IBM Model 1

Lexical probabilities:

- Disregard the position of words in sentences.
- Estimated using Expectation-Maximization Loop.

See the slides by Philipp Koehn for:

- Formulas of both expectation and maximization step.
- The trick in expectation step, swapping sum and product by rearranging the sum.
- Pseudocode.

  Illustration: MT Talk #8 (`https://youtu.be/mqyMDLu5JPw`)