Initial Experiments

1. Description

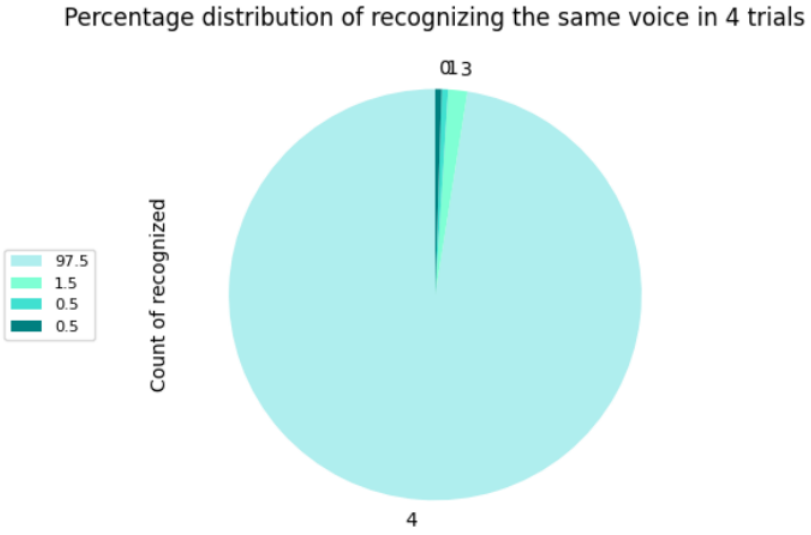
For our initial experiment, we decided to take 200 different speakers with 5 utterances for each one and conduct the research by drawing one utterance and comparing it with the other 4 in the speaker's set and comparing it with a random sample of 4 utterances.

The results are in the following format:

[((a,b), (c,d)), (), ()....]

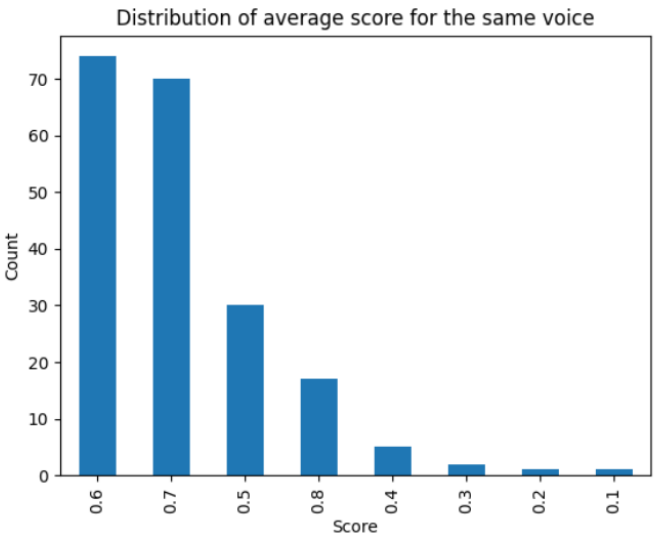
where:
a = how many times a given voice was recognised correctly
b = avg score for how many times a given voice was recognised correctly out of 4 attempts
c = how many times a voice was recognised as not a given one
d = avg score for how many times a voice was recognised as not a given one

2. English Dataset

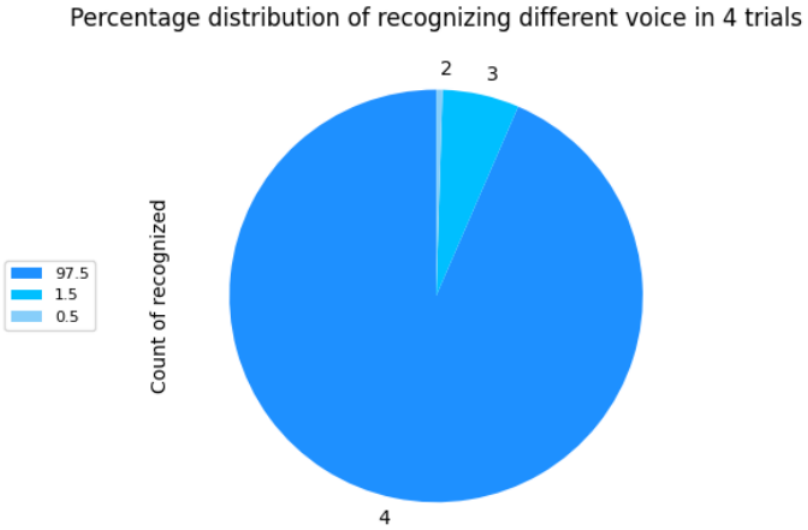


From the plot above, we can see that:

- 4/4 utterances were recognised correctly as the same speaker in 97,5%
- 3/4 in the 1,5% of cases
- 1/4 in the 0,5%
- and 0/4 in the 0,5%

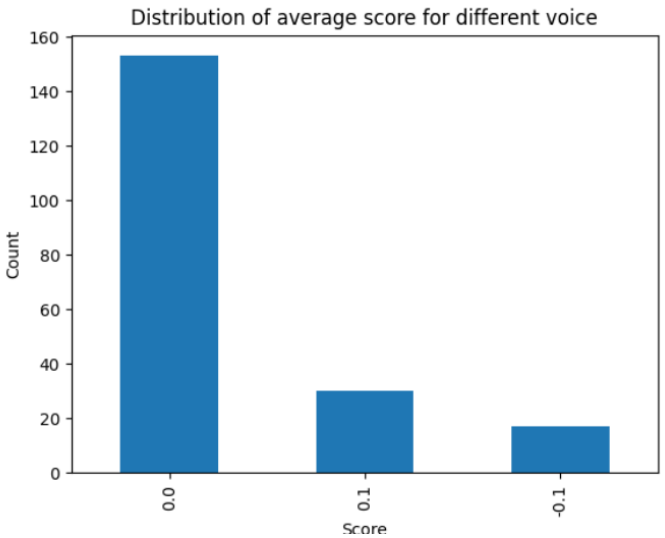


More than 70 voices were recognised correctly with an average score of 0.6.
Around 70 voices with an average score of 0.7.
Then, a huge decrease in the avg score of 0.5 - only around 30 voices.
With the highest average score of 0.8 - only ~15 voices were recognised correctly.
And for the rest of the voices, the average score is below 50%.

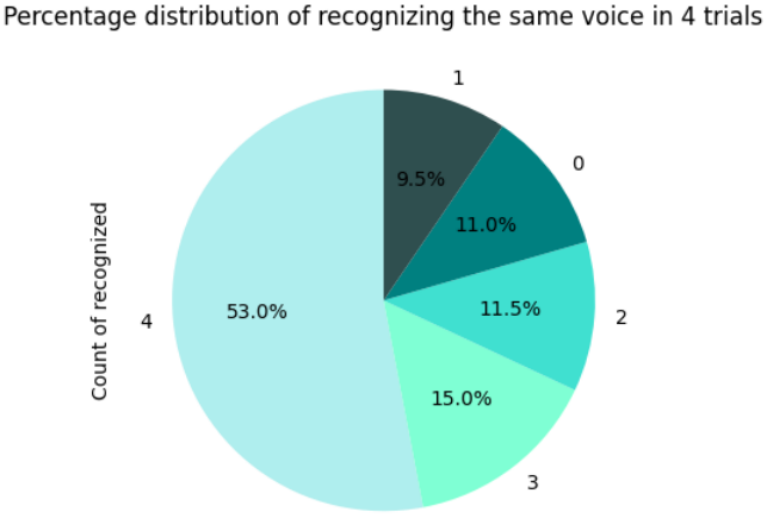


From the plot above, we can see that:

- 4/4 utterances were recognised as not belonging to the testing speaker in the 97,5%
- 3/4 in the 1,5% of cases
- 2/4 in the 0,5%

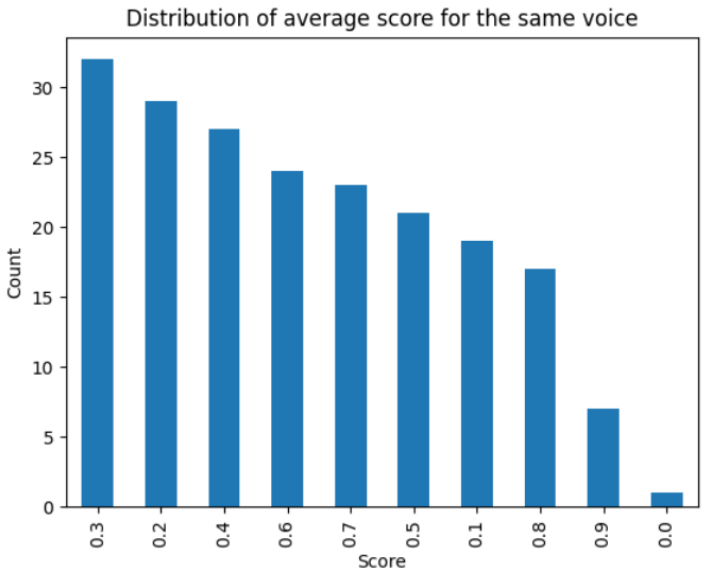


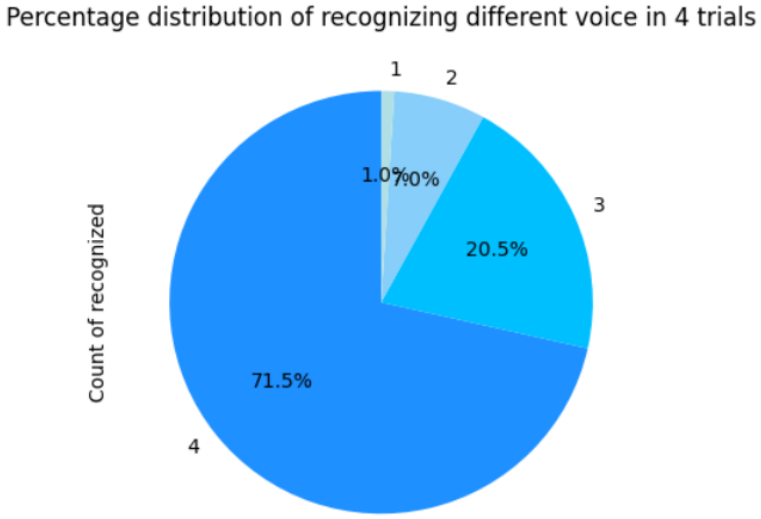
3. Chinese Dataset



From the plot above, we can see that:

- 4/4 utterances were recognised correctly as the same speaker in 53%
- 3/4 in 15% of cases
- 2/4 in 11,5%
- 1/4 in 9,5%
- and 0/4 in 11%





From the plot above, we can see that:

- 4/4 utterances were recognised as not belonging to the testing speaker in the 71,5%
- 3/4 in 20,5% of cases
- 2/4 in 1%
- 1/4 in 1%