Out-of-vocabulary words

Toy train corpus:
This is the house that Jack built.

Toy test corpus:
This is the *malt*.

What's the perplexity of the Bigram LM?

$$p(malt|the) = \frac{c(the\,malt)}{c(the)} = 0$$

$$p(\mathbf{w}_{\text{test}}) = 0$$

$$\mathcal{P} = \inf$$