Elias' Plan of Concentration

Components

“Machine Translation” Paper – 30%

Components

“Machine Translation” Paper – 30%
“The Voynich Manuscript” Paper – 20%

Components

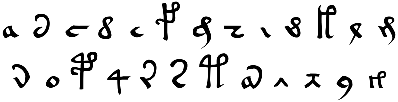
“Machine Translation” Paper – 30%
“The Voynich Manuscript” Paper – 20%
“Sociology of Intonation” Paper – 20%

The Voynich Manuscript

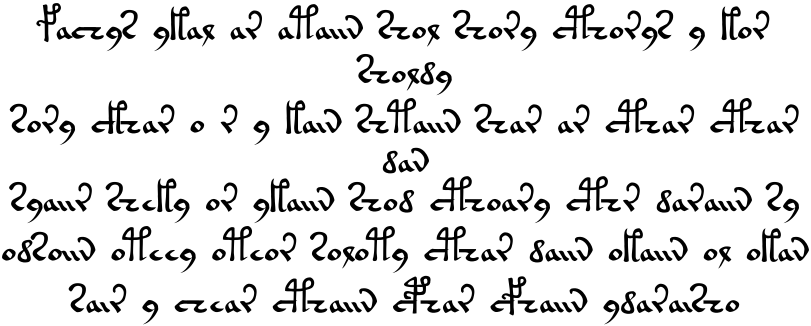
The Voynich Manuscript: a book composed of 116 folios of text in an unknown alphabet.

The Voynich Manuscript

The Voynich Manuscript: a book composed of 116 folios of text in an unknown alphabet.



The Voynich Manuscript



The Voynich Manuscript

fachys ykal ar ataiin shol shory cthorys y kor sholdy
sory ckhar o r y kain shtaiin shar ar cthar cthar dan
syaiir sheky or ykaiin shod cthoary cthr daraiin sy
odsoiin oteey oteor soloty cthar daiin okaiin ol okan
sair y chear cthaiin cphar cfhaiin ydaraisho

The Voynich Manuscript

fa**v**ys ykal ar ata**M** **S**ol **S**ory **T**orys y kor **S**ol**D** sory ckhar o r y kain **S**ta**M** **S**ar ar **T**ar **T**ar dan sya**R** **S**eky or yka**M** **S**od **T**oary **T**r dara**M** sy odso**M** oteey oteor soloty **T**ar da**M** oka**M** ol okan sair y **v**ear **T**a**M** cphar cfha**M** ydarai**S**o

The Voynich Manuscript
Text Analysis

word	freq	word	freq	word	freq	word	freq
<daiin>	489	<dy>	818	<n>	989	the	1008
<chedy>	263	<y>	438	<daM>	650	a	577
	256		434		503	and	570
<shedy>	253	<chy>	349	<veD>	478	of	530
<chol>	238	<qol>	336	<SeD>	370	to	395
<qokedy>	193	<or>	298	<aM>	338	in	350
<dar>	166	<qor>	258	<vol>	327	I	342
<dy>	164	<s>	251	<D>	297	that	227
<qokeedy>	162	<l>	219	<vey>	294	his	214
<qokain>	157	<chey>	216	<s>	285	it	178

Table: Top 10 words: Voynichese, Generated Voynichese, Modified Voynichese, and English

The Voynich Manuscript

Text Analysis

word	freq	word	freq	word	freq
the	1008	de (*of*)	1011	und (*and*)	455
a	577	et (*and*)	432	die (*the*)	328
and	570	des (*of*)	418	der (*the*)	263
of	530	les (*the*)	376	sie (*you*)	212
to	395	à (*to*)	318	er (*he*)	210
in	350	la (*the*)	312	den (*the*)	180
I	342	qui (*who*)	242	in (*in*)	154
that	227	le (*the*)	238	das (*the*)	148
his	214	en (*in*)	230	ich (*I*)	148
it	178	nous (*we, us*)	222	nicht (*not*)	144

Table: Top 10 words: English, French, German

The Voynich Manuscript

Text Analysis

word	freq	word	freq	word	freq
the	1008	**de** (*of*)	1011	**und** (*and*)	455
a	577	**et** (*and*)	432	**die** (*the*)	328
and	570	**des** (*the*)	418	**der** (*the*)	263
of	530	**les** (*the*)	376	**sie** (*they, it*)	212
to	395	**à** (*to*)	318	er (*he*)	210
in	350	**la** (*the*)	312	**den** (*the*)	180
I	342	**qui** (*who, that*)	242	**in** (*in*)	154
that	227	**le** (*the*)	238	**das** (*the*)	148
his	214	**en** (*in, to*)	230	**ich** (*I*)	148
it	178	nous (*we, us*)	222	nicht (*not*)	144

Table: Top 10 words: English, French, German

The Voynich Manuscript
Text Analysis: Zipf's Law

If r is the rank of a word, and f is the frequency,

$$f(r) \propto \frac{1}{r}$$

The Voynich Manuscript
Text Analysis: Zipf's Law

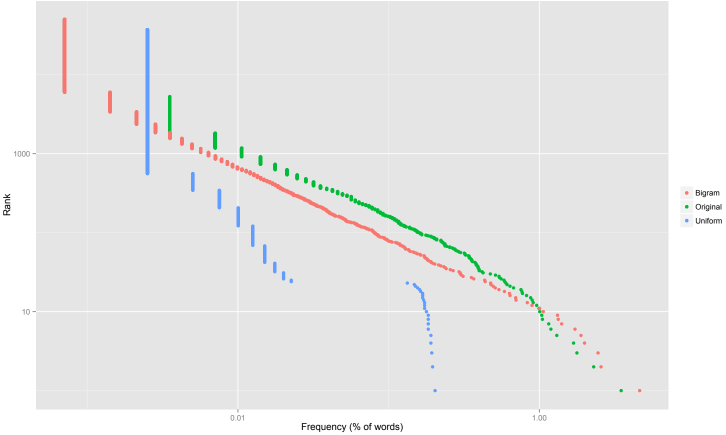


Figure: Zipf's Law Comparison

The Voynich Manuscript
Text Analysis: Characters

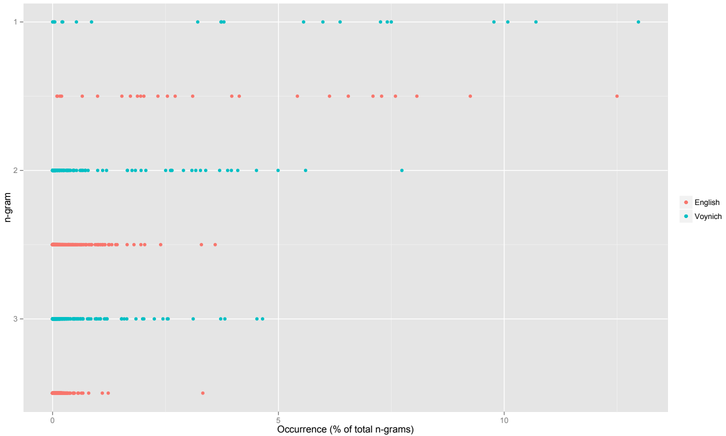


Figure: *n*-gram counts

The Voynich Manuscript
Text Analysis: Characters

What would happen if we replaced some of the outliers?

The Voynich Manuscript
Text Analysis: Characters

What would happen if we replaced some of the outliers?

ch -> v
iin -> M
sh -> S
cth -> T
dy -> D
iir -> R

The Voynich Manuscript
Text Analysis: Characters

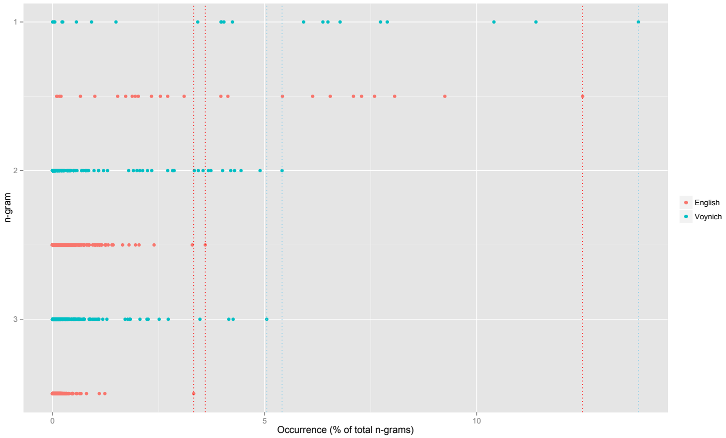


Figure: *n*-gram counts

The Voynich Manuscript
Text Analysis: Characters

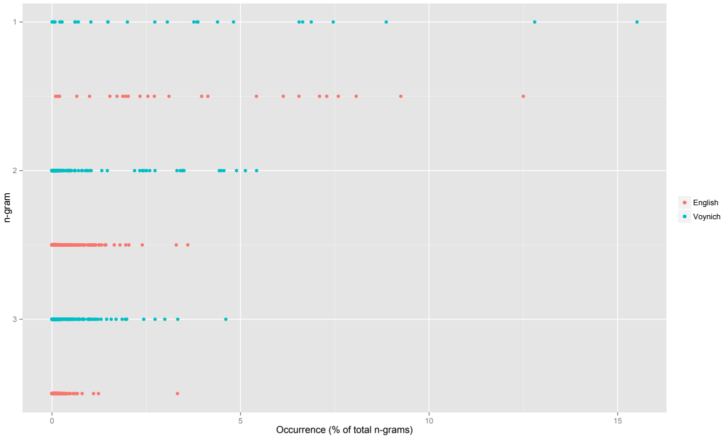


Figure: *n*-gram counts

The Voynich Manuscript
Text Analysis: Characters

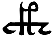
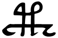
voy	eng	ger	voy	eng	ger	voy	eng	ger
	a	i		k	p		u	o
	b	z		l	u		v	ü
	c	l		m	b		w	k
	d	g		n	a		x	ä
	e	e		o	s		y	w
	f	f		p	v		z	ß
	g	m		q	j			q
	h	d		r	h			
	i	r		s	t			
	j	ö		t	n			

Table: Voynich–English–German equivalencies?

!

<ralmo vDSoe>

Thank you!

Merci!