

Advantages of Explicit Morphology for LM

- LM over morphological tags generalizes better.
 - $p(\text{dvě kočkách}) < p(\text{dvě kočky})$...surely
But we would need to see all combinations of *pruhovaný* and *kočka*!
 \Rightarrow Better to ask if $p(\text{fem-nom fem-loc}) < p(\text{fem-nom fem-nom})$

which is trained on any feminine adj+noun.

- But still does not solve everything.
 - $p(\text{dvě zelené}) \geq p(\text{dva zelené})$... bad question anyway!
Not solved by asking if $p(\text{fem-nom fem-nom}) \geq p(\text{masc-nom masc-nom})$.
- Tagset size smaller than vocabulary.
 \Rightarrow can afford e.g. 7-grams:
 $p(\text{masc-nom fem-nom fem-nom}) < p(\text{fem-nom fem-nom fem-nom})$