Mike Roylance
Ling 572
Homework 9

q3-a)

N	Training accuracy	Test accuracy
1	0.45296296296296296	0.4166666666666667
5	0.6148148148148148	0.6366666666666667
10	0.6840740740740741	0.6966666666666667
20	0.752962962962963	0.73
50	0.8362962962962963	0.7633333333333333
100	0.8951851851851852	0.7933333333333333
150	0.9251851851851852	0.7966666666666666
200	0.9437037037037037	0.8
250	0.9622222222222222	0.7866666666666666

q3-b)

For training, as the number of transformations got larger, the accuracy increased. This makes sense, as the model file was built directly from the training instances.

For test, the “sweet spot” of transformations seems right around 200. Any more appears to be an “over-fitted” model. As can be seen, adding 50 more transformations actually had a more negative affect than remove 100 transformations. The accuracy for 100 transformations was 0.79333 where the accuracy for 250 transformations was 0.78666. So the recommended amount seems to be between 150 and 200, but no more.