Aphasia App

SOFTWARE REQUIREMENTS SPECIFICATION (SRS)

CS4850

Spring 2023

Professor Perry

02/24/23

Michael Clayton, Navonte Riggin, Max Dixon, Patrick Bobbie

Table of Contents

1.0 Introduction

1.1 Overview

This document outlines the requirements and constraints in the development of the Aphasia App.

1.2 Project Goals

The primary goal of the Aphasia App is to improve the speech production and day-to-day activities of people afflicted with aphasia:

- Create a tool that can be readily accessible and used any Android that allows text from outside the app to be inputted
- Detect the user's eye movements to different points on the screen and use this data to detect a particular word that the user is looking at
- Use text-to-speech functionality to read whichever word is being looked at by the user

1.3 Definitions and Acronyms

Aphasia – A condition that causes people to have limited or complete speech deficiency

CNN – Convolutional Neural Network

Deep CNN – A convolutional neural network with several hidden layers

GazeCapture - A large-scale dataset for eye tracking that contains over 2.5 million frames of data from over 1450 different participants

Face mesh – A collection of 468 3D points on the user's face that tracks the user's facial movements in real-time using machine learning allowing it to abstain from requiring a depth sensor

Facial orientation – The general direction the user's face is pointing

Target device – The device that the app will be run on

Music intonation – A pattern of speech that alternates the pitch of successive words to attain effect that is like music

Focus – The word that the user is currently looking

1.4 Assumptions

It is assumed that the user will give the Aphasia App permission to use the camera and storage.

It is assumed that the user will input text into the app.

2.0 Design Constraints

2.1 Environment and System

The application will be run on physical Android devices. The app will be developed specifically on Android version 10, which is also known as Android Q, but it should be compatible with later versions considering Android's backwards compatibility. Android versions 9 and prior will not be tested, so functionality cannot be guaranteed. The app also requires a functional camera to properly track eye movements and a functional speaker for text-to-speech to work properly.

2.2 User Characteristics

In general, the users of this app will be those that have limited speech deficiency in English. Although this includes children that are learning English as their first language and people that are learning English when they are already fluent in another, the primary focus will be on users that suffer from aphasia. The user must be able to hear the text-to-speech to use the app.

2.3 Data Availability

The Aphasia App's functionality regarding the eye tracking will be based on deep learning technology. Due to this, the accuracy of the eye tracking will be strongly reliant on the amount of data that can be collected to train the algorithm used in the app. The GazeCapture dataset is certainly voluminous, but it must be noted that more accurate performance could possibility be attained with an even larger data set. It should also be noted that the GazeCapture dataset was intended for usage on iOS devices.

3.0 Functional Requirements

3.1 Eye Tracking

The Aphasia App will track the user's eye movements as they look at the text on the screen. A facial mesh along with facial orientation will be used to collect data in real-time and processed using a deep CNN and support vector regressor to accurately track where the user's focus is.

3.2 Text-to-Speech

The Aphasia App will use Android's native text-to-speech functionality to read the words that are focused by the user, which will be signaled by the eye tracking portion of the app. The user will also be allowed to repeat certain words to being read aloud by redirecting their focus back to that word.

3.2.1 Music Intonation

The Aphasia App will be able to use music intonation when the text. The user will be able to toggle this to be either on or off.

3.3 Entering Text

There will be a textbox that serves as the primary location for the user to enter text to be read aloud. The user will be able to add and delete text, as well as have copy, cut, paste, and select functionalities.

3.4 Page Navigation

The user will be able to navigate between a home page, the main page which will host the main functionality of the app, and a settings page.

4.0 Non-Functional Requirements

4.1 Capacity

The target device will need to be capable in terms of data processing considering the large amount of data that will be flowing through the eye tracking portion of the app. A reduced frame rate may be used to achieve the real-time requirement. There would also need to be extra storage beyond that of just the app to store the frames taken from the camera to be processed as well as text that is to be read aloud, although this amount will be insignificant.

4.2 Usability

The user interface will be fast and straightforward allowing the user to navigate with ease. The app will require the individual to be in an area of sufficient light without any obstructions between the user and the camera for the eye tracking to work properly. The volume will need to be at a sufficient level compared to any background noise the user may have around them when they use the app. This will be handled externally by the Android OS.

5.0 Risks

4.1 Security Risk

The Aphasia App will not be using facial data for anything beyond determining the focus of the user's eyes, but it cannot be ignored that facial data can be used for other purposes such as facial recognition. Considering this, the Aphasia app will only use facial data for a short amount of time and will delete the data when it is no longer being used. The pictures the facial data is taken from will not be used for any app functionality, and therefore, not be stored in any way. Only the face mesh data points and facial orientation will be used.