Redefining $p(e_1^I|f_1^J)$

What if we modelled $p(e_1^I|f_1^J)$ directly, word by word:

$$
\begin{aligned}
p(e_1^I|f_1^J) &= p(e_1, e_2, \dots e_I|f_1^J) \\
&= p(e_1|f_1^J) \cdot p(e_2|e_1, f_1^J) \cdot p(e_3|e_2, e_1, f_1^J) \dots \\
&= \prod_{i=1}^{I} p(e_i|e_1, \dots e_{i-1}, f_1^J)
\end{aligned}
\tag{13}
$$

...this is "just a cleverer language model:" $p(e_1^I) = \prod_{i=1}^{I} p(e_i|e_1, \dots e_{i-1})$
Main Benefit: All dependencies available.
But what technical device can learn this?