Comparison of Speaker Cues for Deaf Viewers in Virtual Reality

REU Accessible Information and Communication Technologies Site, Gallaudet University

Dr. Raja Kushalnagar, Gallaudet University
Abraham Glasser, Rochester Institute of Technology | National Technical Institute for the Deaf
Edward Riley, Rochester Institute of Technology | National Technical Institute for the Deaf
Kaitlyn Weeks, Gallaudet University

1. Background and Objectives

Mixed Reality technology is growing rapidly. Currently, the infrastructure **lacks accessibility** options for Deaf and Hard of Hearing (DHH) people. Research demonstrates that DHH people experience a cognitive burden in **identifying** and **tracking** multiple speakers while reading captions.

- We aim to establish solutions for this problem by:
  - Creating 3-4 approaches to assist the user in **identifying** the **speaker** among a panel within the virtual reality classroom environment
  - Adding captions to videos

Research Question:

- What is the most efficient approach for Deaf and Hard of Hearing adults to track and distinguish who is **speaking** with their voice in a panel within a **Mixed Reality** environment?

Hypothesis:

- If we include a speaker-identifying **indicator** in Virtual Reality then Deaf and Hard of Hearing users will have an easier time **identifying** the **speaker** instead of a caption standalone setting.

We hope that our study motivates researchers and developers to make **Mixed Reality** technology **accessible** to not only DHH people but also those who rely captions.

2. Methodology

We are comparing the participants' responses on how comfortable they felt with the 4 different approaches.

Technology Used

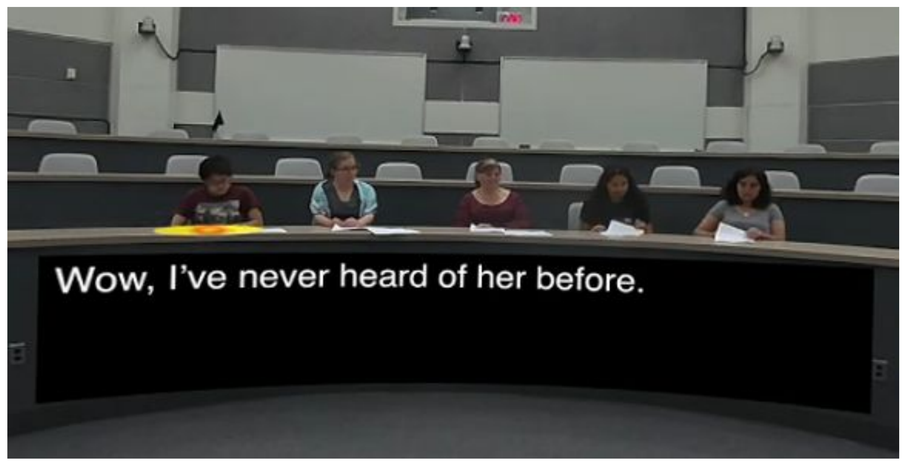
- Adobe Premiere / Adobe After Effects
- Daydream VR headset with Google Pixel phone

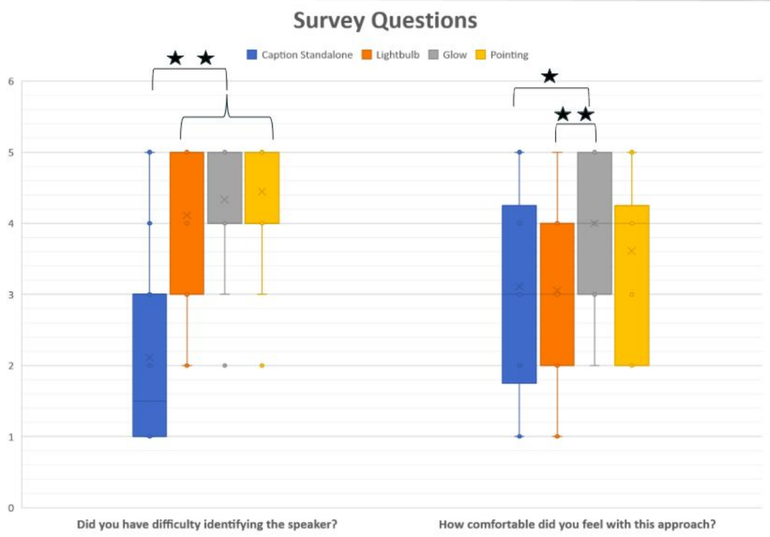
Procedure

1. Developed scripts with multiple speakers in a panel setting.
2. Recorded stimulations.
3. Added captions and indicators to the videos using software.

Participants watched videos and answered survey questions in between.

Figure 1





3. Results

Participants were asked two survey questions in between videos.

"How comfortable did you feel with this approach?"

- The T- test showed us that Caption Standalone and the indicators showed significant P-value. The Caption Standalone was significantly better (t-test, alpha = .01) than each of the indicators.
  - Average of Caption Standalone: 3.167
  - Average of Glow: 4.111
  - Average of Pointing: 3.833
  - Average of Lightbulb: 3.333

"Did you have difficulty identifying the speaker?"

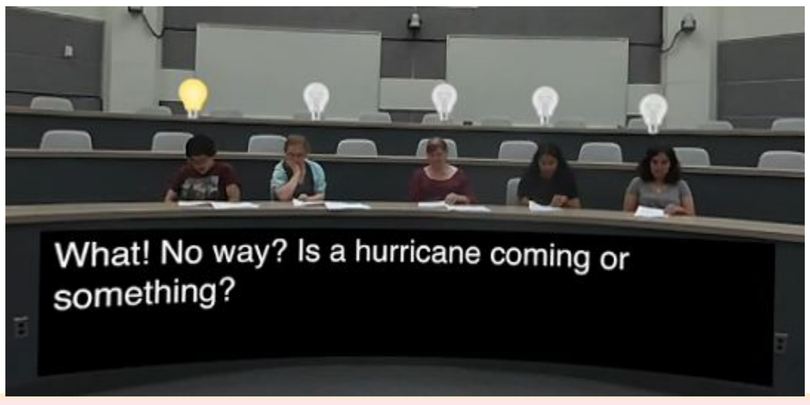
- The T- test showed us the Lightbulb and Glow have a P-value of 0.0069. Though, the T-Test showed us the the Caption Standalone and Glow have a P-value of 0.0761.
  - Average of Caption Standalone: 2.111
  - Average of Glow: 4.333
  - Average of Pointing: 4.444
  - Average of Lightbulb: 4.111

4. Qualitative Feedback

- Concerns with Blurriness of Captioning and Actors - 4
- Captions were too long and too wide to read - 7
  - Some participants expressed their concerns having to move their head to read the captions
- Wish that name/color were added to Caption - 3
  - **Amy:** Hi, my name is Amy and I like programming.
- Good font and color choices - 2

"Awesome experience to see how virtual reality can make a difference towards the Deaf/HOH community."

Figure 2



5. Conclusions

- Indicators are significantly preferred over none.
- **Glow** is generally preferred choice however the data that compares which indicator participants prefer the most is insignificant.
- 16 out of 18 participants wish to see captioning in VR
- 13 out of 18 participants hope to see indicators in VR. The remaining 5 participants said maybe.
- Pursuit of developing indicators for Deaf and Hard of Hearing Virtual Reality users in the future is strongly encouraged.
- Captions should be considered to be customizable as people have various needs and preferences.

6. References

1. Wolfe, K. T., & Dingman, B. (n.d.). RTTD-ID : Tracked Captions with Multiple Speakers for Deaf Students RTTD-ID : Tracked captions with multiple speakers for deaf students.
Brown, A., Turner, J., Patterson, J., Schmitz, A., Armstrong, M., & Glancy, M. (2017). Subtitles in 360-degree Video. 3–8. https://doi.org/10.1145/3084289.3089915
Sharma, R., Huang, T. S., & Pavlović, V. I. (2011). A Multimodal framework for Interacting with Virtual Environments. https://doi.org/10.1007/978-1-4613-1447-9_5
Kushalnagar, R. S., Behm, G. W., Kelstone, A. W., & Ali, S. (2016). Tracked Speech-To-Text Display. (December), 223–230. https://doi.org/10.1145/2700648.2809843
Kushalnagar, R., & Kushalnagar, P. (2012). Computers Helping People with Special Needs. 7382(December 2016). https://doi.org/10.1007/978-3-642-31522-0
Sharma, R., Zeller, M., Pavlovic, V. I., Huang, T. S., Lo, Z., Chu, S., … Schulten, K. (2000). Speech/gesture interface to a visual-computing environment. IEEE Computer Graphics and Applications, 20(2), 29–37. https://doi.org/10.1109/38.824531
Berke, L., Caulfield, C., & Huenerfauth, M. (2017). Deaf and Hard-of-Hearing Perspectives on Imperfect Automatic Speech Recognition for Captioning One-on-One Meetings. 155–164. https://doi.org/10.1145/3132525.3132541
Sharma, R., Huang, T. S., & Pavlović, V. I. (2011). A Multimodal framework for Interacting with Virtual Environments. 53–71. https://doi.org/10.1007/978-1-4613-1447-9_5
Bachmann, D., Weichert, F., & Rinkenauer, G. (2018, July 7). Review of three-dimensional human-computer interaction with focus on the leap motion controller. Sensors (Switzerland), Vol. 18. https://doi.org/10.3390/s18072194
Martins, P., Rodrigues, H., Rocha, T., Francisco, M., & Morgado, L. (2015). Accessible Options for Deaf People in e-Learning Platforms: Technology Solutions for Sign Language Translation. Procedia Computer Science, 67, 263–272. https://doi.org/10.1016/j.procs.2015.09.270
Greenwood, T. A., Village, G., & Data, R. U. S. A. (2009). Metagging of Captions. 2(12).
Morton, D. (1932). Videophone with Continous Speech-to-subtitle translations. 2(12).
Publications, O. (2014). Method and system for testing Closed Caption content of Video Assets. 1(12).
Butler, J. (2019). Perspectives of Deaf and Hard of Hearing Viewers of Captions. American Annals of the Deaf, 163(5), 534–553. https://doi.org/10.1353/aad.2019.0002
Butler, J. (2019). The Visual Experience of Accessing Captioned Television and Digital Videos. Television and New Media, 1–18. https://doi.org/10.1177/1527476418824805
Butler, J. (2018). Integral Captions and Subtitles: Designing a Space for Embodied Rhetorics and Visual Access. Rhetoric Review, 37(3), 286–299. https://doi.org/10.1080/07350198.2018.1463500
Remote, M. F. O. R. (2004). System and method for remote audio caption visualizations. 1(19).
Pavlovic, V. I., Sharma, R., & Huang, T. S. (2002). Gestural interface to a visual computing environment for molecular biologists. 30–35. https://doi.org/10.1109/afgr.1996.557240
Chang-Hyun, P., Gilsoo, J., & Young-Ho, C. (2006). Development of a Virtual Reality Training System for Live-Line Workers. International Journal of Human-Computer Interaction, 20(3), 285–303. https://doi.org/10.1207/s15327590ijhc2003_7
Wu, Y., Huang, T. S., Institute, B., & Mathews, N. (n.d.). Vision-Based Gesture Recognition: A Review.
Lwo, L., & Chia-Tzu Lin, M. (2012). The effects of captions in teenagers' multimedia L2 learning. ReCALL, 24(2), 188–208. https://doi.org/10.1017/s0958344012000067

8. Acknowledgments

The contents of this poster were developed in part under a grant from the National Science Foundation, grant #1757836 (REU AICT) and under a grant from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR grant number #90DPCP0002). NIDILRR is a Center within the Administration for Community Living (ACL), Department of Health and Human Services (HHS). The contents of this poster do not necessarily represent the policy of NIDILRR, ACL, HHS, and you should not assume endorsement by the Federal Government.

Director of REU Accessible Information and Computational Technologies, Raja Kushalnagar
Graduate Assistant, Abraham Glasser

Contacts:
Edward Riley, mason.rileyzero@gmail.com
Kaitlyn Weeks, kaitlyncweeks@gmail.com