Bigram language model

So that's what we get for n = 2:

$$p(\mathbf{w}) = p(w_1)p(w_2|w_1)\ldots p(w_k|w_{k-1})$$

Toy corpus:

This is the malt
That lay in the house that Jack built.

p(this is the house) = p(this) p(is| this) p(the| is) p(house| the)