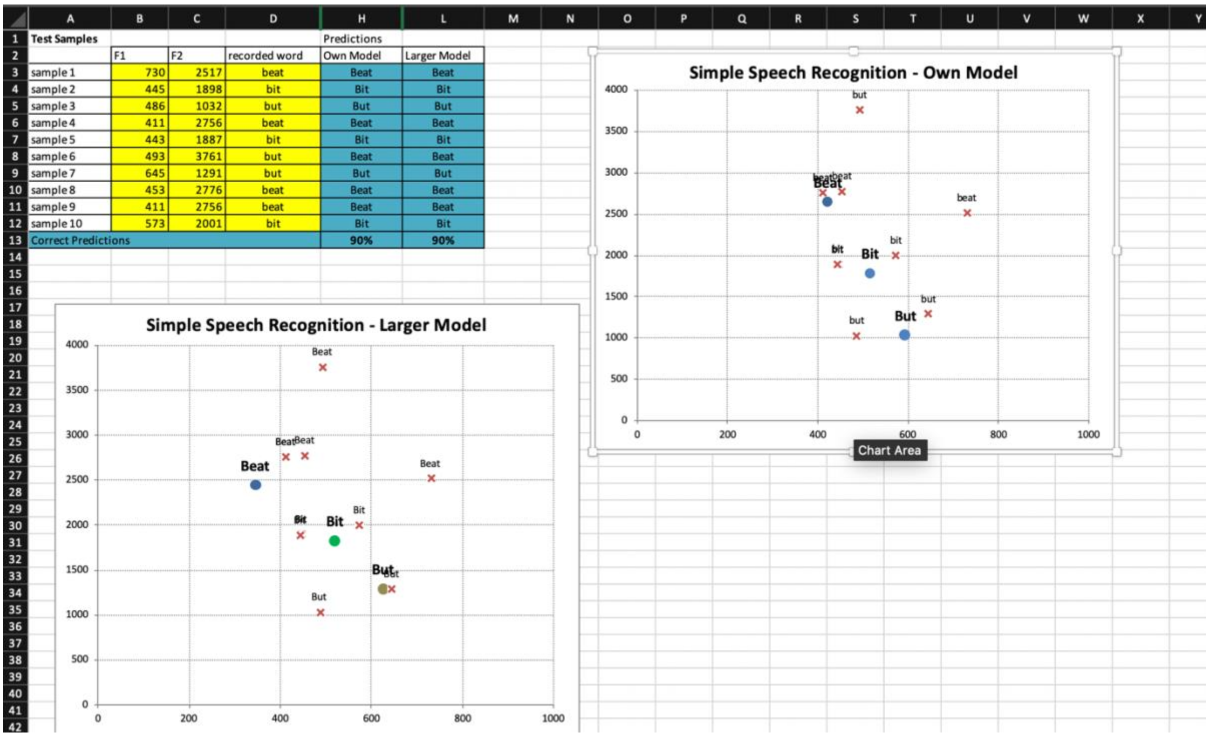
a) Describe your experience and your results (use screenshots where necessary). Which model performed better? Which words were most often confused?



Test Samples				Predictions	
	F1	F2	recorded word	Own Model	Larger Model
sample 1	730	2517	beat	Beat	Beat
sample 2	445	1898	bit	Bit	Bit
sample 3	486	1032	but	But	But
sample 4	411	2756	beat	Beat	Beat
sample 5	443	1887	bit	Bit	Bit
sample 6	493	3761	but	Beat	Beat
sample 7	645	1291	but	But	But
sample 8	453	2776	beat	Beat	Beat
sample 9	411	2756	beat	Beat	Beat
sample 10	573	2001	bit	Bit	Bit
Correct Predictions				90%	90%

Both models preformed equally good. This might be due to the fact that I only used male voices without a thick accent so they sounded similar to me.
When I used a Chinese person ,which had a thick accent, as a sample both models failed to successfully identify the word (sample 6).

b) In your opinion, when and why are larger training sets necessary? What are the drawbacks/benefits of using recordings from a group of people as opposed to several examples from the same person? Can you think of ways the model could make better use of the larger amount of data rather than simply averaging over the samples?

Larger training set are necessary when we want to open the model to the general public. they are necessary because different people speak differently ,for instance females tend to have a voice that has a higher pitch.

The drawback of having a group of people instead of a single person is that there will be people that have voice that is too different to the others and will manipulate the data in the wrong direction. But the advantage of having a large enough group of people is that the model can predict more accurately for more people while the other model can predict accurately for the same person.

An alternative approach to this problem is to have several averages, e.g. one for females ,one for Irish people with a thick accent, one for Africans with a thick accents , one for Iranians and etc.

c) Would this method be useful to distinguish between the words beat and beet? How would you solve this problem?

No, because they sound exactly the same and this this model cannot understand the context around the word. The only way that I can think of to solve this problem is to have a big enough neural network that can understand and predict the meaning of a whole sentence.