Boston University
Electrical & Computer Engineering
EC463 Senior Design Project

First Semester Report

Speech Interactive Therapy Application

Submitted to

Andrey Vyshedskiy
(617) 817-1916
vysha@bu.edu

Team 13 SITA

Team Members
Jenni Norell: janorell@bu.edu
Diana DeLeon: ddel@bu.edu
Zhe Deng: thezhe@bu.edu
Swathi Jaisankar: swathij@bu.edu
Cindhuja Sramasamy: cindhu25@bu.edu

Submitted: 10 December 2019

Table of Contents

Executive Summary

Speech Interactive Therapy Application (SITA)
Team 13 – Speech Therapy

Our team is determined to make a speech interactive therapy application to help young, nonverbal children with autism spectrum disorder (ASD). The game begins with videos of a speech pathologist pronouncing phonemes. As the child's skills progress, the game will play more complex and advanced levels, up to sentence-level exercises. Parents or guardians will also have the choice to record their own video levels. For every video, the child will receive 2D visual rewards based on an accuracy score. A mix of envelope matching and our client's own child vocal isolation algorithm will give us an accuracy function tailored for young, nonverbal learners. All of SITA's components will be built in Unity 3D and deployed to iOS and Android. Our goal is to provide a very simple, yet dynamic therapy environment that gradually builds necessary speech skills.

1.0 Introduction

Our Project is on a Speech Therapy App. Our team is developing a game that will assist in the speech therapy of kids with autism spectrum disorder. We are developing this application for non-verbal children between the ages of 2 to 5. It is a dynamic speech therapy game for language cognition and speech training.

Our App is called the SITA - Speech Interactive Therapy Application. It is mainly designed to provide an early therapy action to the ASD kids in the absence of a therapist in a game-like environment.

Design Plan of SITA:

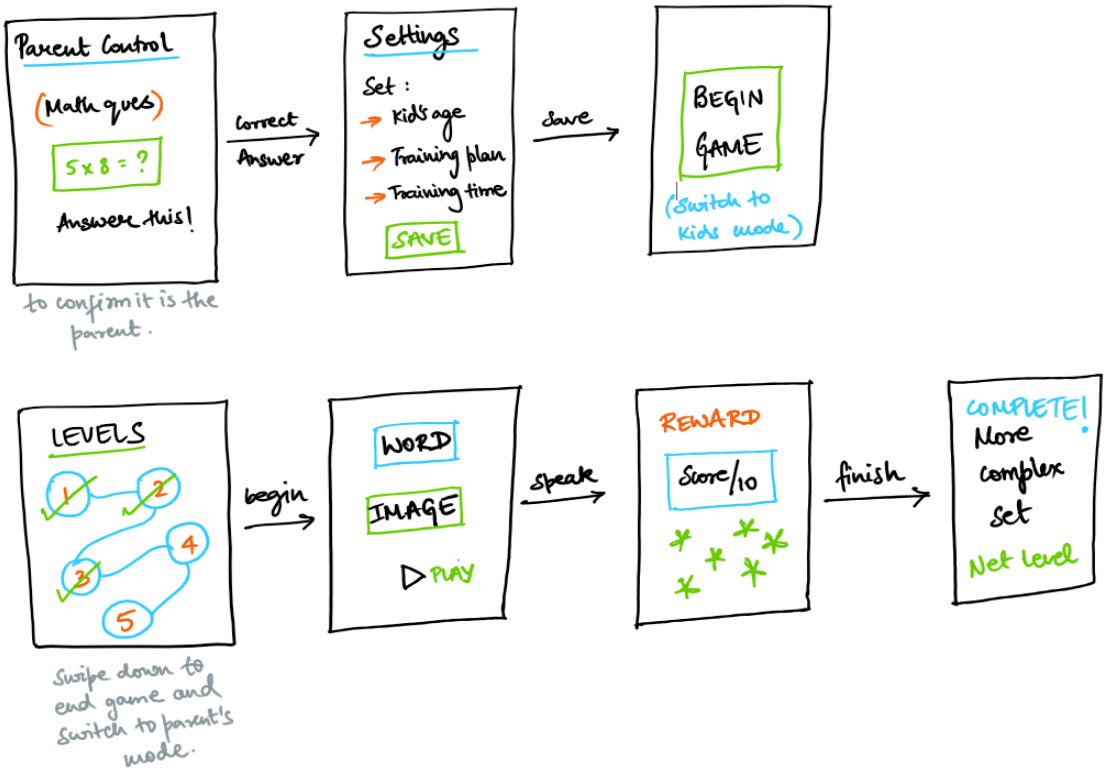
- ❑ Developed in Unity 3D as proposed by the client.
- ❑ User Interface made colorful and interactive in an appealing way to the kids.
- ❑ Created with only 2D effects to avoid longer loading time and lagging when in kids mode to keep them engaged.
- ❑ Access control provided to parents and kids only for privacy reasons and everything is run offline.
- ❑ Database consists of a speech pathologist recorded model word videos organised based on phonemes.
- ❑ Customization for kids based on their improvement in speech in terms of difficulty.

Goals of SITA:

- To provide speech therapy access to all ASD kids at minimum costs.
- Make the speech training more fun filled in a gaming atmosphere.
- Ensure the training session is accessible to kids at their convenient times and places.
- Maintains privacy in terms of kids voice, training and progress.
- Helps keep a regular check on their improvement and customize the training sessions.
- Developed on Unity 3D which can be made both Android and iOS compatible.

Current flow of SITA:

- ➔ Opens with the lock screen in which the parent answers a math question to ensure it is not the child logging in.
- ➔ Goes to the levels page where the difficulty level for the game can be chosen by the parent for their kids.
- ➔ Control can be transferred to the settings page where the parent can track the progress of their child and perform additional settings.
- ➔ After the level is chosen, the app will switch to the kids mode and a video of a speech pathologist pronouncing a word will be played to the child and the child will be asked to repeat it.
- ➔ Once the child repeats the word, the kid will be rewarded with some visual 2D effects and some music later on.


Parent Control
(Math ques)
5 x 8 = ?
Answer this!
to confirm it is the parent.
correct Answer
Settings
Set :
Kid's age
Training plan
Training time
SAVE
save
BEGIN GAME
(Switch to kids mode)
LEVELS
1
2
3
4
5
Swipe down to end game and switch to parent's mode.
begin
WORD
IMAGE
PLAY
speak
REWARD
Score/10
finish
COMPLETE!
More complex set
Net level

2.0 Concept Development

Our client, Andrey Vyshedskiy, is the founder and CEO of ImagiRation LCC. His company currently focuses on Mental Imagery Therapy for Autism which is an early-intervention application for children with Autism Spectrum Disorder (ASD). The application is filled with bright and interactive puzzles designed to help children learn how to mentally process an image or object. Therefore, the application is designed to improve a child's visual skills and language skills through image processing.

Our customer now wants an application that will help a nonverbal child with ASD improve their vocalization skills. Currently, there aren't any speech therapy applications available for nonverbal children with ASD. Tectus's line of applications is tailored to adults with ASD who have some capability of speech and vocalization skills. These applications are also bland and not engaging. The voice they use to pronounce the words is robotic and most likely autogenerated, so it sounds like Siri or Alexa talking.

The speech interactive therapy application (SITA) we are developing is going to be based on MITA's user interface, since the children will learn and focus more if the application is set up as a bright and colorful game. Our goal is to improve these children's language cognition and vocalization skills. The game will have levels that will increase in difficulty as the child progresses. In order to keep the child engaged and encouraged, we will focus on keeping a positive environment by giving the child rewards through special effects. The amount of rewards given will depend on how accurate the child said the phenom or word the pathologist said. All of our videos will be recorded by a speech pathologist who will accurately pronounce the phenom or word.

Engineering Requirements of the Customer:

- ➢ App to be developed on Unity 3D
- ➢ App has to run offline no internet or web server usage.
- ➢ Colorful and interactive User Interface
- ➢ Videos have to be organized based on phonemes.
- ➢ Should run on two modes parents mode and kids mode.
- ➢ Progress tracking must be saved to database.
- ➢ Voice recognition needs to be integrated in Unity.
- ➢ Parents must be able to record their own video and sound.
- ➢ Easy accessibility of the app from anywhere.

Reasons and Conceptual Approach:

Our application is being developed on Unity 3D, since it's iOS and Android compatible. Unity3D is also the platform that MITA was developed on, and it allows a wider range of people to have access to our application. It was decided that SITA will run without wifi or bluetooth This allows the application to run offline and on multiple devices at once. This maintains the privacy of the kids, and it prevents lagging and excessive loading time that an online game dependent on wifi would have. The rewards we are implementing include music and special effects that will be used to visually aid and motivate the kids. We will be creating a phonemes scene, and linking it to the levels scenes. We will also be tracking the child's speech improvement progress by logging it into a database under the settings scene.

In order to keep the child from messing around with the settings or from leaving a game when he or she didn't mean to, we decided to verify it's the parent by using a math question instead of a password. This makes the login easier, and it avoids the necessity of remembering a passcode.

Our approach for recognizing the child's voice has changed recently. We initially used the Windows in-built voice recognition algorithm, but now we've decided to change it to a unified voice recognition algorithm. This decision was made because this approach works on any platform while the Windows in-built voice recognition doesn't. We also discussed how we are going to record and store parents customized videos and sounds. We plan on having a camera within the application where the parent can record and then upload the video all within the application. We also decided that this approach is best because it allows the children to practice their speech and vocalization skills in any language.

3.0 System Description

What we will provide as a solution is an application built in Unity3D. This application has various stages, or scenes. Each scene has its own functionality and connects to at least one other scene. Below is a flowchart of how these scenes interact with each other. The interrupt lock scene is the first page seen when opening the application. This is a simple math problem to verify that an adult is trying to access the application rather than a child, and it has two states. Its first state is its initial state, in which there are unlimited tries to access the application. The second state is when we enter the interrupt lock after exiting the video scene. From here, the user has three attempts to solve a math problem, and if they get it incorrect each time they are locked on the video scene for 60 seconds.

The next scene is the LevelMenu scene. This scene is, after getting passed the interrupt lock, the main menu. There are various buttons that connect to the video level scene and one button that allows us to navigate into the settings page (the parent setup scene). The Parent Setup scene currently has four panels, each of which will contain relevant information once the final prototype is built and a button leading back to the level scene. The video scene currently connects to a video and when clicked, plays it. After the video is played, there is a pause in which the application is running a voice recognition algorithm. If the correct words are spoken, there is a reward element that appears on the scene. At any point in this scene, the user can press a button that will lead to the level menu scene (after getting through the interrupt lock).

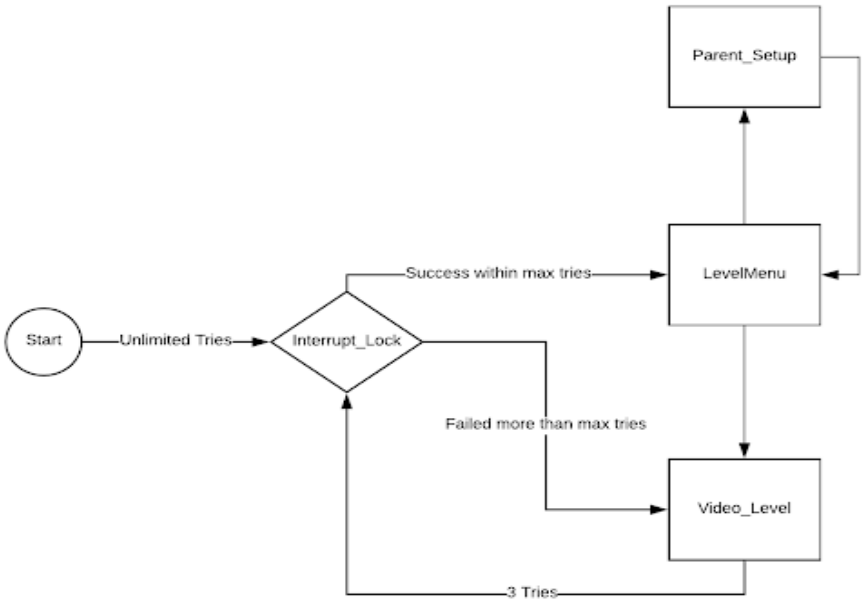


Figure 1. Block Diagram of First Prototype

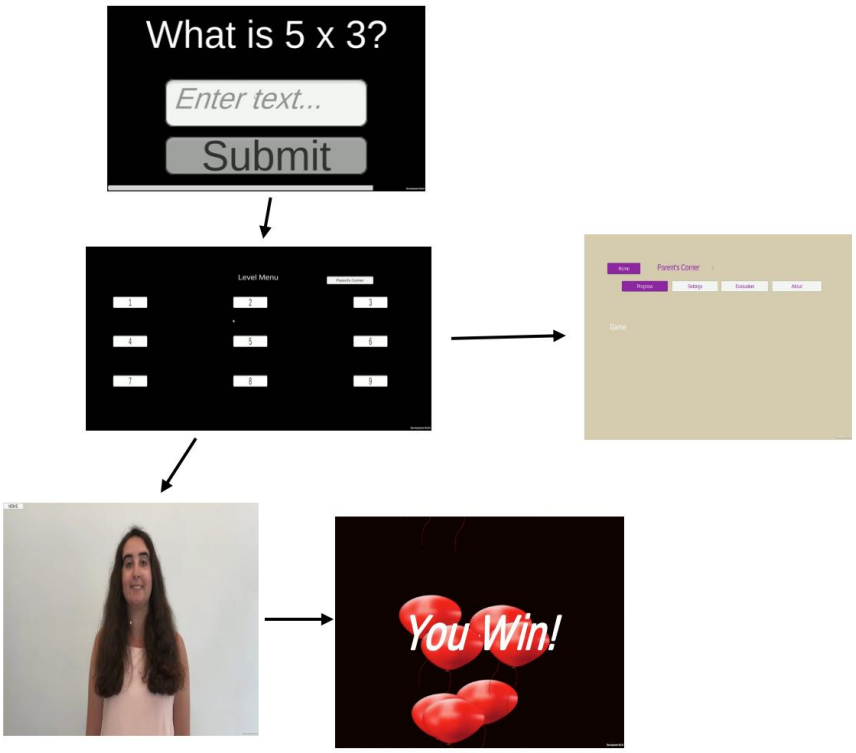


Figure 3. Screenshots from Current Prototype

This next section details what the final prototype will look like. The flowchart itself does not look that different, with only one added scene, but certain aspects of the currently existing scenes will be changed. The Parent Setup scene will allow us to select which levels we want to display. For example, if we choose “easy”, then only the “easy” levels will be shown. It will also be keeping track of the progress of the levels. There will be an added scene that connects to the Parent Setup scene. In this scene the user will have the ability to record their own videos that will then be stored in a database. There will be an option to display these homemade videos on the levels menu.

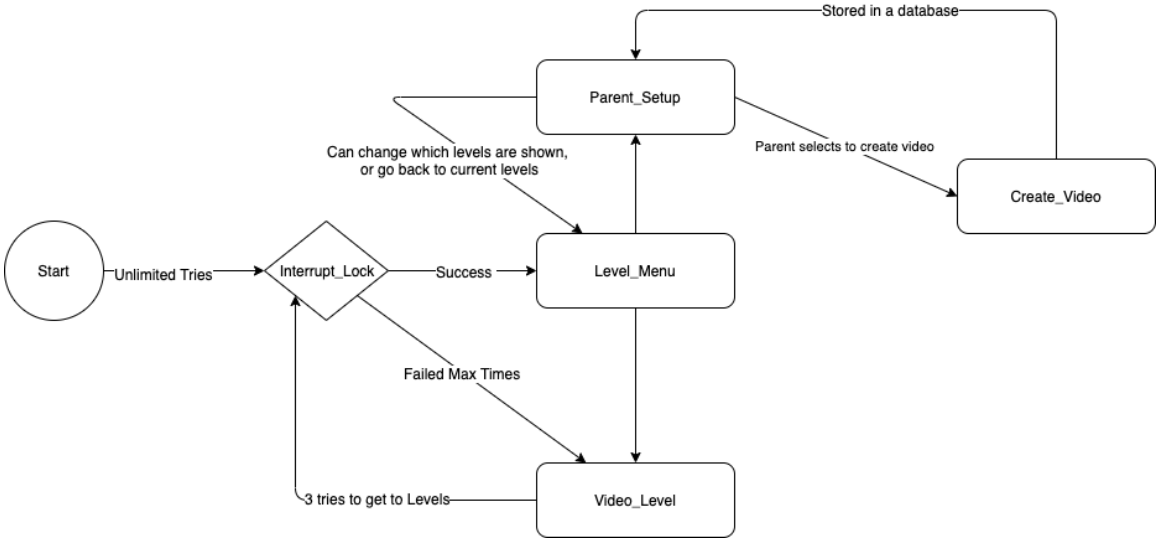


Figure 2. Block Diagram of Final Prototype

4.0 First Semester Progress

Key Testing Results:

1. Unity sutters more than expected (avoid expensive memory allocation and scene switches)
2. voice recognition depends on environment noise (make sure to include client's vocal isolation algorithm before any processing of mic input)
3. Most voice recognition libraries for C# and Unity are deprecated or severely limited
4. App needs to look more professional before release

Zhe Deng:

- Scene Manager
  - Scripted and made a prefab for an in-editor Scene Manager object
  - Uses asynchronous operation to load scenes on a separate thread
- Lock Screen (Interrupt_Lock)
  - Made logic to create math questions and test answers
  - Connected script to Scene Manager for tamper-proofing (locking video scene after 3 failed unlock attempts)
  - Has a timer and a submit button delay for each new question
- Effects Scripting
  - Trigger ParticleSystems via code and sync with AudioClips
  - Will take in accuracy scores from 0 to 1 to determine specific effect parameters
- Voice Algorithm
  - Experimented with native Unity Voice Recognition and Spectrum Analysis (AudioSource.GetSpectrumData)
  - Modified microphone script to read constant buffers
  - Wrote custom FFT routine in C# which gets spectrum data directly from microphone buffers

Cindhuja SRamasamy:

- Rewards Scene Effects:
  - ★ Effect objects created in adobe photoshop
    - Three different png images of effects with balloons, stars and smileys created.
    - Made colourful on transparent background to be integrated in the unity effects.
  - ★ Effect scenes on unity
    - Integrated the textures from adobe photoshop and created the materials for the effects to be triggered.
    - Three scenes for different accuracy ranges initially created with the materials drafted in the assets.
    - Particle effects were triggered with the different textures and materials to try how the emission generated 2D effects.
  - ★ Texts for the Rewards

    - Text Mesh Pro was used for texts in the scenes in between the effects along with the particle system effects.

Diana DeLeon:

- Video Scene
    - Created a Canvas with a RawImage component
    - Adjusted Texture parameter to Main Camera
- 'videoController' Script
    - Parameters included: Image, playIcon, videoToPlay, audioSource and videoSource
    - 'IEnumerator playVideo()' function retrieves the video, buffers the video, then plays the video
    - Loop: Plays the video once, then waits 10 sec for the user to repeat after the pathologist. This is done 3 times if no response is given. If a response is given, the loop is stopped.
- Integration
    - Connecting Level Menu to individual videos based on difficulty
    - Creating a new script
        - Will contain an array of videos
        - Videos chosen based on a number system

Jennifer Norell:

- Level Scene
    - Created a canvas that was overlaid with various buttons
- Level Buttons
    - Created 3 rows of buttons (3 buttons total) that can navigate to the Video Scene
- Parent's Corner Button
    - Created a TextMeshPro button that can navigate to the Parent's corner
- Integration
    - Connecting level menu to specific levels dependent on which level is selected. (Working with Diana)
    - Connecting level menu to parent's corner to pass the highest level the child has achieved in order to track progress (Working with Swathi)

Swathi Jaisankar:

- Parents' Corner
    - My main responsibilities have been in figuring out what exactly to incorporate into the Parents' Corner to make it best suitable for the child and the parent.
- Parent_Setup Scene
    - Made it accessible after the math question is answered correctly
    - Created a canvas, scene manager, and event system
    - Canvas includes multiple buttons, panels, and text boxes:
        - Home Button that connects to the Home Page
        - Progress Button that gets data from the Levels Menu and has an evident bar to track progress (need to work on it further with Jenni so I can access the data generated from the levels she created)
        - Settings Button for parents to customize the app in the best suitable way for their child (Client will give us input on specific phoneme options)

- Evaluation and About Buttons to give the user information about the application

5.0 Technical Plan

Task 1: Video Recorder
12/16-1/13
The video recordings should be inputted for each level accordingly while taking into account the Parent's customized settings / preferences from Parent's Corner. After the video plays and child repeats it, the rewards should show up. Lead: Diana, Cindhu

Task 2: Make Level Menu Phoneme-Oriented
12/23-12/30
Going off of what our client suggested, the levels will have to be sorted by phonemes rather than by numbers. The levels will increase in difficulty by phonemes. Lead: Jenni

Task 3: Sync Level Menu with Videos/Database
12/16-1/20
The level menu will have to be synced with all the videos which will have to be called from the database. The scenes will all have to flow into one another at real time and not be delayed especially as that could cause the child to get bored and lose focus on the game. Lead: Diana

Task 4: Voice Envelope Processing
12/16-1/6
Approximately 10 seconds of samples from the microphone input and all of the samples from the therapist/parent/guardian pronunciations will be converted into down-sampled envelopes using a moving RMS block. Then, both envelopes will be normalized and cross correlated. Local max values of the cross correlation will serve as the accuracy scores. This algorithm is planned to work in real time to provide dynamic rewards feedback. Lead: Zhe

Task 5: Effects Assets and Sync with Sound
12/23-1/13
Particles will be reworked so that they leverage more components from the Unity particle system (i.e. more control on particle diffusion and scaling as well as better optimization). Audio assets will sync with triggers and then performance testing will be done to determine an appropriate number of triggers which does not negatively affect frame rate. Lead: Cindhu

Task 6: Parents Corner Sync Progress
12/23-1/20
For the Parents Corner all 4 panels must be completed. Progress, Settings, Evaluation and About. The Progress Panel will showcase exactly how far the child is getting by being connected to the Levels Menu so parents can easily keep track of their child's progress. The Settings Page is for parents to be able to set and customize the application in the best way that suits their child. This includes being able to add in specific words or phonemes for their child to focus more on. The Evaluation Page includes a form to help SITA with its ongoing research to develop the most effective interventions for children with Autism. Finally the About Page gives parents detailed

insights into the purpose of the app and gives them step by step on how to navigate through it. All these aspects must be fully functional and integrated. Lead: Swathi

Task 7: Optimize and Test for Mobile
1/20-2/10
Our existing project will have to be tested to be the perfect fit for a mobile device. It should be working on both iOS and Android platforms so we will have to optimize it accordingly for both. Lead: Swathi, Jenni

Task 8: Integrate Client Voice Algorithm/ Vowel Detection
1/20-2/17
The client will give us a working voice algorithm that we can integrate into our application in order for it to detect the child's voice as close as possible. This way an accurate rewards system can be figured out based on how closely the child matches the speech pathologist. Lead: Zhe

Task 9: Test with ASD/nonverbal Children
3/16-3/30
We will be testing our application on actual children with ASD / nonverbal kids in order to test if it actually works. We will be able to measure the success of our application by keeping track of the progress that the child is making from the start of the use of the app to the end. Lead: Swathi

Task 11: Final Report
4/20-4/27
Finally, we will put together all of combined research, data, and findings onto our Final Report. This will be the conclusion of our project for Senior Design.

6.0 Budget Estimate

Our current budget estimate is zero dollars. Our Speech Interactive Therapy Application is a software development based project. We do not have any hardware components, and we do not have to purchase any software, platforms or tools in order to complete our project. All of the tools we are using to develop our application are free.

7.0 Attachments

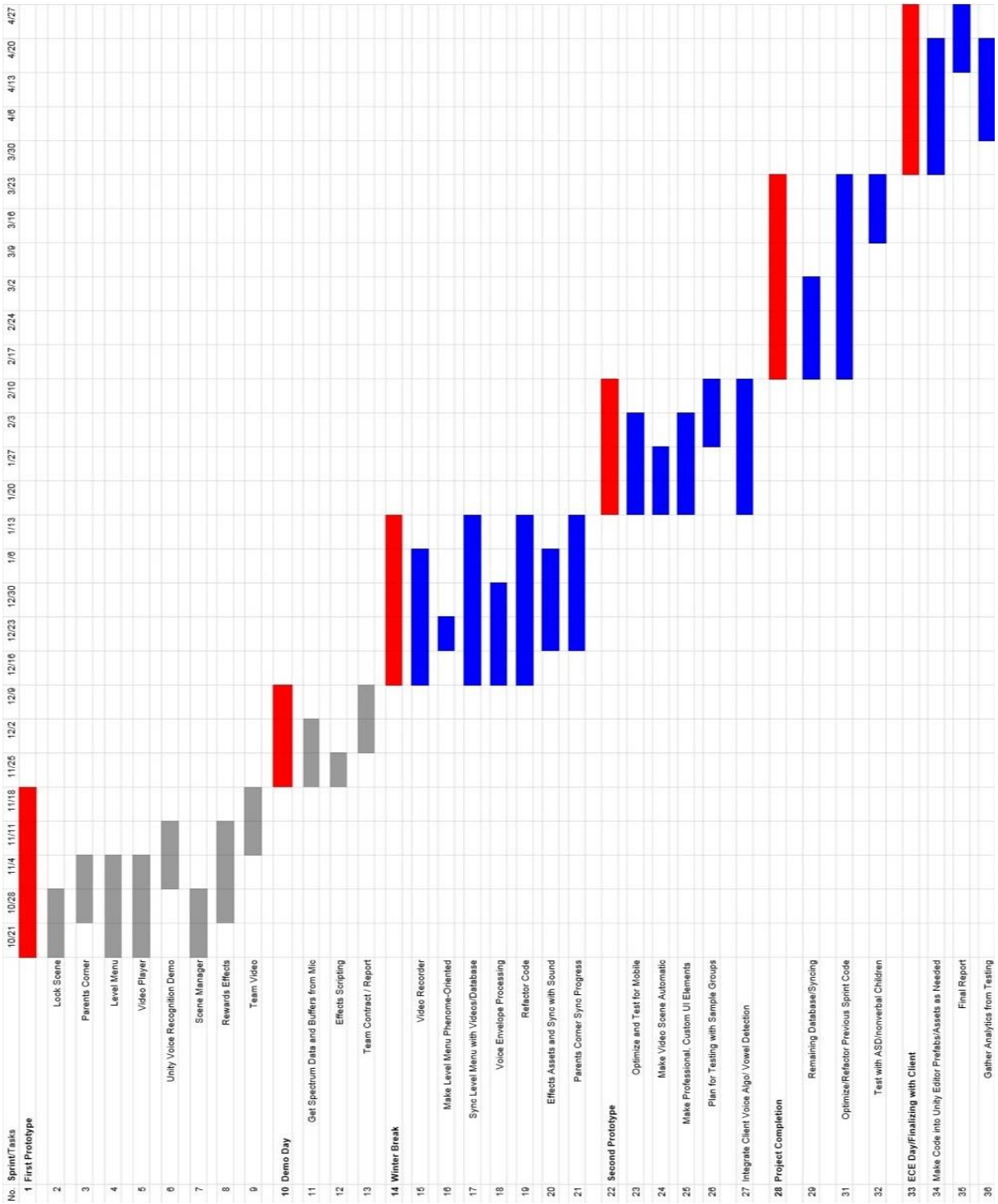
7.1 Appendix 1 – Engineering Requirements

Team # 13___ Team Name: Speech Therapy _____

Project Name: Speech Interactive Therapy Application (SITA)_____

Requirement	Details
Target Platforms	iOS, Android
Offline	Must not require internet and store parent recorded videos locally
Tamper-proofing	Parent settings require simple math questions to access; video scene locks after 3 unlock tries
FPS	App must run consistently at a minimum of 60 frames per second, without noticeable stuttering
Envelope Detection	The envelope detection algorithm must be able to take at least 10 seconds of microphone input and convert it into a down-sampled envelope for cross correlation with a target envelope. Both envelopes should be in RMS and normalized.
Aspect Ratio	Must support common mobile aspect ratios with minimal rescaling
Touch Compatible	App must be compatible with mobile keyboards and support gestures
Spectrum Analysis	Must be able to get microphone data at a sampling rate of 16kHz and buffer sizes of at least 4096 (i.e. FFT resolution is at least 4Hz).

7.2 Appendix 2 – Gantt Chart



7.3 Appendix 3 – Other Appendices

References

"MITA News." *ImagiRation*, imagiration.com/autism/mita-news. Accessed 4 Oct. 2019.

"The Best Autism Apps of the Year." *healthline*, 2017, www.healthline.com/health/autism/top-iphone-android-apps. Accessed 4 Oct. 2019.

"LetMeTalk: Free AAC Talker." www.letmetalk.info. Accessed 4 Oct. 2019.

Helter, Heather. "8 Reasons Why Tactus Therapy Apps Work Well for Autism." *Tactus Therapy*, 1 April 2015, tactustherapy.com/apps-for-autism. Accessed 4 Oct. 2019.

"Lets See Why You Need Comprehension Therapy." *Tactus Therapy*, tactustherapy.com/app/comprehension. Accessed 4 Oct. 2019.

Hagen, Mark. "Speech Sounds Visualized - Helping People Understand and Be Understood."

Closing The Gap, 23 April 2019. www.closingthegap.com/speech-sounds-visualized-helping-people-understand-and-be-understood. Accessed 4 Oct. 2019.

"What Are the Symptoms of Nonverbal Autism?" *Healthline*, 2019, What are the symptoms of nonverbal autism?