



Aim

Crack initial digits of the Voynich manuscript and determine the possible letters which may stand for digits.

Method

- In order to determine which language this manuscript uses, we study it from the perspective of possible written numbers. Moreover, we search for digits possibly related enumerated items in the manuscript's illustrations by employing data mining techniques.
- Matlab

Significance

The project group provide a possible method to crack the Voynich manuscript. In addition, this method can be expand to decoding area.

Background

The Voynich manuscript is a mysterious 15th century book. No one today knows what it says or who wrote it. In addition, the Voynich manuscript involves many peculiar illustrations and symbols.

Phase 1: Text Investigation

By comparing with books in known language, some potential relationship can be found from the following figures. Figure 1 shows the ratio of unique words/ total words. Figure 2 shows the word length. Figure 3 shows the percentage of words appear more than once. Result: There is the highest correlation between the Voynich manuscript and German.

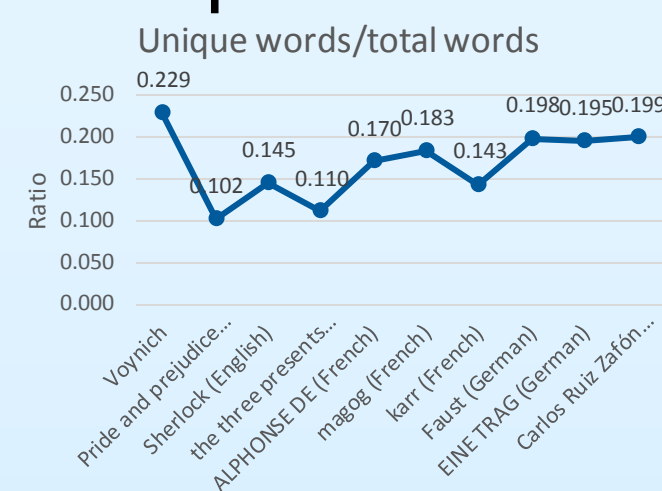


Figure 1

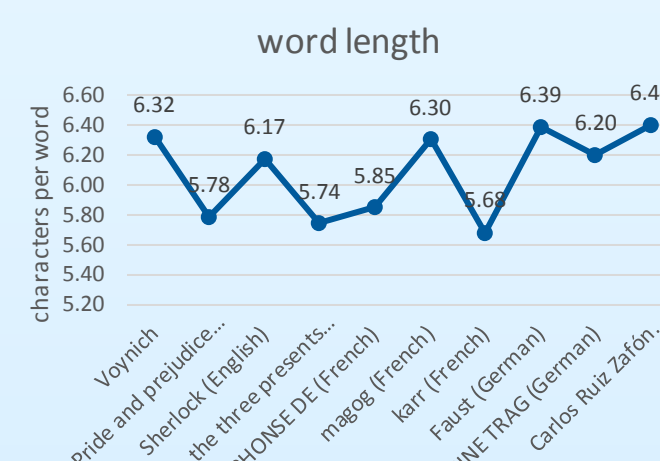


Figure 2

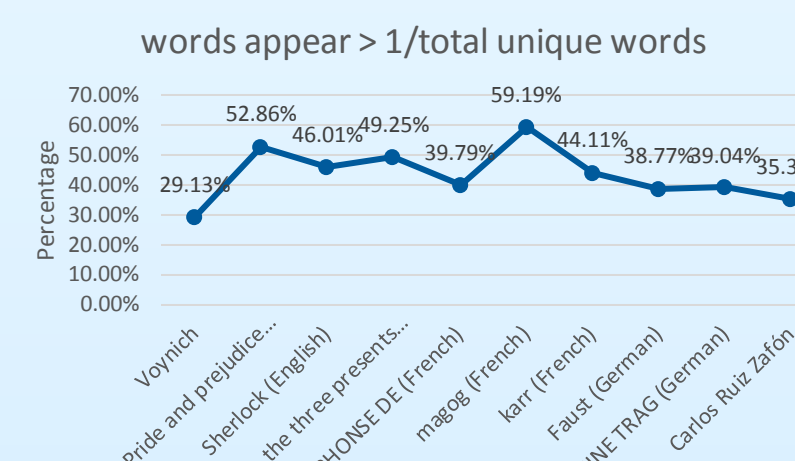


Figure 3

Results

- The most Possible language: German
- The most Possible digits: 'o' and '1'
'ol' and '13'
'or' and '13'
- Possible digits: 's' and '2', '6'
'y' and '2', '6'
'a' and '1'
'r' and '1'

Phase 2: Illustrations Research

Search potential digits from illustrations in the Voynich manuscript by using data mining.

| letter | number | frequency |
|--------|--------|-----------|
| o | 1 | 24 |
| | 2 | 20 |
| | 3 | 16 |
| | 4 | 14 |
| | 5 | 12 |
| ol | 10 | 14 |
| | 13 | 12 |
| | 12 | 11 |
| or | 10 | 19 |
| | 12 | 13 |
| | 13 | 12 |
| os | 19 | 11 |
| | 12 | 4 |
| | 13 | 4 |

Figure 1

| | | |
|---|---|----|
| r | 1 | 48 |
| | 2 | 26 |
| | 3 | 21 |
| | 5 | 19 |
| | 4 | 14 |
| s | 2 | 54 |
| | 1 | 46 |
| | 3 | 41 |
| y | 5 | 32 |
| | 2 | 36 |
| | 1 | 30 |
| | 3 | 29 |
| | 5 | 20 |

Figure 2

Result: There is a potential relationship between:

- 'o' and '1'
- 'ol' and '10'
- 'ol' and '13'
- 'ol' and '12'
- 'or' and '10'
- 'or' and '12'
- 'or' and '13'
- 'os' and '19'
- 'r' and '1'
- 'r' and '2'
- 's' and '2'
- 's' and '1'
- 's' and '3'
- 'y' and '2'
- 'y' and '1'
- 'y' and '3'

Phase 3: Marginal symbol Investigation

Search potential illustrations in the Voynich manuscript by using data mining.

| letter | Pages (25) | The digit which letter may stand for | Pr |
|--------|------------|--------------------------------------|--------|
| y | 18 | 5 | 16.67% |
| | | 6 | 38.89% |
| | | 7 | 55.56% |
| | | 8 | 27.78% |
| | | 9 | 11.11% |
| | | 5 | 38.46% |
| | | 7 | 53.85% |
| | | 8 | 30.77% |
| | | 9 | 15.38% |
| l | 13 | 5 | 38.46% |
| | | 7 | 53.85% |
| r | 7 | 5 | 28.57% |
| | | 6 | 40% |
| | | 7 | 30% |
| s | 10 | 6 | 40% |
| | | 7 | 30% |
| | | 8 | 50% |
| o | 8 | 8 | 25% |
| | | 9 | 25% |
| | | 10 | 18.18% |
| ar | 22 | 12 | 4.46% |
| | | 13 | 27.27% |
| | | 14 | 9.1% |
| | | 15 | 13.64% |
| | | 16 | 9.1% |
| | | 19 | 9.1% |
| | | 10 | 19.05% |
| al | 21 | 13 | 28.57% |
| | | 14 | 9.52% |
| | | 15 | 19.05% |
| or | 17 | 10 | 17.65% |
| | | 13 | 29.41% |
| | | 16 | 11.76% |
| ol | 21 | 10 | 14.29% |
| | | 12 | 14.29% |
| | | 13 | 23.81% |
| am | 12 | 12 | 16.67% |
| | | 19 | 16.67% |
| | | 16 | 9.52% |
| dy | 6 | 14 | 33.33% |
| | | 19 | 33.33% |
| | | 16 | 40% |

Result: From the figures, there is a

potential relationship between:

- 'y' and '6'
- 'y' and '7'
- 'l' and '7'
- 'l' and '5'
- 'r' and '5'
- 's' and '6'
- 'o' and '6'
- 'o' and '1'
- 'ar' and '13'
- 'ar' and '15'
- 'a' and '13'
- 'a' and '10'
- 'a' and '15'
- 'or' and '13'
- 'ol' and '13'
- 'am' and 12
- 'am' and '19'
- 'dy' and '14'
- 'dy' and '19'
- 'om' and '16'