As the number of degrees past the highest probability guess increased: the text became increasingly
nonsensical. Somewhere before losing coherence: around degree 40 in my experiment, the text reached
the ideal mix of absurdism while maintaining coherence. For instance in:

Of the pine-trees crusted with seeds
And have been cold a long Time
...
he spruces rough in the distant years
...
Of any misery in the sound of the old

A theme of age adds to the snowy atmosphere of the poem, and repeats appropriately throughout.
While iterating in degrees higher still, by degree 69 the model loses coherence:

One must have a mind of at
To regard the frost and the p
Of the pine-trees crusted with golden

Although the last line is interesting, the first two don't mean much of anything. The model is reduced to low probability fragment tokens whose purpose is to connect to a new idea, rather than predicting a token representing a meaningful word, or idea.

To train a model to find the nouns within a poem is challenging because the model does not have a
method to classify symbolic abstractions like nouns. Also, words should be classified differently
depending on their grammatical contexts.
Because the model is already doing the work of establishing a context through
its tokenizations, an additional layer could be added after the tokenizations, which maps each group
of tokens associated with a word, to one type of word: nouns, adjectives, verbs, adverbs. Each type
could be associated to a number.

The model could then reprocess the input by assigning probabilities of each token-group's association to each word type. The output would not perfectly classify the grammar, but with a large enough dataset, could get extremely close.