

**Thesis Portfolio**

**ASR-based System for Speech Therapy in Adults**  
(Technical Report)

**An Analysis of Advancements in Technology and Its Social Impact on the Audiology  
Landscape**  
(STS Research Paper)

An Undergraduate Thesis

Presented to the Faculty of the School of Engineering and Applied Science  
University of Virginia • Charlottesville, Virginia

In Fulfillment of the Requirements for the Degree  
Bachelor of Science, School of Engineering

Brandon Williams  
Spring, 2021

Department of Computer Science

## **Table of Contents**

Sociotechnical Synthesis

*ASR-based System for Speech Therapy in Adults*

*An Analysis of Advancements in Technology and Its Social Impact on the Audiology Landscape*

Thesis Prospectus

## **Sociotechnical Synthesis**

Children and young adults who suffer from hearing loss or auditory processing problems may experience a delay in the development of their speech given that they cannot hear quiet sounds such as “s”, “sh”, “f”, “t”, and “k”, which lead to speech impairments. The problem is that children often discontinue speech therapy once they enter upper-level schooling as these minor speech issues start to become more permanent, leaving them without access to a speech therapy solution. Previous apps found in speech therapy work by helping to fix single words or syllables at a time without considering the inclusion of others' perception of what speech-impaired users may sound like to them. In transcripts of Zoom's closed captioning services this school year, their platform occasionally outputted words that did not align with what the professor may have been attempting to say, inadvertently pointing out “quirks” in their pronunciations. This topic is worth examining in that this added context may identify words and sounds that users may not have realized they struggled with, a key mechanism which would provide more input in the speech correction process.

Applying this concept to an automated form of speech therapy, I will be proposing a mobile application would prompt the user to read aloud from a pre-selected script into their device, feed this speech signal into a highly accurate open-sourced audio-to-text algorithm that would transcribe the audio, and flag words that did not match the original transcript as “mispronounced”. Users will be given the opportunity to see what the program interpreted their “unclear” speech as, equipping them with a unique insight into “how” they may be mispronouncing sounds through the design of a human-computer friendly interface and a unique application of open-source deep learning audio-to-text algorithms.