

# 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.*

$$p(\text{house} \mid \text{this is the}) = \frac{c(\text{this is the house})}{c(\text{this is the ...})} = \frac{1}{4}$$