- Consider the word York

  - fairly frequent word in Europarl corpus, occurs 477 times
  - as frequent as foods, indicates and providers
  → in unigram language model: a respectable probability

- However, it almost always directly follows New (473 times)

- Recall: unigram model only used, if the bigram model inconclusive

  - York unlikely second word in unseen bigram
  - in back-off unigram model, York should have low probability