Code

We performed three studies to verify the robustness of our methods.

1. First, we evaluated the accuracy over 100 **bootstrap iterations**: in each iteration the original data is replaced by a random sample over lemmas with repetitions. The resulting accuracy is 90\% with standard error across iterations of \colorbox{yellow}{(std)}. This check proposes that the method does not heavily rely on specific features and thus robust against small changes in the data.
2. Next, we raised the question \textit{is the method data dependent?} We applied **bagging** by randomly selecting the train and test chapters, and evaluating the attribution accuracy which resulted in 91\%. In addition, we verified that \textit{the accuracy is robust with respect to the selected feature}, i.e. single term, bigram, trigram.
3. Finally, we evaluated the \textit{**accuracy of the HC-attribution with respect to the number words**} in a tested text. We see that already chapters of length 700 words result in a 90\% accuracy, which is rather good for such a low amount of words. For additional details about these tests see Supporting Information.
4. Confidence intervals – didn’t send you the code, but its basically base on #1

Data need for Figures/ tables

1. HC values of all the tested data vs the 3 corpora (used for: Table 1, 3D plot)

Both the training data

And unk

1. Probability attribution (figure 5)
   1. Your method
   2. Gaussian attribution
2. Separation words – need to think what exactly to put here, I think we should put the following:
   1. Separation between the 3 reference corpora
   2. Select 1-2 chapters, and provide the separation from the 3 ref data
   3. The rest put online – so that everyone can check on their own