Sign Language in Social VR - Challenges and Adaptations

Dylan R. Fox[1]
[1]XR Access, D.R. Fox Design

***Keywords*—Social Virtual Reality, VR, Sign Language, Deaf, Hard of Hearing, Mozilla Hubs, Accessibility, Communication**

Social VR systems like Mozilla Hubs offer a powerful new platform for socialization and collaboration. However, the avatar and voice-based systems threaten to leave behind people that rely on sign language, lip reading, and other non-auditory means to communicate.

At the XR Access Symposium on July 20, 2020, we attempted to set up the Hubs system to accommodate sign language interpretation. In this flash talk, I will cover some of the challenges we encountered, workarounds we developed, and suggestions for future updates so that social VR can be accessible to everyone.

Fig. 1. Virtual reality office environment featuring three people using robot avatars, with floating heads and torsos.

I. The Challenge

Deaf and hard of hearing people face many challenges in synchronous digital environments. Many social experiences rely on voice chat, which they can neither contribute nor listen to. Even when speakers are captioned and text chat is also available, typed out responses often come out too slowly to keep up with spoken conversation. Video conferencing platforms such as Zoom can enable the use of sign language, and spoken communication with non-signers via interpreters, but come with their own challenges.

Mozilla Hubs offers a new way to meet and communicate socially, and can be used on phone, desktop, or on a VR headset. However, it still largely relies on voice and typed chat. While the rough position of hands is shown for users with motion-tracked controllers, it's not enough to communicate in sign language, nor is there a quick way to communicate facial expressions that form an important subtext to signing.

II. The Adaptations

While we were able to rig a workaround in Hubs using webcam feeds displayed in summoned panels, here are some adaptations that could make social VR more accessible to people who rely on sign language.

Improved hand tracking. This is the single biggest improvement that deaf attendees asked for. Improving WebXR to capture articulated hand tracking would enable sign language speakers to communicate much more naturally. See here for how some people have adapted sign language for use in VR. [1]

Ability to communicate facial expression. Similarly, facial expression offers important subtext to sign language. Whether via facial recognition or some kind of quick select wheel, it would be an important asset for naturalistic communication.

Pinning camera displays to avatars. By allowing users to have their webcam feed be connected to and travel with their avatars, deaf users would be able to communicate in sign while retaining the benefits of a mobile avatar in a 3D world.

Interpreter co-pilots. Training sign language interpreters to use Hubs is challenging and often requires a back-channel to the deaf user anyway. Enabling a remote interpretation system, in which the interpreter sees and hears through the deaf user's avatar and can respond through their microphone, akin to Berke et al.'s "Chat in the Hat" system, could be an alternative. [2]

Automatic captioning. Using speech to text to automatically capture words spoken by other users or summoned videos would enable deaf users to understand spoken conversations. Ideally, visual indicators would point towards sources, or they could appear as speech bubbles tied to their source.

III. References

[1] Hamilton, Ian. 2020. "Sign Language in VR 'Worth Exploring' As Hand Tracking Improves." *UploadVR*, June 17, 2020. https://uploadvr.com/sign-language-vr-asl/

[2] Berke, Larwan, William Thies, and Danielle Bragg. "Chat in the Hat: A Portable Interpreter for Sign Language Users." In The 22nd International ACM SIGACCESS Conference on Computers and Accessibility, pp. 1-11. 2