Translation model: p(f|e)

We could learn translation probabilities for separate words:

