perspectives on Using Automatic Speech Recognition in Conversation. In *Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility* (ASSETS '17). ACM, New York, NY, USA

[2] Duerstock, Bradley S., Rohit Ranchal, Yiren Guo, Teresa Taber Doughty, J. Paul Robinson, and Keith Bain. "Assistive Notetaking Using Speech Recognition Software."

[3] Kafle, S., Huenerfauth, M. (2016) Effect of Speech Recognition Errors on Text Understandability for People who are Deaf or Hard of Hearing. Proc. SLPAT 2016 Workshop on Speech and Language Processing for Assistive Technologies, 20-25,

[4] Apone, Tom, Brad Botkin, Marcia Brooks, and Larry Goldberg. "Caption Accuracy Metrics Project." 1-16.

[5] Larwan Berke, Sushant Kafle, and Matt Huenerfauth. 2018. Methods for Evaluation of Imperfect Captioning Tools by Deaf or Hard-of-Hearing Users at Different Reading Literacy Levels. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (CHI '18). ACM, New York, NY, USA, Paper 91, 12 pages

7. Acknowledgments

Contacts:
Promiti Datta, promitid@gmail.com
Amit Balchandani, amit.balchandani@gallaudet.edu
Pablo Jakubowicz, pablo.jakubowicz@gallaudet.edu