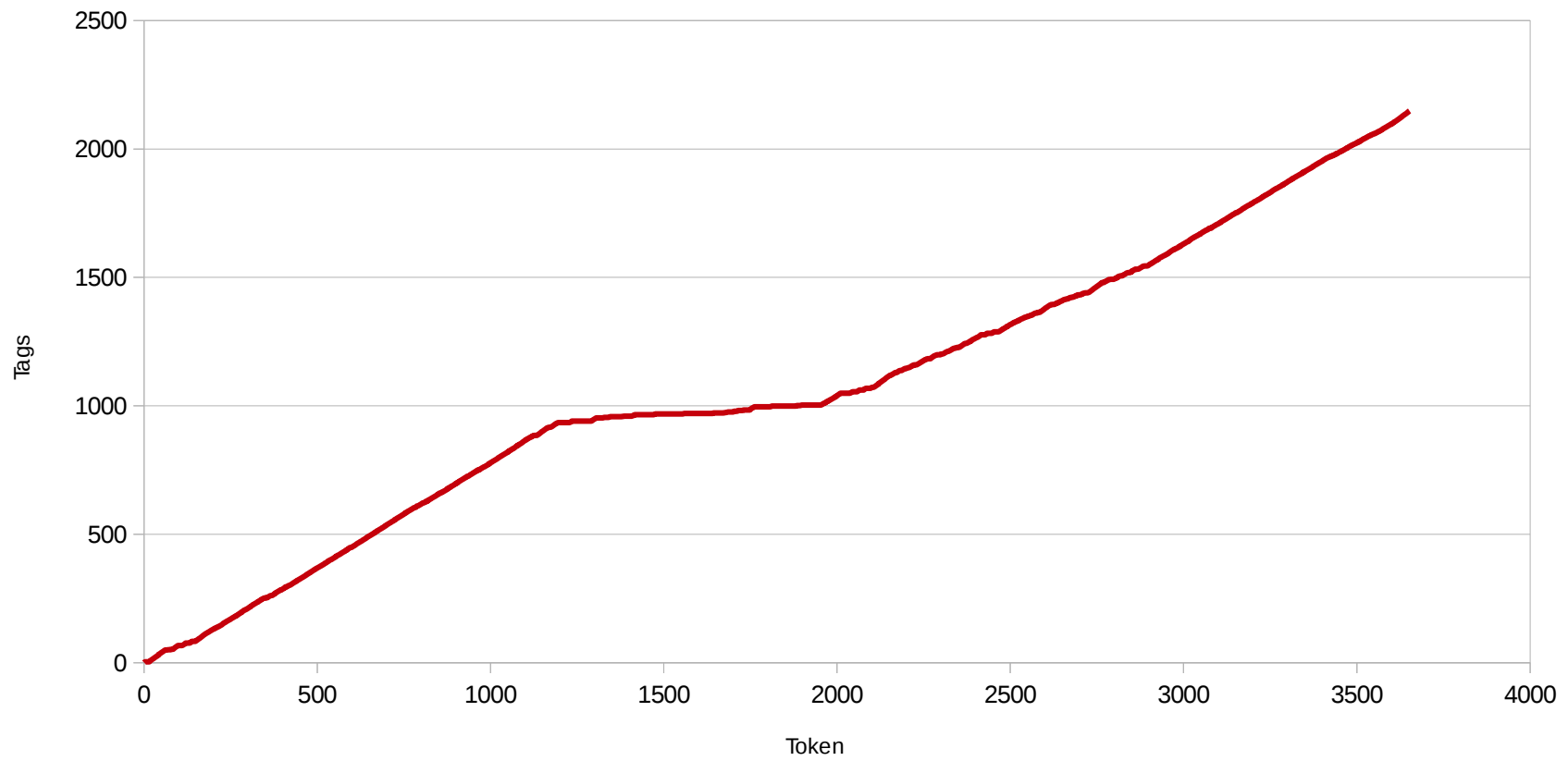


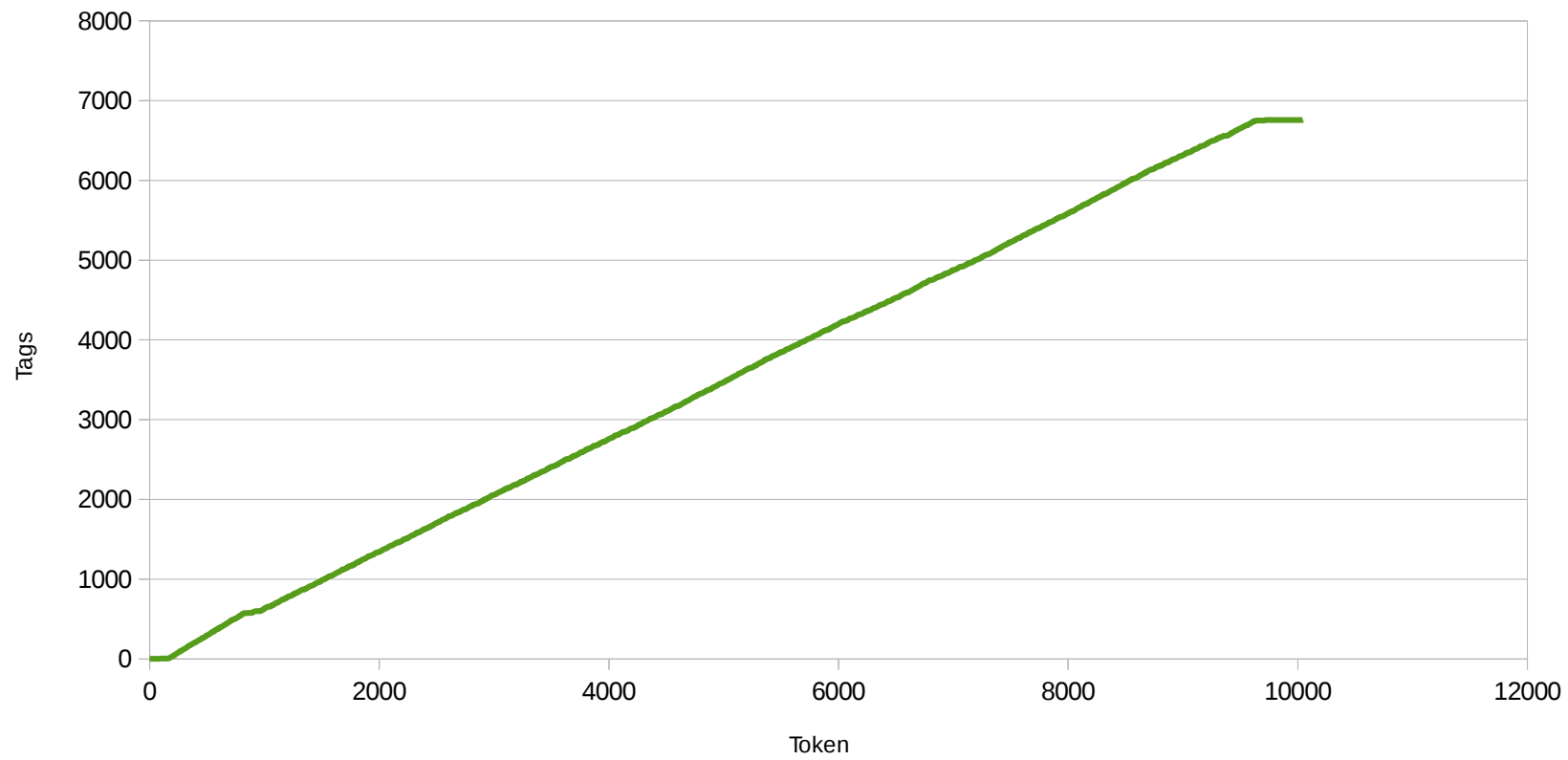
# Task 1

Spiegel online



# Task 1

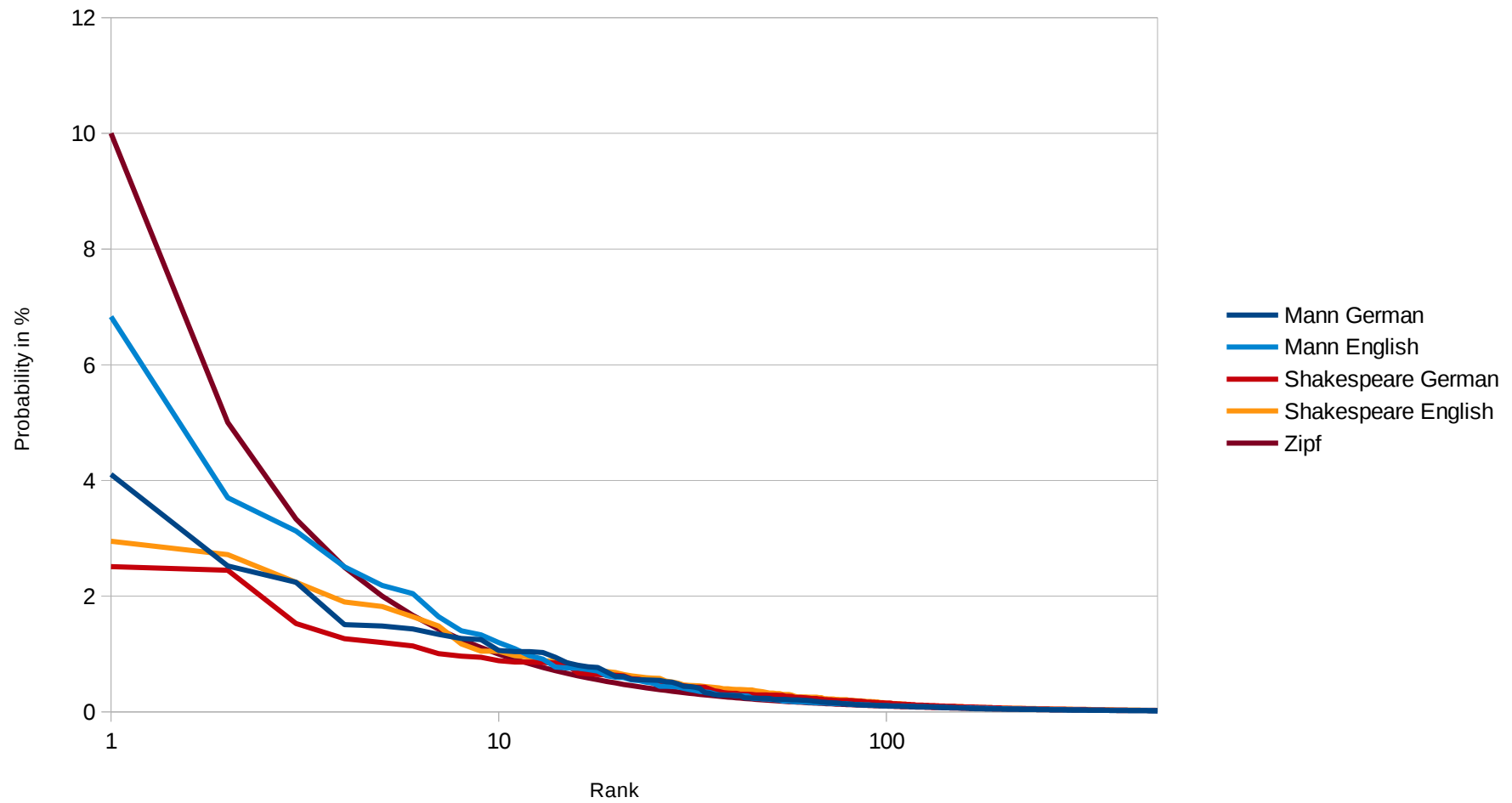
YouTube



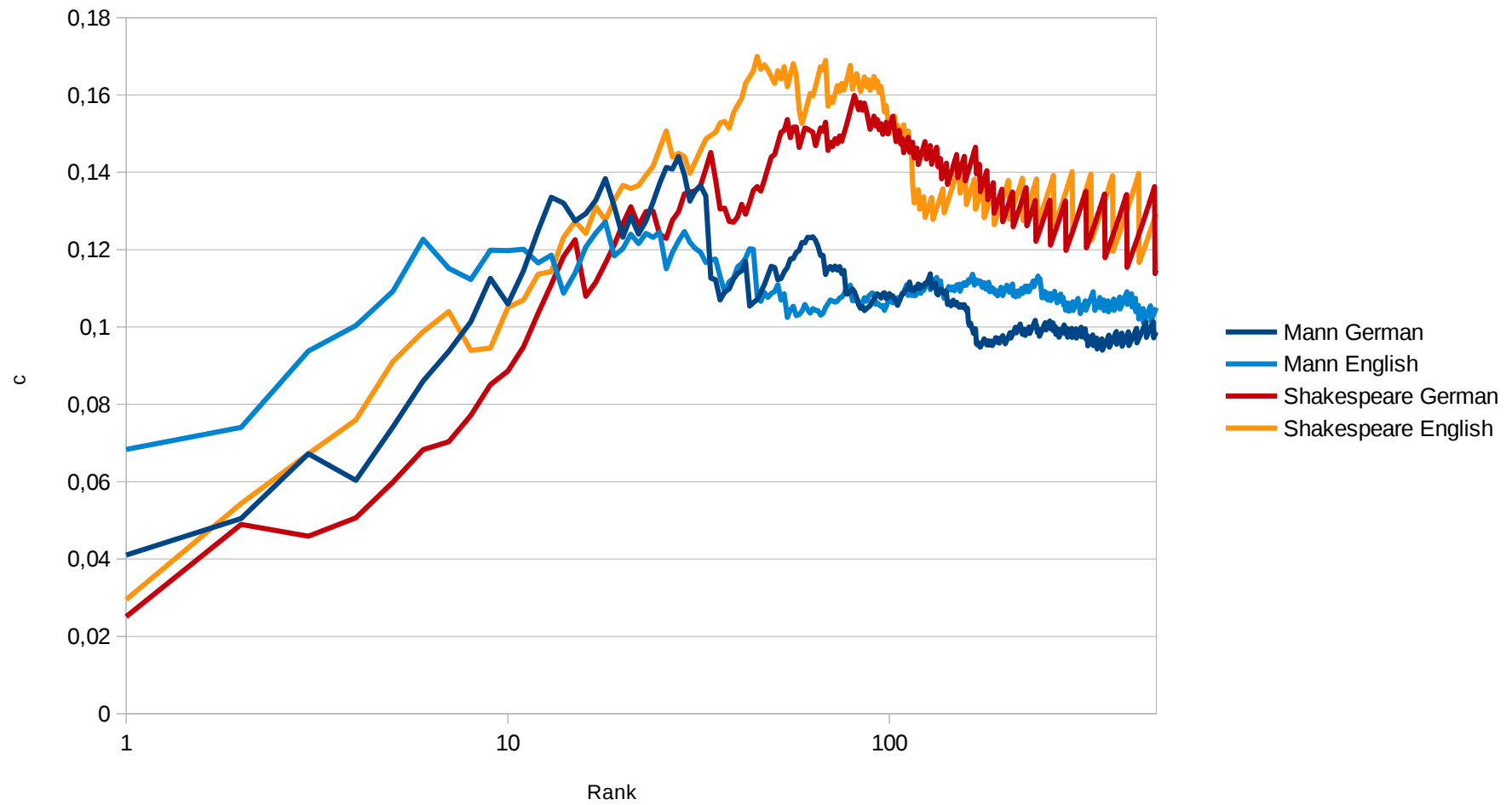
# Task 2

- Byte sum has same output for documents with the same characters in different order

# Task 4



# Task 4



# Task 6

stopword list: `src/exercise_6/english.stop`

source:

`http://jmlr.org/papers/volume5/lewis04a/a11-smart-stop-list/english.stop`

# Task 6

- behind  
star wars behind the scenes
- without  
cocktail without alcohol
- changes  
changes of birth rate
- welcome  
welcome present
- will  
napoleon last will

# Task 7

- Thomas Mann: Royal Highness
  - 119.035 Words
- Stemmers of snowball implementation
  - „english“ stemmer: 7.821 stems
  - „porter“ stemmer: 8.282 stems