**CSC 745 Advanced Multimedia Programming**

**Exercise No. 56, September 25, 2018**

**Speech Recognition**

**Goal**

Learn how to write code to use speech recognizers

**Background**

There are a number of free speech recognizers available on the web. In this exercise you’ll learn to write code to connect to a recognizer and perform speech to text transcription.

**Procedure**

Do the speech recognition tutorial at <https://realpython.com/python-speech-recognition/>. The link to the audio file used in the tutorial seems to be broken; there is a directory on the class server (Data/Speech) holding a number of utterances recorded in different languages. The tutorial also asks you to use a file that has speech mixed with the sound of a jackhammer; the file *hammered\_speech.wav* is a similar file.

The speech recordings on the class server are from the Harvard Sentences collection – a standard dataset for testing speech coding, transmission, and recognition systems. The file *HarvardSentences.docx* has a transcription of the full set of sentences. The recordings are a subset of the full dataset.

Go through the tutorial. Try recognizing speech in some other languages. You’ll have to find the language code for other languages and, unfortunately, I couldn’t find transcriptions for all the non-English language sentences – so you should try languages you understand.

If you have time, try recognizing microphone input. You may find the quality of the microphones on our lab systems makes this a difficult task.