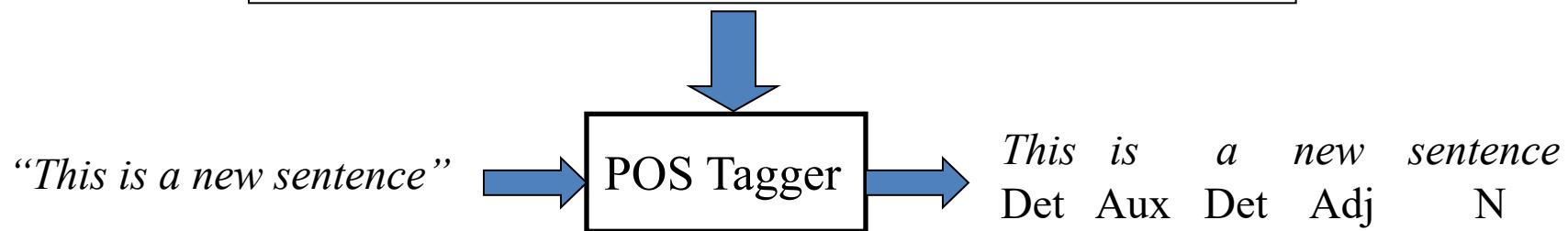


Using Language Models for POS Tagging

Training data (Annotated text)

<i>This</i>	<i>sentence</i>	<i>serves</i>	<i>as</i>	<i>an</i>	<i>example</i>	<i>of</i>
Det	N	V1	P	Det	N	P
<i>annotated</i>	<i>text...</i>					
V2	N					



Consider all possibilities,
and pick the one with
the highest probability

<i>This</i>	<i>is</i>	<i>a</i>	<i>new</i>	<i>sentence</i>
Det	Det	Det	Det	Det
....
Det	Aux	Adj	Adj	N

{ V2 V2 V2 V2 V2 }

$$p(w_1, \dots, w_k, t_1, \dots, t_k) = \prod_{i=1}^k p(w_i | t_i) p(t_i | t_{i-1})$$

Method 1: Independent assignment
Most common tag

Method 2: Partial dependency

$w_1 = "this"$, $w_2 = "is"$, ..., $t_1 = Det$, $t_2 = Det$, ...,