1 Lab 3

I thought about a few different phrases that seemed difficult to recognized, spoke them into the default ASR system from Microsoft Azure, and logged the recognized utterance and confidence value. The results are shown in the following table.

Spoken Phrase	Recognized Phrase	Confidence
Brassica Oleracea	Brassica ulirakia	0.21973889
Rutilus rutilus	Route to loose route Telus	0.3586412
Der fliegende Holländer	Their flag in the hole under	0.44981894
Väinö Linna	Venerlina	0.14243059
Humanisten	Humanist	0.34891802

None of the phrases were recognized correctly. I don't think this is due to my accent; it seems to be recognized correctly when I say more common phrases, and the proposed utterances from Azure speech recognision seem fairly close to the correct ones in a phonetic sense.

I then created a text file with the expected target strings, uploaded it to Azure, and trained a new model using that. This did not improve the results at all, which surprised me.

Finally, I recorded sound files for all the words, paired them with the expected output, and trained a third model. This model was evaluated as before, with the following results:

Spoken Phrase	Recognized Phrase	Confidence
Brassica Oleracea	Brassica olyroakia	0.29313663
Rutilus rutilus	Route to Lus Rutilus	0.24030894
Der fliegende Holländer	Their fleek in the Holelander	0.5087921
Väinö Linna	Venerlina	0.07821497
Humanisten	Humanist	0.34319595

The table shows no improvement! This was *very* surprising. Unfortunately, I could not find out what the problem was. One possible explanation is that my application connects to the wrong endpoint, but this is not the case according to the logs. Another possibility is that the training data was bad, but it seems reasonable enough as far as I can tell.