

# Group: Project 10

## Project specification on DT2112 - Speech Technology

Erifili Ichtiaroglou	Razan Jaberibraheem	Panagiotes Mousikidis	George Zervakis
erifili@kth.se	email@kth.se	email@kth.se	zervakis@kth.se

February 23, 2018

## Shadow speaking and chanting

**Supervisor: Jens Edlund**

### Description

Shadow speaking is a technique where a speaker attempts to trace the speech of another as closely as possible. It can be used for a wide variety of purposes, ranging from speech and cognitive research through data collection to entertainment. Current technology handles sound in ways that allows us to quickly implement systems with shadow speaking that would previously have taken great effort. The task here is to set up a pure HTML5 environment for shadow speaking and to run one or several experiments in it for proof of concept.

The scope of this project is to build an HTML5 environment for shadow speaking and observe differences in the reaction time (time between hearing a word and pronouncing it). To that end we will make use of javascript libraries available for HTML5 audio and we will run a number of different experiments, where the participants will repeat certain utterances immediately after hearing them. The experiment results will provide insight on which are the most difficult utterances to be pronounced, or how easily a human can follow a natural or a synthetic stimulus.

## Division of work

- Literature study, understanding the theoretical aspects of the topic and examine methodically implementation tutorial/guides to begin with the set up
- Implementation on HTML5
- Experimenting and inference
- Final assembly

Of course, all members will spent time in all parts, but the ones with e.g. better programming skills will help the others on the critical points and vice versa. Having that in mind, we propose the following timetable:

## Time - Line

### Feb 22 - 8 Mar

- Day 1-2: **Literature Study**

[HTML5 WebAudio API W3C specification](#) - Erifili

[WebAudio API overview with examples](#) - Razan

[Old \(2011\) introduction at HTML Rocks](#) - Panagiotes

[An exemplified intro at CSS-tricks](#) - George

Indicative Literature: Lee, Carson. "Speech Shadowing Support System in Language Learning." (2017) - The whole group

- Day 3-7: **Implementation**

### **Task 1: Create a webpage with a list of predefined audio files**

1. Configure Firebase to be used as a host for our website.
2. Configure the Firebase Storage and connect it with our project.
3. Store predefined audio files.
4. Create simple webpage to list the audio files on the webpage by retrieving the files from the firebase storage.
5. Add play button, where the user can play a selected audio file from the list shown on the webpage.
6. Responsibles: Razan Jaberibraheem & Panagiotes Mousikidis

**Task 2: Simple webpage to record the audio voice**

1. Add record button to record user voice.
2. On click the Recording WebAudio starts recording user voice.
3. Store the recorder audio to the Firebase Storage.
4. Responsible: Erifili Ichtiaroglou

**Task 3: Analyse the user reordered voice and the predefined recordings**

1. Use the WebAudio API to analyse the user recorders voice.
2. Measure the similarity degree berate the predefined recordings and the user recorded voice.
3. Responsible: George Zervakis

- Day 8: **Presentation**

- Day 8-12: **Experiments and inference**

We will evaluate our website using a controlled experiment. First we will conduct a pilot study with four users, and then we will run our experiment study. We will conduct the study with ten participants, our participants will be master students their age is between 23-30. The goal of the study is to evaluate the voice recording that we will get from the users with the predefined recordings that we have in our website. We would like to investigate the shadow of user speech on the predefined recordings. We will analyze the reaction time of users shadowing the audio they listen to.

- Day 13-14: **Report hand-in**

**Grade Target = A**