Memo – Second Deliverable Test Plan

**To:** Professor Pisano, Professor Kia

**From:** The Karty Boyz

**Team:** 15

**Date:** 2/10/2014

**Subject:** Second Deliverable Testing Plan

# **1. Overview**

1.1

The second deliverable test aims to demonstrate the functionality of the preliminary steps in processing external gameplay audio. Specifically, the test will involve processing a .wav file containing a known number of utterances and show that we can identify the starting and ending boundaries of these utterances accurately.

1.2

This test is important because its success or failure will offer evidence as to whether or not we will be able to deliver our planned “Tag It” feature. Without the ability to break up an audio stream into discrete utterances, it will be impossible to apply the planned hidden markov model phrase detectors to the audio. Once the audio has been broken into its constituent utterances we are confident that we will be able to accurately classify the type of utterance as either “Tag it” or “Watch this.”

# **2. Equipment and Setup**

2.1

The following equipment will be used in the conducting of this test

|  |  |
| --- | --- |
| Microphone | For recording test speech |
| Computer | For processing speech output |
| USB Audio Interface | For connecting mic to computer |

2.2

The microphone will be connected to the computer via the USB audio interface. A speech sample will be recorded before the test occurs and used during the actual test. During the test the audio recording method will be demonstrated, and the sample audio played for comparison. The processing algorithms will then be applied to the pre-recorded audio.

# **3. Measurements Taken**

3.1

The primary measurement that we are interested in is the total number of utterances detected. This will correspond to the number of discrete speech items in the audio file.

3.2

In addition, we will be interested in knowing where the utterances occurred in the audio stream. This includes the time at which the utterance began and the time at which it completed.

3.3

The recorded data will also be presented to a human who will count the number of utterances and record at what time each utterance occurred.

# 4. Criteria for Success

4.1

The number of utterances must be within 10% of the counted number of utterances. This is to account for the fact that upon listening “watchthis” and “watch. this” may be ambiguities between what the processing detects and what the human ear detects.

4.2

A human listening to the audio stream must be able to identify the utterance at some time in between the detected start and end time or up to ⅓ of a second outside of those detections. This error bound is to account for human reaction time.

4.3

The test will be considered a success if the collected data is within the parameters specified above.

**5. Conclusion**

5.1

Once this test has finished we will have demonstrated the feasibility of implementing the “Tag it” feature. Utterances are one of the more challenging aspects of this feature and a satisfactory test will allow us to be confident that we will successfully complete this feature.

5.2

Once utterances are correctly detected we can move on to classifying the utterances using hidden markov models.