Graphical Frequency Analysis

N-gram analysis that works interactively and graphically

- **Analysis**
- **Description**
- **Background**
- **Language frequency**

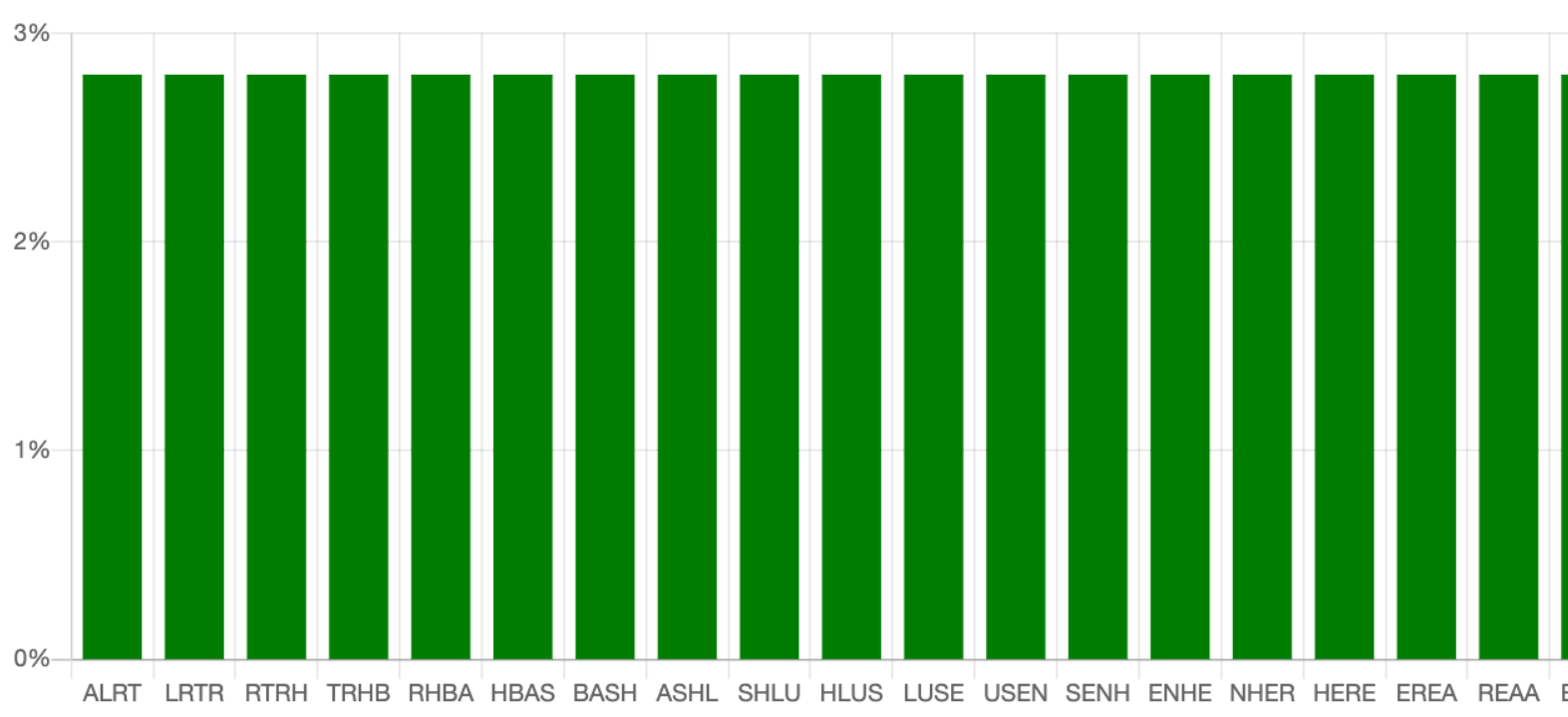Text to analyze [Restrict to English alphabet] **English German**

```
ALRTRHBASHLUSENHEREAAEC
LTCMHIJNDHEIENGG
```

Text length: 39

N-gram: [-] 4 [+]

Number of bars:

26 ⦿ Sort by frequency ○ Sort alphabetically

⬇ N-gram download

Author: Jean Michel Polak

This tool allows you to analyze the frequency of letters in your text and display them clearly in a diagram. The analysis can be performed for unigrams (single letters), bigrams (two-letter groups), trigrams (three-letter groups) or longer.

You can choose between the German and English alphabet. Only the letters from the selected alphabet are analyzed. You can also download the result of the analysis in CSV-format.
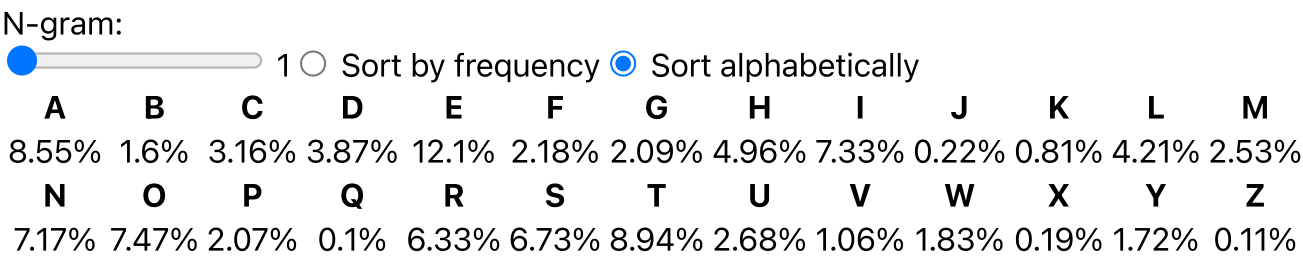
In each language, the individual letters occur in a sufficiently long, natural text at a frequency characteristic of the language. In English the most common letter is E with a frequency of about 12%, followed by the letter T with about 9%.

A monoalphabetic substitution cipher can be easily broken with a frequency analysis. Such a cipher can be recognized by the fact that never two plaintext characters are mapped by the same ciphertext character.
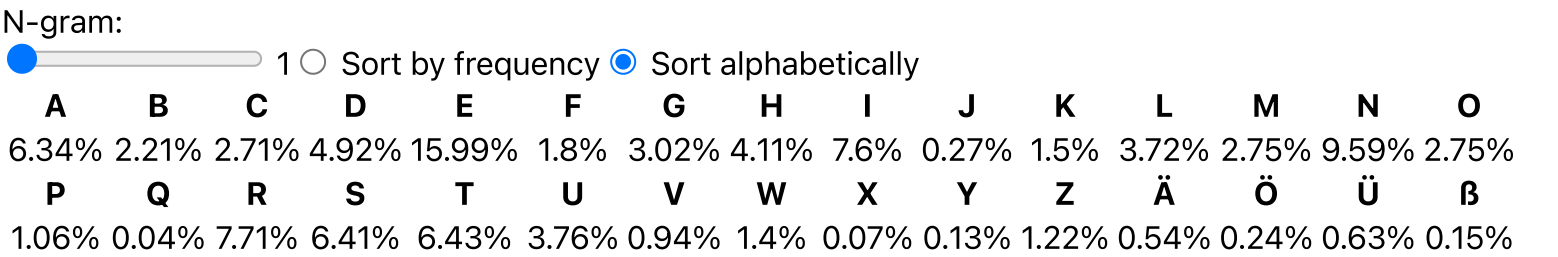
Each plaintext character is assigned one or more ciphertext characters (in this case the frequency analysis is much more difficult). To crack the cipher, the occurring ciphertext characters are counted and compared with the average frequency of plaintext letters in the corresponding national language.

In this tab you will find the frequency of 26 and 30 single letters for German and English respectively. If you move the slider for N-grams, you can also display the 26 or 30 most frequent N-grams for N = 2, 3, 4 and 5 (Default is 1).

Frequency for the English language

N-gram:

1 ○ Sort by frequency ● Sort alphabetically

A	B	C	D	E	F	G	H	I	J	K	L	M
8.55%	1.6%	3.16%	3.87%	12.1%	2.18%	2.09%	4.96%	7.33%	0.22%	0.81%	4.21%	2.53%
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
7.17%	7.47%	2.07%	0.1%	6.33%	6.73%	8.94%	2.68%	1.06%	1.83%	0.19%	1.72%	0.11%

Frequency for the German language

N-gram:

1 ○ Sort by frequency ● Sort alphabetically

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
6.34%	2.21%	2.71%	4.92%	15.99%	1.8%	3.02%	4.11%	7.6%	0.27%	1.5%	3.72%	2.75%	9.59%	2.75%
P	Q	R	S	T	U	V	W	X	Y	Z	Ä	Ö	Ü	ß
1.06%	0.04%	7.71%	6.41%	6.43%	3.76%	0.94%	1.4%	0.07%	0.13%	1.22%	0.54%	0.24%	0.63%	0.15%

ference

requencies were used from: **practicalcryptography**

⬀ Share link