

I lost track of the problem statement for this project as well but we were given training corpus' in French and English then told to find a training corpus for the third language ourselves. Then we were to build bigram and unigram models and predict the language of some test sentences based on the probabilities from our models. I did the unigram part and my group partner did the bigram.