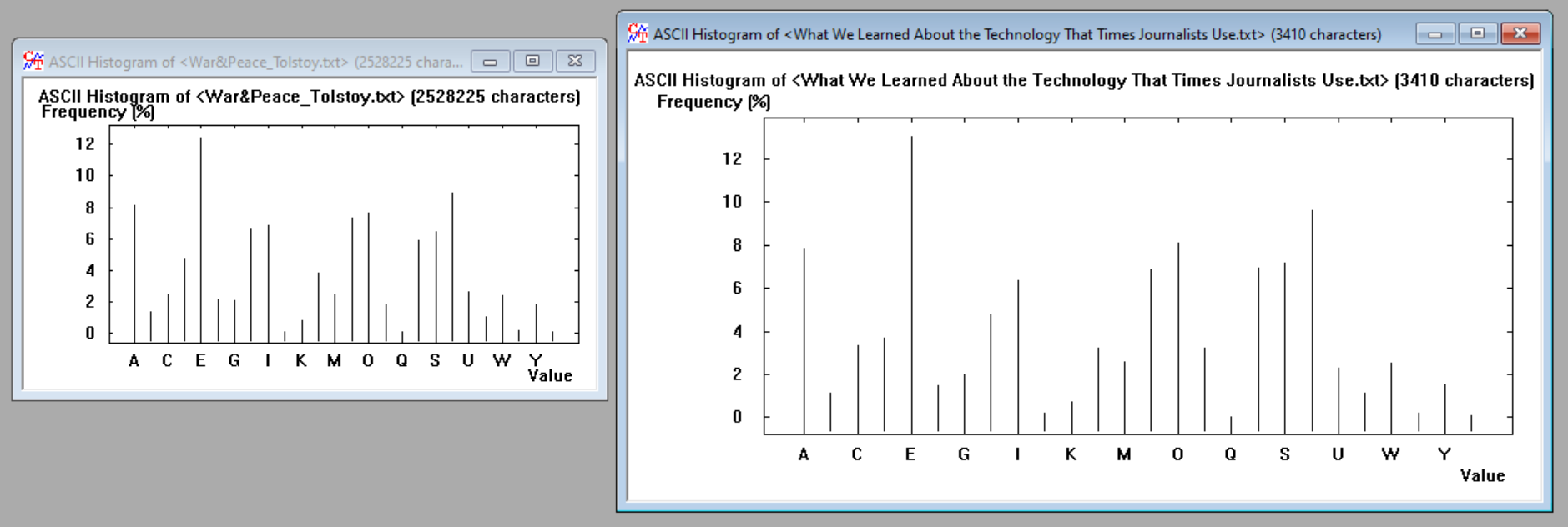
**Homework 5 – Cryptography**

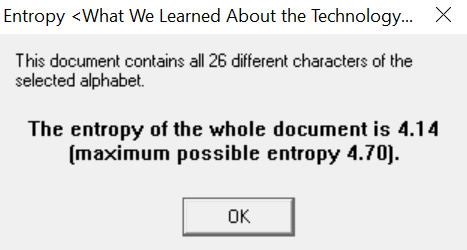
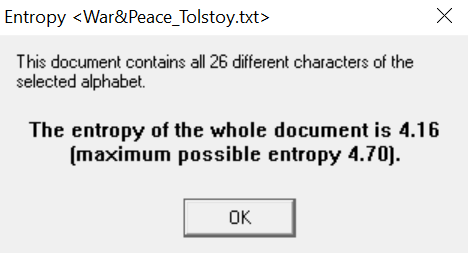
|  |
| --- |
| * This is an individual assignment, and is worth 20 points. * The due date is Wednesday, October 7 midnight. * You need to provide your answers using the Outcome file. * Follow the usual naming convention. * Do not make screenshots too small. |

## Task 1 (4 points)

* (1.5 points) Create a histogram for each text that display the relative frequency of letters in a graphical form. For this, go to Analysis > Tools for Analysis. Provide the two histograms in screenshots.



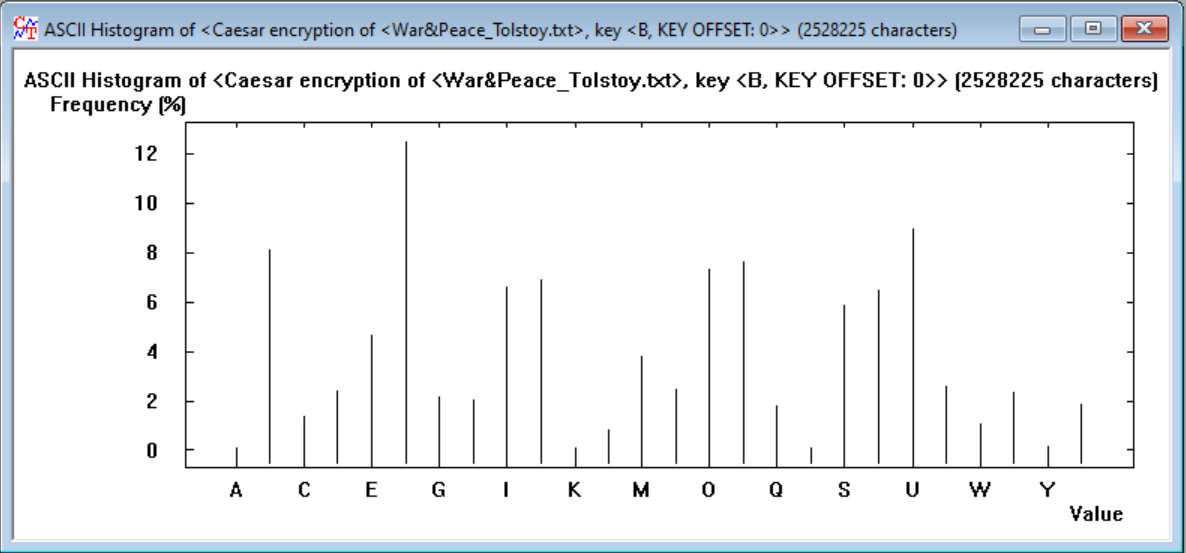
* (1.5 points) Do the two histograms depend significantly on the texts you have provided?
  + No, the English language has a letter frequency which is simply the amount of times letters of the alphabet appear on average in written language. The histograms look nearly identical despite being on different subjects and one being much longer in length.
* (1 point) Calculate the entropy of each text. For this, Analysis > Tools for Analysis > Entropy. What would you conclude from the comparison of the entropies?



* + The chosen texts are 0.2 points apart. War & Peace is longer and has a higher entropy, this means it is harder to convert into bits. It’s interesting because it is considerably longer, but not far off from the NYT article.

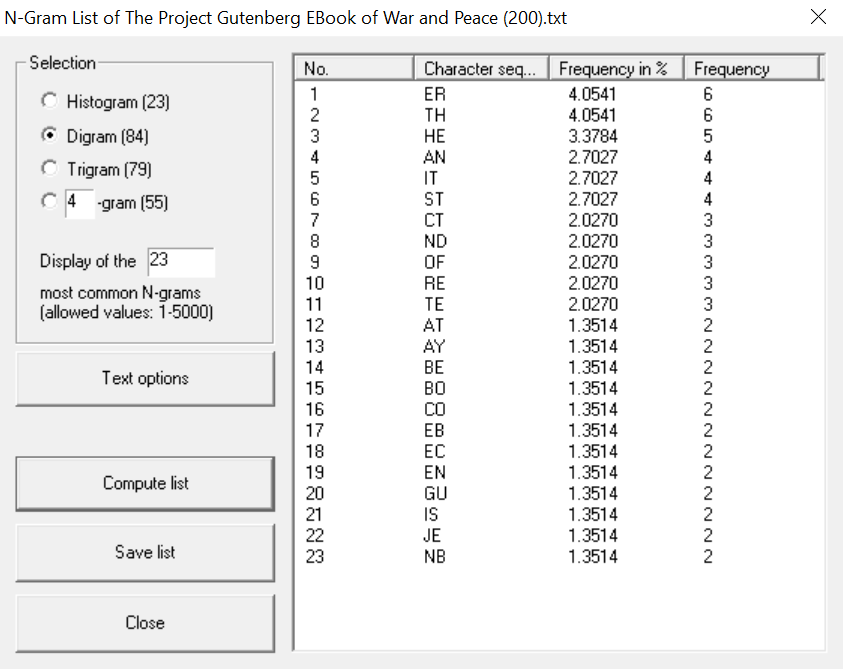
## Task 2 (4 points)

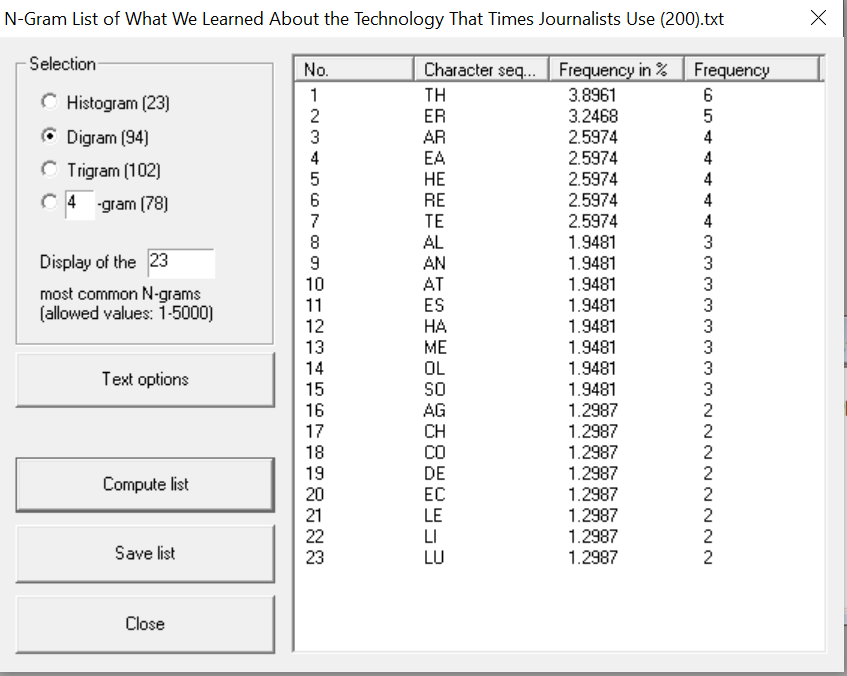
* (2 points) What are the characteristic features of the obtained distribution compared with the original text? Provide a screenshot to support your answer.



* + Everything shifted over one place to the right because I selected B as the key entry. This explains everything moving over by one since one was the number value of the shift.
* (2 points) How would you apply the features you have discovered in cracking the key?
  + Visually you can see that the frequencies have shifted, cracking a simple encryption attempt like this wouldn’t be too difficult with the assistance of both histograms. Without both, this tool does not provide much help.

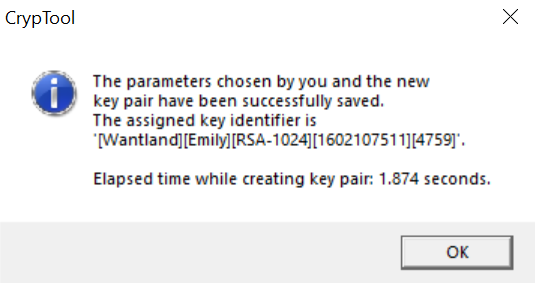
## Task 3 (4 points)

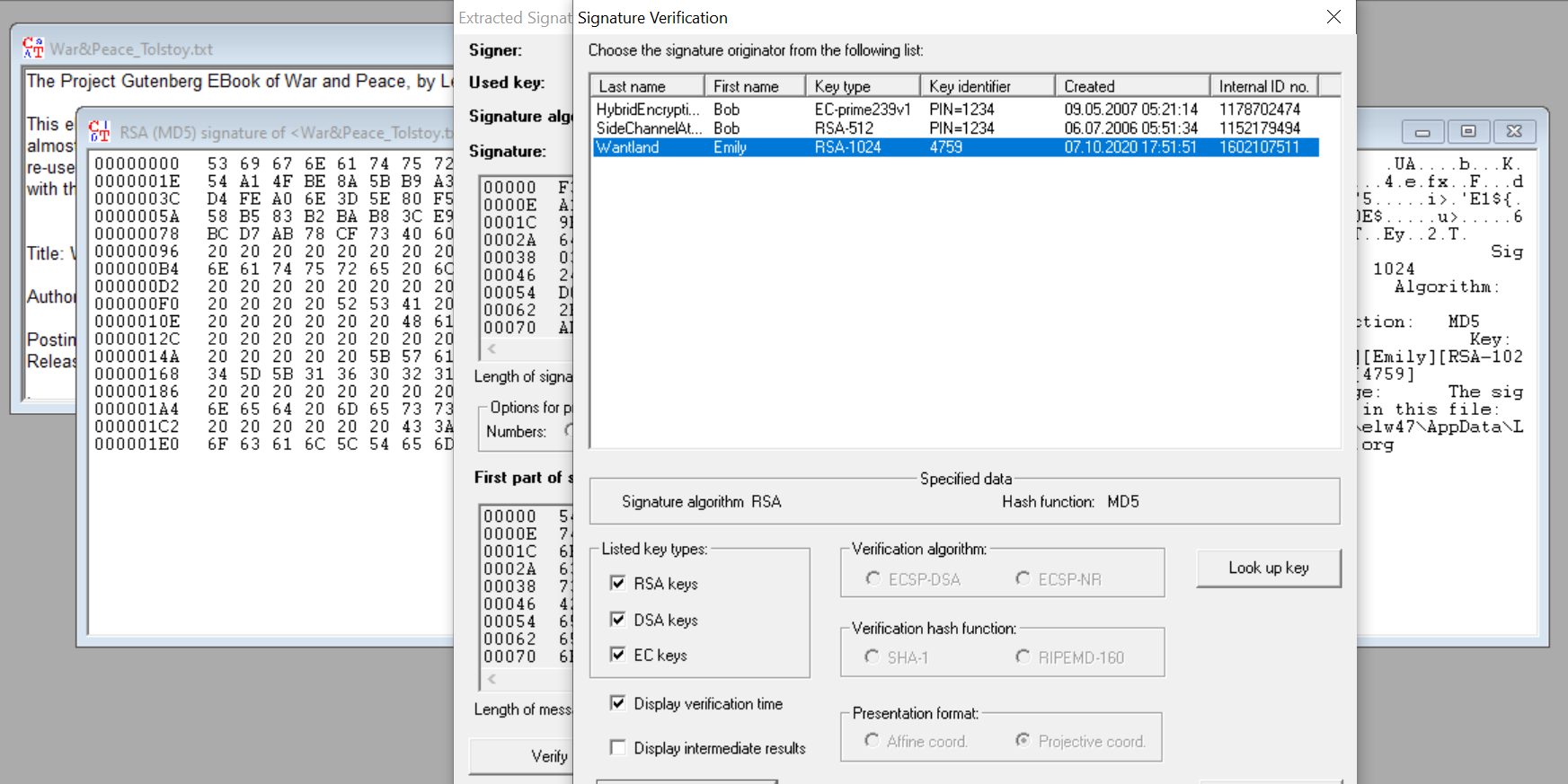
* (1.5 points for War&Peace\_Tolstoy.txt) Summarize your observation of the frequency distribution. Also, provide a screenshot of the frequency distribution.
* 
  + I’m surprised ER has the greatest frequency, I thought it would be SE.
* (1.5 points for a NYT article) Summarize your observation of the frequency distribution. Also, provide a screenshot of the frequency distribution.



* + TH makes sense to me, “the” is an extremely common word so it being the most frequent is understandable.
* (1 point) Compare and contrast the two frequency distributions.
  + These are similar to the histograms, in the top five “ER”, “TH”, and “HE” appear in both. The letter combinations present are the letters that appear to have higher columns in the histograms. Even though both texts are very different, they still have similar character sequences, though the frequency differs.

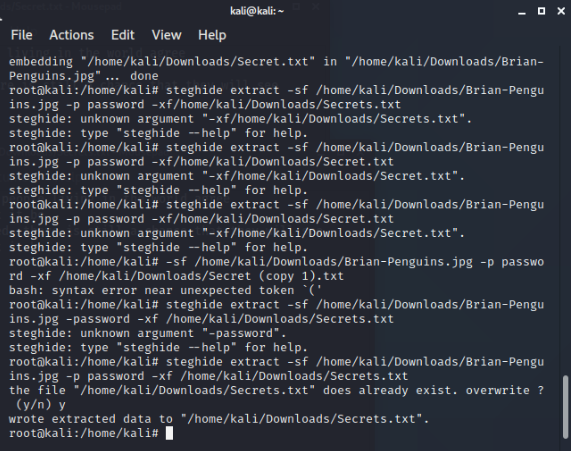
## Task 4 (4 points)

* (2 points) Attach a screen shot that shows the successful creation of the key pair.
* 
* (2 points) Attach a screenshot that displays the signature verification.



## Task 5 (4 points)

* (2 points) Follow the steps in the video and show the recovered secret text file as below.



* (2 points) Show in screenshots that the contents of the two files are the same.

|  |
| --- |
| The image of the original file |
|  |
| The image of the recovered file |