



Le génie pour l'industrie

Département de génie logiciel et des TI

## Devoir 4

*Traitement des langues naturelles*

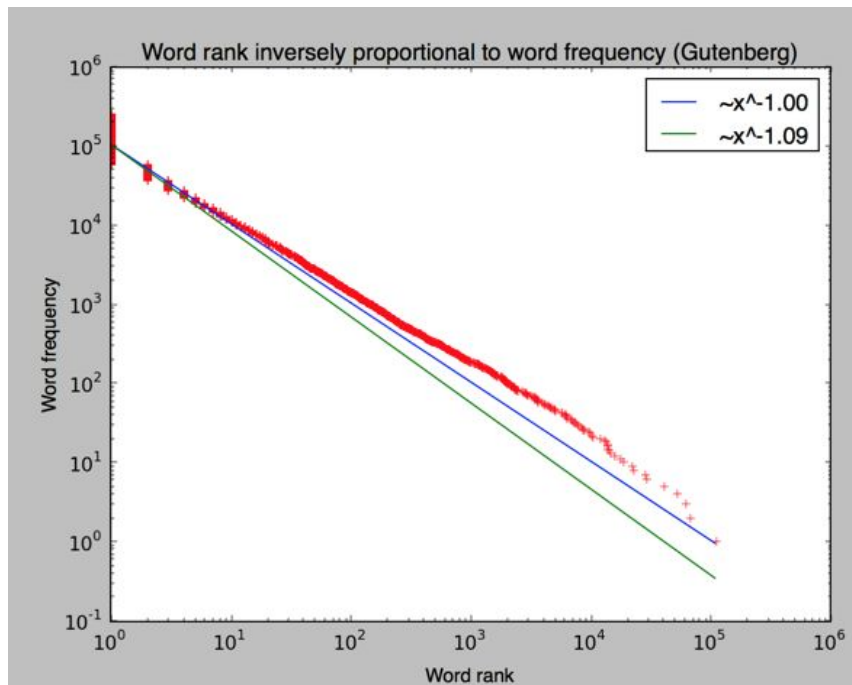
<b>No du devoir</b>	4
<b>Étudiant(s)</b>	Dominique JACQUES-BRISSETTE
<b>Code(s) permanent(s)</b>	JACD16069105
<b>Cours</b>	MTI830
<b>Session</b>	Été 2015
<b>Professeur</b>	Sylvie Ratté
<b>Date</b>	10 juin 2015

## A- Graphiques de la loi Zipf

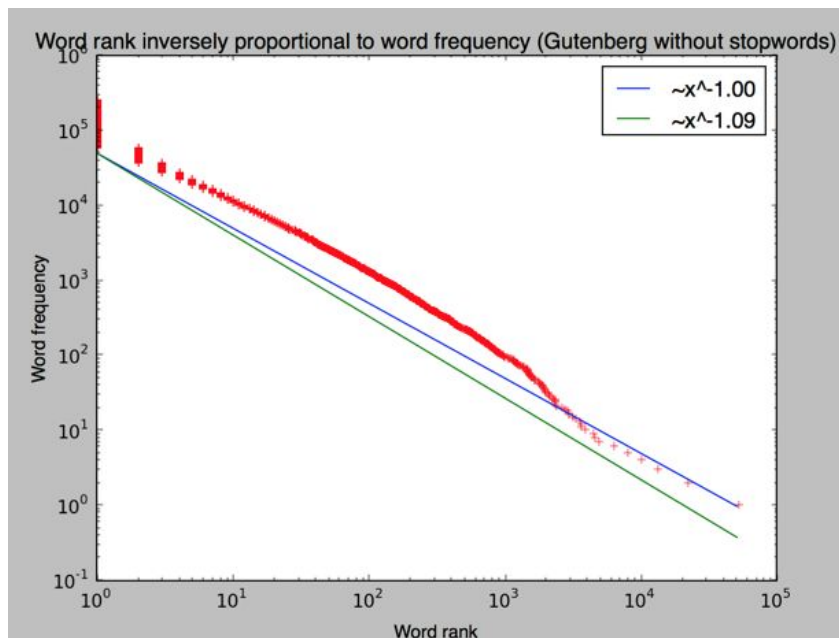
Dans tous les cas, à l'aide des graphiques on peut confirmer visuellement que le “rank” est inversement proportionnel à la fréquence des mots.

### Gutenberg

Avec “stopwords”

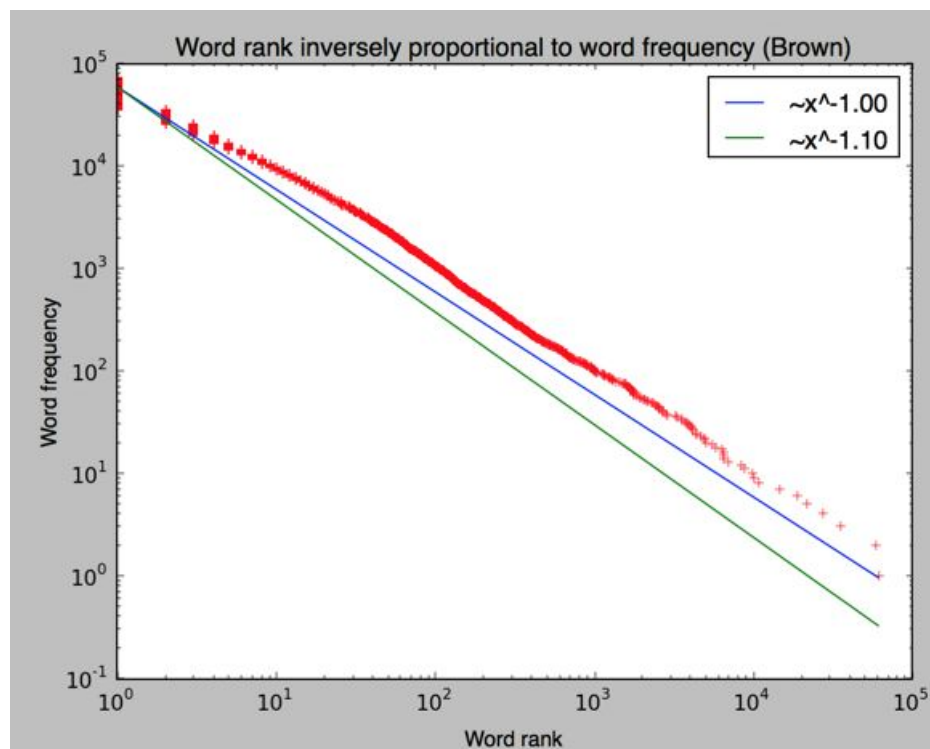


Sans “stopwords”

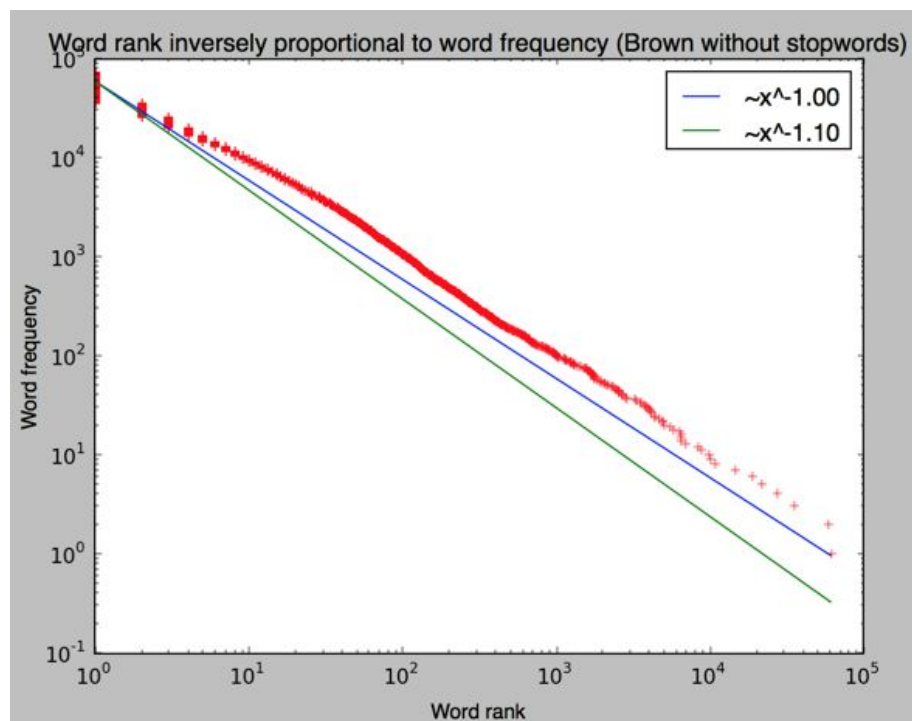


## Brown

Avec “stopwords”



Sans “stopwords”



## B- Mauvaise analyse des rôles sémantiques.

Voici la phrase en anglais choisie :

The complex houses married and single soldiers and their families .

Où “complex” n’est pas un adjectif mais le nom “complexe” comme synonyme de bâtiment, et “houses” ne veut pas dire “maisons” mais bien “héberge”. Une traduction simple en français du message à transmettre pourrait ressembler à ceci : “Le complexe héberge des soldats mariés et célibataires ainsi que leur familles”.

Comme le site de l’université d’Illinois ne fonctionne pas

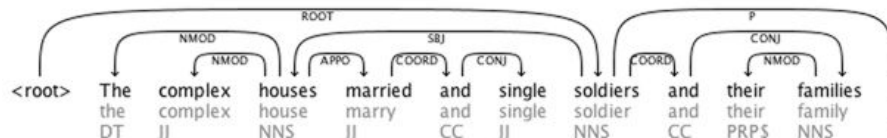
([http://cogcomp.cs.illinois.edu/page/demo\\_view/srl](http://cogcomp.cs.illinois.edu/page/demo_view/srl) ne retourne aucun résultat.)

j’ai utilisé le site suivant : <http://barbar.cs.lth.se:8081/parse>

Voici le résultat :

	The	complex	houses	married	and	single	soldiers	and	their	families	.
family.01									A2		

Parsing sentence required 14ms.



ID	Form	Lemma	PLemma	POS	PPOS	Feats	PFeats	Head	PHead	Deprel	PDeprel	IsPred	Pred	Args: family.01
1	The	the	the	DT				3			NMOD			
2	complex	complex	complex	JJ				3			NMOD			
3	houses	house	house	NNS				7			SBJ			
4	married	marry	marry	JJ				3			APPO			
5	and	and	and	CC				4			COORD			
6	single	single	single	JJ				5			CONJ			
7	soldiers	soldier	soldier	NNS				0			ROOT			
8	and	and	and	CC				7			COORD			
9	their	their	their	PRP\$				10			NMOD			A2
10	families	family	family	NNS				8			CONJ	Y	family.01	
11	.	.	.	.				7			P			

Comme on peut voir, l’outil de traitement ne trouve aucun verbe principal de la phrase (en fait, il est incapable de trouver un verbe tout court). Le verbe dans ce cas-ci est “houses”, mais il l’identifie comme étant un nom.