We would like

$$\sum_{n=1}^{\infty} \sum_{w_1 \dots w_n} p(w_1 \dots w_n) = \sum_{n=1}^{\infty} \sum_{w_1 \dots w_n} g(w_1 \dots w_n, n) \times 0.5^n = 1$$

where we can choose the function $g(w_1 \dots w_n, n)$.
Note that we have 3 words in the vocabulary $\mathcal{V}$, so there are 3^{n-1} sequences of the form $w_1 \dots w_n$. If we set

$$g(w_1 \dots w_n, n) = \frac{1}{3^{n-1}}$$

then

$$\begin{aligned}
\sum_{n=1}^{\infty} \sum_{w_1 \dots w_n} g(w_1 \dots w_n, n) \times 0.5^n &= \sum_{n=1}^{\infty} 0.5^n \underbrace{\sum_{w_1 \dots w_n} g(w_1 \dots w_n, n)}_{=1} \\
&= \sum_{n=1}^{\infty} 0.5^n = 1
\end{aligned}$$

We have

$$
\begin{aligned}
\sum_{u\in\mathcal{V},v\in\mathcal{V}} P(X_1=u, X_2=v) &= \sum_{u\in\mathcal{V},v\in\mathcal{V}} P(X_1=u)\times P(X_2=v) \\
&= \sum_{u\in\mathcal{V}} P(X_1=u) \underbrace{\sum_{v\in\mathcal{V}} P(X_2=v)}_{=1} \\
&= \sum_{u\in\mathcal{V}} P(X_1=u) \\
&= 1
\end{aligned}
$$

We have

$$
\begin{aligned}
& \sum_{u \in \mathcal{V}, v \in \mathcal{V}} P(X_1 = u, X_2 = v) \\
= & \sum_{u \in \mathcal{V}, v \in \mathcal{V}} P(X_1 = u) \times P(X_2 = v | X_1 = u) \\
= & \sum_{u \in \mathcal{V}} P(X_1 = u) \underbrace{\sum_{v \in \mathcal{V}} P(X_2 = v | X_1 = u)}_{=1} \\
= & \sum_{u \in \mathcal{V}} P(X_1 = u) \\
= & \ 1
\end{aligned}
$$

$$
\begin{aligned}
& p(\text{He saw their was a football in the park ?}) \\
= \; & q(\text{He}) \times q(\text{saw}) \times q(\text{their}) \times q(\text{was}) \times \ldots
\end{aligned}
$$

$$
\begin{aligned}
& p(\text{He saw there was a football in the park ?}) \\
= \; & q(\text{He}) \times q(\text{saw}) \times q(\text{there}) \times q(\text{was}) \times \ldots
\end{aligned}
$$

$$
\begin{aligned}
& p(\text{He saw their was a football in the park ?}) \\
> \; & p(\text{He saw there was a football in the park ?})
\end{aligned}
$$

if and only if

$$q(\text{their}) > q(\text{there})$$

$$p(\text{He saw their was a football in the park ?})$$
$$= q(\text{He}) \times q(\text{saw}|\text{He}) \times q(\text{their}|\text{saw}) \times q(\text{was}|\text{their}) \times \ldots$$

$$p(\text{He saw there was a football in the park ?})$$
$$= q(\text{He}) \times q(\text{saw}|\text{He}) \times q(\text{there}|\text{saw}) \times q(\text{was}|\text{there}) \times \ldots$$

- Model is now *sensitive to context* (word before or after *their* or *there*)
- But if Count$(w_{i-1}, w_i) = 0$ for any pair of words, then $p(w_1 \ldots w_n) = 0$, which will cause problems.

1a) The dog in the park was big

1b) The dogs in the park were big

2) The dog which the cat saw is big

There are many other examples