

Analyse de Fréquence de la Langue Française

Les fréquences d'apparition des différentes lettres en langue française sont données par le tableau suivant

Table 1: Fréquence des lettres en Français en %

Lettre	Fréquence	Lettre	Fréquence
a	8,25	n	7,25
b	1,25	o	5,75
c	3,25	p	3,75
d	3,75	q	1,25
e	17,75	r	7,25
f	1,25	s	8,25
g	1,25	t	7,25
h	1,25	u	6,25
i	7,25	v	1,75
j	0,75	w	0,00
k	0,00	x	0,00
l	5,75	y	0,75
m	3,25	z	0,00

De la même manière on a une fréquence d'apparition des bigrammes (couple de lettres). Les bigrammes qui apparaissent le plus souvent sont donnés dans le tableau ci-dessous.

Table 2: Les bigrammes les plus fréquents (en nombre d'apparition sur 10 000 lettres)

es	305	te	163	ou	118	ec	100	eu	89	ep	82
le	246	se	155	ai	117	ti	98	ur	88	nd	80
en	242	et	143	em	113	ce	98	co	87	ns	79
de	215	el	141	it	112	ed	96	ar	86	pa	78
re	209	qu	134	me	104	ie	94	tr	86	us	76
nt	197	an	30	is	103	ra	92	ue	85	sa	75
on	164	ne	124	la	101	in	90	ta	85	ss	73
er	163										

A noter, que les bigrammes, constitués par deux consonnes, les plus fréquents sont :

nt(197) *tr*(86) *ns*(79) *st*(61).

Ceux constitués de deux voyelles sont :

ou(118) *ai*(117) *ie*(94) *eu*(89) *ue*(85) *ui*(68) *au*(64) *oi*(52) *io*(49)

Les bigrammes formés de deux lettres identiques les plus fréquents sont

<i>ss</i>	<i>ee</i>	<i>ll</i>	<i>tt</i>	<i>nn</i>	<i>mm</i>	<i>rr</i>	<i>pp</i>	<i>ff</i>	<i>cc</i>	<i>aa</i>	<i>uu</i>	<i>ii</i>
73	66	29	24	20	17	16	10	8	3	3	2	1