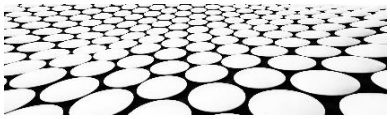
SEG3125 User Interface Design and Analysis



MODULE 10 - ACTIVITY

Speech Recognition in a Vocal Assistant

GOALS

One of the important steps of a voice assistant is the speech recognition step, which is speech to text transcription (ASR). We have seen that several factors can influence recognition rates, namely factors related to the speaker, the utterance, the discourse content, and the environment.

In this activity, you will test the voice recognition module of a personal assistant of your choice. If you don't have access to any device, you can test Google's Web Speech API.

We will use a set of test sentences, and try to test speech recognition by modifying various factors.

SUBMISSION DATE

- Tuesday, July 28, 11:30 pm

SUBMISSION METHOD

- In Brightspace, your Module 10 checklist contains a link to submit your experimentation report.

INSTRUCTIONS

Step 1 - Choose 10 sentences

There is a set of phrases, called Harvard Sentences, which are often used to perform recognition tests or simply tests for the quality of speech transmission in various communication devices.

Here are several sentence lists from Harvard Sentences. From these lists, choose 10 sentences (they can come from different lists).

These sentences are in English. I think the fact that you may not be a native English speaker will affect recognition rates and it will be interesting.

Here ... we must admit that we are introducing a negative bias in the test, because our devices do not expect these Harvard sentences! These are not usual commands… All the same, let's look at how the voice recognition component of our intelligent personal assistant will manage!

Step 2 - Choice of voice assistant

Decide on a device you want to test. It can be your phone, or a tablet, or if you don't have access to any device, you can test Google's Web Speech API. It's hard to use smart speakers for our test, because they don't show us the transcript of what is said, and that's what we want to assess here.

Step 3 - Choice of various conditions to compare

Here are some conditions you can test. You don't have to do all of them, **choose at least 5**. Or suggest other conditions.

1. Speaker (ask friends or family to help you)
    - Different age
    - Man / Woman
    - Different native language
2. Utterance
    - Isolated versus connected words
    - Whisper versus talk normally
3. Environment
    - Indoor (in the house) versus outdoor
    - With or without music in the room
    - With or without someone else talking in the background
    - Speak directly to your device, or stand with your back to the device
    - Speak near or far from the device

Step 4 - Perform your experiment

For each of your 10 sentences, test a comparison of factors. Write down what is recognized by the assistant.

Please note. We are not interested in the response generated which comes from the fact that the voice assistant does not know what to do with what you are saying. For example, if I test Siri with "A pot of tea helps to pass the evening", it will answer "I didn't get that" or "I don't know how to respond to that" (obviously !!) but I'm not interested ... For this lab, we're testing speech transcription and then I'm just going to look at how the assistant transcribed the speech. For this example, the transcript is "A part of tea helps to pass the evening". This shows that voice recognition recognized "a part" when I said "a pot".

You should create a small table like this with 10 comparative tests.

	Harvard Sentence	**Condition 1**	**Transcription from Siri**	**Condition 2**	**Transcription from Siri**
1	*The small pup gnawed a hole in the sock.*	John. Male. 40. Silent room. Utterance with connected words.	*Smart but not a hole in the sock*	John. Male. 40. Silent room. Utterance with isolated words.	*The small nod a hole in the sock*
2	*Press the pants and sew a button on the vest*	Lucie. Female. 13. Silent room. Normal voice.	*Press the pants and sewer button on the best*	Lucie. Female. 13. Silent room. Whispering.	*Press the pants at so I put it on the first*
3					
4	…				
	…				

Step 5 - Write a small report

In your report, put the information corresponding to each of steps 1 to 4.

- Which phrases did you choose? (at least 10)
- What device are you testing?
- What conditions are you testing?
- Include your result table.
- A small conclusion. What do you think of the recognition results? Are you satisfied? Disappointed? Are there factors that seem problematic for recognition?

☆☆☆

EVALUATION

- This activity is worth 1.5%.
- The assessment is based on effort, so any student who has made an effort to perform experimentations and vary conditions as mentioned in the requirements will get full marks. Your report must include everything required (step 5 above).