

Toy corpus

*This is the house **that Jack** built.*

This is the malt

***That** lay in the house **that Jack** built.*

This is the rat,

***That** ate the malt*

***That** lay in the house **that Jack** built.*

This is the cat,

***That** killed the rat,*

***That** ate the malt*

***That** lay in the house **that Jack** built.*

bigrams

$$p(\text{Jack} \mid \text{that}) = \frac{c(\text{that Jack})}{c(\text{that...})} = \frac{4}{10}$$