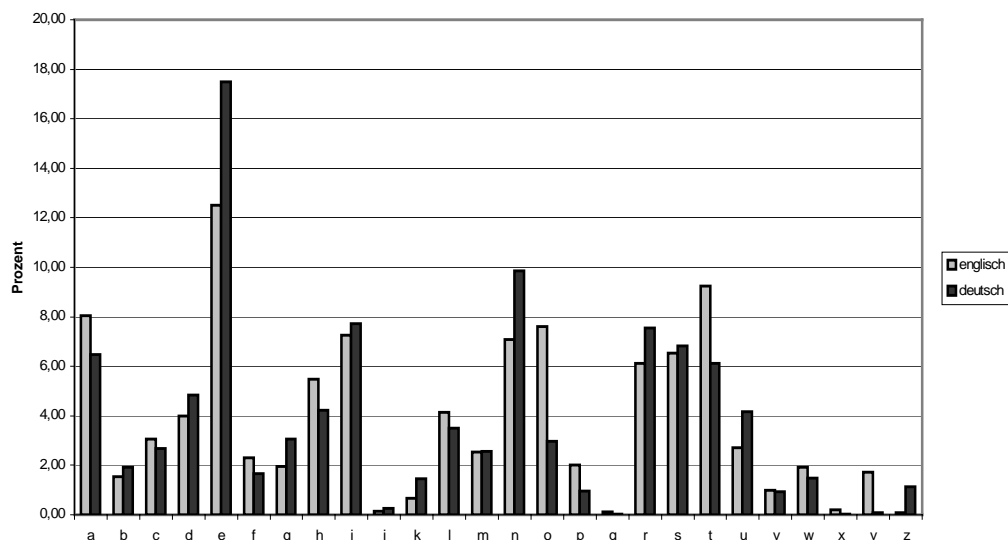


Hilfstabellen zur monoalphabetische Entschlüsselung

Hypothetische Zeichenwahrscheinlichkeiten im Englischen und im Deutschen:

	<i>englisch</i>	<i>deutsch</i>
a	8,04 %	6,47 %
b	1,54 %	1,93 %
c	3,06 %	2,68 %
d	3,99 %	4,83 %
e	12,51 %	17,48 %
f	2,30 %	1,65 %
g	1,96 %	3,06 %
h	5,49 %	4,23 %
i	7,26 %	7,73 %
j	0,16 %	0,27 %
k	0,67 %	1,46 %
l	4,14 %	3,49 %
m	2,53 %	2,58 %

	<i>englisch</i>	<i>deutsch</i>
n	7,09 %	9,84 %
o	7,60 %	2,98 %
p	2,00 %	0,96 %
q	0,11 %	0,02 %
r	6,12 %	7,54 %
s	6,54 %	6,83 %
t	9,25 %	6,13 %
u	2,71 %	4,17 %
v	0,99 %	0,94 %
w	1,92 %	1,48 %
x	0,19 %	0,04 %
y	1,73 %	0,08 %
z	0,09 %	1,14 %



Für die deutsche Sprache erhält man die im wesentlichen schon von *Lange-Soudart* 1935 angegebene Partition in die Cliquen

{e} {nirsatdhu} {lgocmbfwkz} {pvjyxq} oder etwas feiner zerlegt
 {e} {n} {irsat} {dhu} {lgocm} {bfwkz} {pv} {jyxq},
 die bei längeren Texten vielleicht noch unterteilt werden kann in
 {e} {n} {ir} {sat} {dhu} {lgo} {cm} {bfwkz} {pv} {jyxq} oder
 {e} {n} {irs} {at} {dhu} {lgc} {om} {bfwkz} {pv} {jyxq}.

Für die englische Sprache bestätigt sich die schon von *L. D. Smith* 1943 angegebene Partition in die Cliquen

{etoani} {rsh} {dl} {ucwmfygpb} {vk} {xjqz} oder etwas feiner zerlegt
 {e} {t} {oani} {rsh} {dl} {ucwm} {fygpb} {vk} {xjqz}, die bei längeren Texten vielleicht
 noch unterteilt werden kann in
 {e} {t} {oa} {ni} {rsh} {dl} {uc} {wm} {f} {ygp} {b} {v} {k} {xjqz}.

Häufigkeiten von Bigrammen im Deutschen und Englischen

Bigramm dt.	Häufigkeit
en	3,88 %
er	3,75 %
ch	2,75 %
te	2,26 %
de	2,00 %
nd	1,99 %
ei	1,88 %
ie	1,79 %
in	1,67 %
es	1,52 %

Bigramm engl.	Häufigkeit
th	3,15 %
he	2,51 %
an	1,72 %
in	1,69 %
er	1,54 %
re	1,48 %
on	1,45 %
es	1,45 %
ti	1,28 %
at	1,24 %

Häufigkeitsverteilung von Trigrammen

Trigramm dt.	Häufigkeit	Trigramm engl.	Häufigkeit
ein	1,22 %	the	3,53 %
ich	1,11 %	ing	1,11 %
nde	0,89 %	and	1,02 %
die	0,87 %	ion	0,75 %
und	0,87 %	tio	0,75 %
der	0,86 %	ent	0,73 %
che	0,75 %	ere	0,69 %

Häufige Viergramme

deutsch	icht, keit, heit, chon, chen, cher, urch, eich, ...
----------------	--

Aufschluß über den Ursprung eines Textes kann auch die mittlere Wortlänge geben:

Sprache	Zeichen
deutsch	5,9
englisch	4,5
französisch	4,4
italienisch	4,5
spanisch	4,4
russisch	6,3

Die zehn häufigsten Wörter:

Sprache	Häufigste Wörter
deutsch	die, der, und, den, am, in, zu, ist, dass, es
englisch	the, of, and, to, a, in, that, it, is, I
französisch	de, il, le, et, que, je, la, ne, on, les
italienisch	la, di, che, il, non, si, le, una, lo, in
spanisch	de, la, el, que, en, no, con, un, se, sa