Language model: p(e)

$$p(\mathbf{e}) = p(e_1)p(e_2|e_1)\ldots p(e_k|e_1\ldots e_{k-1})$$

N-gram models or neural networks:

Have	a	good	day
<EOS>	Have	a	good