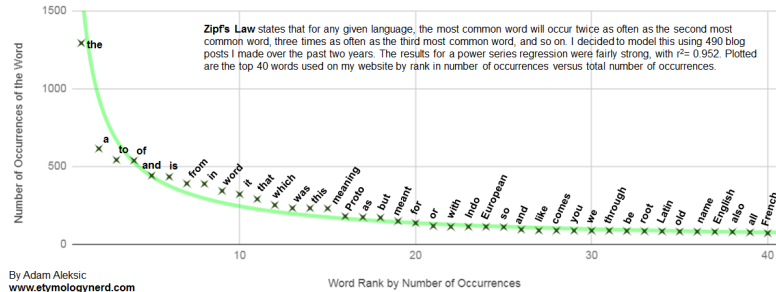


Loi de Zipf

Zipf's Law and My Blog on Language



11 / 29

bataille (35,2)
politique (11,5)
amour (19,94)
voyage (17,44)

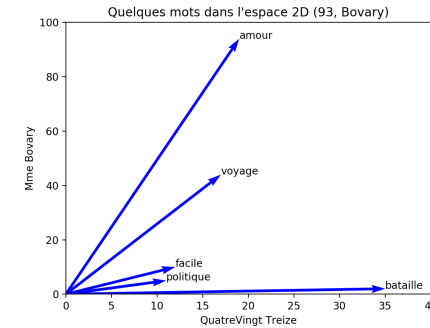


Figure – représentation graphique des mots dans le plan (93, Bovary)

13 / 29

KWIC voyageur

je connaissais le voyageur : Différent du compagnon à un transport / Hébergement Le plus vite votre le monde ... 6ème autres , les grands fév 2008 Ces parmi les meilleurs les 15 principaux ses amies des foyers ...) et des lancé vers des clients chez un des plus grands voyage à cheval . clients chez un sur cette destination hôtel , à votre voyageur qui les amène ou tour-opérateur . En Animaux domestiques : Lorsqu' , légalement , aucune n' FRAM propose les . Le Comité scientifique européen , Kuoni est essayent de prendre , qui opèrent principalement en ligne . Service en ligne en locaux qui fêtent . La commercialisation Le susceptibles d' offrir - Développement d' un européens dont ils spécialiste du voyage - Développement d' un Catalogue destination du , à votre ambassade

frWaC, consultation 2021-08-11

14 / 29

	bataille	voyage	homme	femme
arriver	246	470	1 819	890
tomber	100	83	1 205	384
habiller	2	4	339	384
mourir	180	116	339	1 088
	55 331	208 520	668 289	346 093

Table – Matrice Terme-Terme obtenue dans frWaC. Contextes en ligne.

15 / 29

f	aardvark	computer	data	pinch	result	sugar
apricot	0	0	0	1	0	1
pineapple	0	0	0	1	0	1
digital	0	2	1	0	1	0
information	0	1	6	0	4	0

17 / 29

f	computer	data	pinch	result	sugar	Σ
apricot	0	0	1	0	1	2
pineapple	0	0	1	0	1	2
digital	2	1	0	1	0	4
information	1	6	0	4	0	11
Σ	3	7	2	5	2	19

18 / 29

p_{ij}	computer	data	pinch	result	sugar	p_{a*}
apricot	0/19	0/19	1/19	0/19	1/19	2/19
pineapple	0/19	0/19	1/19	0/19	1/19	2/19
digital	2/19	1/19	0/19	1/19	0/19	4/19
information	1/19	6/19	0/19	4/19	0/19	11/19
p_{*b}	3/19	7/19	2/19	5/19	2/19	19/19

19 / 29

p_{ij}	computer	data	pinch	result	sugar	p_{a*}
apricot	0,00	0,00	0,05	0,00	0,05	0,11
pineapple	0,00	0,00	0,05	0,00	0,05	0,11
digital	0,11	0,05	0,00	0,05	0,00	0,21
information	0,05	0,32	0,00	0,21	0,00	0,58
p_{*b}	0,16	0,37	0,11	0,26	0,11	

20 / 29

<i>ppmi</i>	computer	data	pinch	result	sugar
apricot	-	-	2,25	-	2,25
pineapple	-	-	2,25	-	2,25
digital	1,66	0,00	-	0,00	-
information	0,00	0,57	-	0,47	-

21 / 29

<i>f</i>	computer	data	pinch	result	sugar
apricot	2	2	3	2	3
pineapple	2	2	3	2	3
digital	4	3	2	3	2
information	3	8	2	6	2

22 / 29

<i>p_{ij}</i>	computer	data	pinch	result	sugar	<i>p_{a*}</i>
apricot	0,03	0,03	0,05	0,03	0,05	0,20
pineapple	0,03	0,03	0,05	0,03	0,05	0,20
digital	0,07	0,05	0,03	0,05	0,03	0,24
information	0,05	0,14	0,03	0,10	0,03	0,36
<i>p_{*b}</i>	0,19	0,25	0,17	0,22	0,17	

23 / 29

<i>ppmi</i>	computer	data	pinch	result	sugar
apricot	-	-	2,25	-	2,25
pineapple	-	-	2,25	-	2,25
digital	1,66	0,00	-	0,00	-
information	0,00	0,57	-	0,47	-
<i>ppmi</i> [add2]	computer	data	pinch	result	sugar
apricot	0,00	0,00	0,56	0,00	0,56
pineapple	0,00	0,00	0,56	0,00	0,56
digital	0,62	0,00	0,00	0,00	0,00
information	0,00	0,58	0,00	0,37	0,00

Figure – Comparaison de mesures de *ppmi* sur l'exemple de (Jurafsky & Martin, 2019), avec ou sans lissage laplacien +2

24 / 29