

# Project Proposal: ChatAssist

Andrew Wei, James Kim, Laurissa Lam, Ron Fan  
nowei@cs.washington.edu, jamesk67@cs.washington.edu,  
laurissa@cs.washington.edu, ronbo@cs.washington.edu

April 11, 2019

## Summary

ChatAssist is an augmented reality application that displays speech bubbles above people's heads to show what they're saying and records all spoken words that are detected. ChatAssist is specifically designed for people with impaired hearing, people who are non-native speakers of a language, or for use in situations like business meetings where an automatic conversation log is useful.

## Project Description

Our goal is to connect speech to text through AR in a way that is both meaningful and accessible. Many people have trouble hearing spoken words clearly, whether due to hearing impairments or unfamiliarity with a language. Our application aims to help these people by showing a transcript of the conversation in real time so that they can still interact smoothly with others despite their various circumstances.

We plan to use a Magic Leap, Google's Cloud Speech-to-Text API, and possibly QR codes in order to accomplish this goal. We will record speech using Magic Leap's built-in microphones and convert the speech audio to text. We will then display it as a text bubble that the user can see and read. Our first milestone will be to display nearby speech as text in a virtual display at an arbitrary location. Then, we will work on picking up speech directionally so that we can draw speech bubbles with text where the speech is coming from. Our end goal is to display text next to a nearby person who is talking. We will attempt to do this using the Magic Leap's built-in microphones, but if that proves insufficient, we will use separate standalone microphones to pick up speech audio. We may use wearable visual markers (like QR codes) to help pinpoint locations to draw speech bubbles at. We hope that our application will make interpersonal communications more accessible to everyone.

## Hardware Platform/Device

We believe the Magic Leap is the most suitable device for our project from a design perspective: since talking requires focus, it's ideal for speech bubbles to appear automatically, instead of having to physically do something to see the bubbles (like holding up a phone camera). The Magic Leap also fits our hardware requirements, with its multiple built-in microphones and computer vision processor. We believe this project is possible with ARCore, but it seems less ideal since it requires constant cumbersome interaction for a process that should feel natural and automatic.