

Cours : Ingénierie de la langue

Auteur : Kouamé Kouassi

**Rapport d'étude : L'usage du verlan dans un
corpus de tweets**

Date : 30/04/2020

Professeurs : Gaël Lejeune, Yoann Dupont



Lettres Sorbonne Université

Sommaire

1. Présentation du sujet.....	3
2. Présentation du corpus.....	3
3. Présentation des outils.....	4
4. Définition des termes.....	4
A. Corpus.....	4
B. Unité lexicale.....	4
C. Fréquence.....	4
D. Indice.....	4
E. Cooccurrences dans TXM.....	5
F. Requête CQL dans TXM.....	5
G. Concordances.....	5
H. N-Gramme dans Antconc.....	5
I. Key Word In Context d'Antconc (KWIC).....	5
5. Analyse des données.....	5
a. Verlan 1 : <i>teuf</i> > <i>fête</i>	5
b. Verlan 2 : <i>meuf</i> > <i>femme</i>	6
c. Verlan 3 : <i>goléri</i> > <i>rigoler</i>	8
d. Verlan 4 : <i>keuf</i> > <i>policier</i>	9
e. Verlan 5 : <i>reuss</i> > <i>sœur</i>	10
6. Conclusion.....	11

1. Présentation du sujet

Le sujet choisi concerne l'usage du verlan dans les tweets. Le domaine du langage argotique est une sphère très intéressante du point de vue linguistique parce qu'on y observe une certaine dynamique linguistique. C'est-à-dire que les éléments du langage argotique évoluent, changent indépendamment des normes rigides de la grammaire. Cette évolution est très perceptible dans l'usage des verlans à l'écrit comme à l'oral dans la société française, plus précisément chez les jeunes, où chacun use du mot verlanisé comme il l'entend.

Qu'est-ce que donc le verlan ? Le verlan est un type particulier d'argot qui consiste au codage lexical par inversion des syllabes, insertion de syllabes postiches, suffixation et infixation systématique.

L'analyse des usages de différents mots verlanisés de la langue française permettra de faire ressortir un certain nombre de règles implicites répétitives auxquelles obéissent ces mots.

Avec les outils informatiques dont nous disposons, nous serons en mesure d'observer par exemple quels sont les éléments qui coexistent avec les verlans et voir pour chaque expression verlanisée choisie, quels éléments les accompagnent la plupart du temps à chaque usage.

2. Présentation du corpus

L'analyse est menée sur un corpus de tweets recensé automatiquement mais aussi à la main pour les verlans choisis sur lesquels le travail doit être mené. Avec TXM, le nombre d'unités lexicales (mots) est $t = 280325$, un nombre de valeurs $v = 21556$, une fréquence minimum $f_{min} = 1$, une fréquence maximale $f_{max} = 25589$. Nous avons 4 propriétés de mots qui sont : word, flemma, fpos et lbn. Au total, le nombre de corpus de tweets agrégé est de 4.

Les mots verlanisés qui sont analysés sont au nombre de 5 :

- teuf > fête
- keuf > policier
- goleri > rigoler
- meuf > femme
- reuss > sœur

Il y a bien évidemment un nombre conséquent de verlans, mais ceux-ci ont été choisis du fait de leur popularité et de leur utilisation, car ils sont plus utilisés que d'autres ce qui facilite leur observation.

Le corpus est en format texte et la langue est le français. Les personnes à l'origine de ces tweets sont bien-sûr des utilisateurs du réseau social Twitter pris au hasard par la section recherche avancée du réseau mais aussi ciblés.

Toutefois, le constat est que la récupération automatique des tweets sur des utilisateurs ciblée s'avère moins précise que celle faite à la main parce que la plupart des mots verlanisés sur lesquels travailler n'y figurent presque pas exceptés 1 ou 2 mots, tandis que la recherche du mot « teuf » par exemple fait apparaître tous les tweets contenant ce mot, donc est plus précis. Toutefois, ces corpus récupérés automatiquement augmentent dans une moindre mesure le nombre d'occurrences de verlan de notre liste.

3. Présentation des outils

Les outils utilisés pour l'analyse du corpus sont Antconc et TXM. Ces outils sont adaptés aux études statistiques sur les textes¹ dont les fonctionnalités vont nous permettre d'aboutir à nos analyses comme l'usage des expressions rationnelles avec TXM parce qu'il possède un moteur d'expression rationnelle² et de l'Analyse Factorielle des Correspondances et encore bien d'autres.

Antconc a été surtout utilisé pour l'analyse en contexte gauche et droit par la fonctionnalité KWIC Sort (Key Word In Context), où des verlan ont été triés selon les 2, 3, 4 mots avant ou après le mot lui-même pour observer les parties du discours qui l'entourent le plus. L'outil TXM a fait la plus grande partie du travail avec :

- L'analyse des cooccurrences
- La fréquence
- Les verbes les plus utilisés pour les verlan choisis du corpus
- Le temps des verbes
- Les spécificités de chaque verlan

4. Définition des termes

A. Corpus

Un corpus est une collection de données langagières sélectionnées et organisées selon des critères linguistiques explicites³.

B. Unité lexicale

C'est lorsque chaque texte est composé de mots pouvant avoir des propriétés comme la catégorie grammaticale.

C. Fréquence

C'est l'effectif d'une unité lexicale dans le corpus.

D. Indice

Lorsque l'on fait une requête CQL dans TXM, les cooccurrents qui apparaissent sont calculés en fonction de l'indice de spécificité qui est un indicateur statistique de présence.

1 Textométrie : <http://textometrie.ens-lyon.fr/>

2 Source Wikipedia : https://fr.wikipedia.org/wiki/Expression_régulière

3 Habert, Nazarenko *et* Salem, 1997.

E. Cooccurrences dans TXM

Les cooccurrents sont des éléments lexicaux situés autour d'un motif lexical qu'on a recherché par une requête CQL.

F. Requête CQL dans TXM

Le Corpus Query Language est un langage d'expression de requêtes permettant de rechercher un motif linguistique qui peut être un mot ou une suite de mot.

G. Concordances

Cette commande construit une concordance kwic à partir des résultats de recherche correspondant à une requête CQL sur un corpus ou un sous-corpus sélectionné⁴.

H. N-Gramme dans Antconc

Fonctionnalité qui permet par exemple de donner la probabilité qu'un mot apparaissent après un autre mot.

I. Key Word In Context d'Antconc (KWIC)

Analyse d'un mot en contexte par tri selon le contexte gauche ou droit.

5. Analyse des données

a. Verlan 1 : *teuf* > *fête*

Le nombre d'unités lexicales est $t = 27$ pour une fréquence maximale $f_{\max} = 27$.

Les cooccurrents de ce verlan triés selon le plus grand indice (18) sont :

- **teuf** : fréquence = 27, indice = 18
- **grosse** : 2e place, fréquence = 43, indice = 12. [frpos="ADJ"][word="teuf"]
- **mega** : 3e place, fréquence = 8, indice = 9. [frpos="NOM"][word="teuf"]
- **la** : 4e place, fréquence = 1875, indice = 8. [frpos="DET:ART"][word="teuf"]
- **une** : 5e place, fréquence = 797, indice = 7. [frpos="DET:ART"][word="teuf"]

Elite. C'est la teuf Ha ouai	grosse teuf	alcool chichon tous à volonté ...
fois. WoW Eh venez on orga ...	grosse teuf	à la fin de ce fucking confine...
que tout soit finie pour faire un...	grosse teuf	des familles, des potes, du so...
merde est finis, on fait la plus	grosse teuf	jamais crée pour fête ça, oblig...
la fin du confinement on fait la...	grosse teuf	du siècle genre tous ensembl...
un virus, j'ai organisé une super	grosse teuf	avec tous mes potes à la mai...

Cooccurrent "grosse" de Teuf - TXM

4 Manuel TXM : <http://textometrie.ens-lyon.fr/files/documentation/Manuel de TXM 0.7 FR.pdf>

On peut voir aussi qu'avec le KWIC Sort de Antconc, « teuf » est parfois précédé de la préposition « en » selon 1 mot trié à gauche, [frpos="PRP"][word="teuf"] dans TXM. « Grosse », « la », « le », et « mega » sont toujours dans le contexte gauche de ce mot. Pour les cooccurrents droits, les mots qui se répètent plus d'une fois sont plus des prépositions : « après », « avec », « à », [word="teuf"][frpos="PRP"] dans TXM.

Le verbe le plus utilisé selon le KWIC Sort est « faire » qui apparaît toujours dans le contexte gauche et selon au moins 4 mots triés à gauche. Il est soit utilisé à l'infinitif, conjugué au présent et au futur. Cela peut supposer que « teuf » est utilisé comme une **locution verbale** « *faire la teuf* » ou « *faire une teuf* » comme en français *faire la fête*.

b. Verlan 2 : meuf > femme

Le nombre d'unités lexicales dans TXM est $t = 95$ pour une fréquence maximale $f_{\max} = 95$. Les cooccurrents de ce verlan sont triés selon le plus grand indice (25) et ayant une fréquence élevée et pertinents :

- 1^{er} : **une**, fréquence = 797, indice 27. [word="une"][word="meuf"]
- 2^e : **a**, fréquence = 1261, indice 12. [word="a"][frpos="DET:ART"][word="meuf"]
- 3^e : **elle**, fréquence = 498, indice 11. [word="meuf"][word="elle"]. Il n'y a pas que cet index qu'on peut formuler mais aussi [word="elle"][word="m"][frpos="VER:pres"][word="ptdr"] [frpos="DET:ART"][word="meuf"]. Cela fera long si toutes les occurrences de « elle » doivent être recherchées.
- 4^e : **qui**, fréquence = 812, indice = 8. [word="meuf"][word="qui"]
- 5^e : **la**, fréquence = 1875, indice = 5. [word="la"][word="meuf"]
- 6^e : **ma**, fréquence = 465, indice = 4. [word="ma"][word="meuf"]

Avec Antconc, le constat le plus évident selon 1 mot trié à gauche, le verlan « meuf » est généralement précédé d'un déterminant, plus précisément d'un article défini « la » et indéfini « une ». Cela est dû au fait que dans la plupart des tweets qui contiennent « meuf », la « meuf » dont on parle est connue de la personne qui s'exprime. Ce qui explique l'usage de **l'article défini**.

Par contre pour l'usage de l'article indéfini « une », la « meuf » dont on parle n'est pas forcément une « meuf » qu'on ne connaît pas. Donc si l'on doit généraliser, on peut dire que ce mot est aussi utilisé avec un déterminant la plupart du temps. Mais pas que ; car on remarque aussi

```
regarder l'équipe Elle me fait trop goleri la meuf qui traduit elle est grave impliquée Les lives in
. Restez chez vous. RESTEZ CHEZ VOUS. Vraiment la meuf la plus relou de TikTok je peux pas mla
buzzer pcq y'a une histoire de viol. la meuf à aucun cas parler de viol hein idiot</tweet> <
eet> RT @riodejaneiro69: @Ancass_Elle a gechan la meuf a rami</tweet> <tweet> RT @bndrs8: @Ancass_Ahhh
3h48 et que jme rapelle du tweet de la meuf qui a dit " les meufs qui se maquille la
que dans la version blanche de l'image la meuf soit...</tweet> <tweet> RT @Many_Waves_: Tony dans
????? 1 MILLIARDS tu le tue et apres tu fais la meuf en deuil tchiiip concentre toi</tweet> <tweet>
autant manq...</tweet> <tweet> RT @badgaloumi: La meuf qui avait tweeker: « Parfois Dieu met un garçon d
.co/hQknAXo3dS</tweet> <tweet> RT @Locamami: La meuf n'a jamais été vilaine</tweet> <tweet> RT @Chris
re</tweet> <tweet> RT @_laevd: wesh wesh wesh la meuf elle twerk sans jambes ou je rêve ?!</tweet> <t
au...</tweet> <tweet> RT @babyfr3sh4: «Tu vois la meuf derrière? Soit discrète quand tu regarde » Moi:
être comme ça</tweet> <tweet> C'est elle la meuf moins chère. Elle mérite que des grandes baffes d
> RT @majestueuse: Moi je me demande c ki la meuf qui a dmd a mamadi d'envoyer son tdc 8'97</
co/BkNLxBbFtl</tweet> <tweet> RT @badgaloumi: La meuf qui avait tweeker: « Parfois Dieu met un garçon d
, il va se vexer jusqu'à faire culpabiliser la meuf alors qu'il sai...</tweet> <tweet> RT @Bakabraa:
s surtout au niveau sentimental quand il kiffe la meuf la mddr</tweet> <tweet> @Mas_sadji En sah j'
GbPjZNG</tweet> <tweet> Elle me fait penser à la meuf qui a fait la vidéo ou à la fin
ie ...</tweet> <tweet> @FahiimUchiwa78 Wallah la meuf elle a fait une chute de 1m50 en mode
```

dans une moindre mesure l'usage de **déterminants possessifs** « ma », « sa », « ta ». En syntaxe, on peut voir que ce mot est parfois en position de complément d'objet direct ou indirect.

Ce même résultat sur les déterminants peut être obtenu avec la fonctionnalité **Concordances** de TXM avec un tri selon le contexte gauche. « DET:ART » pour déterminant de type article et « DET:POS » pour déterminant de type possessif.

Contexte gauche	Pivot
NOM la_DET:ART	meuf
res une_DET:ART	meuf
NOM la_DET:ART	meuf
RP une_DET:ART	meuf
ADV les_DET:ART	meuf
NOM la_DET:ART	meuf
ADV la_DET:ART	meuf
RP ma_DET:POS	meuf
per une_DET:ART	meuf
SENT la_DET:ART	meuf
mp une_DET:ART	meuf

Parties du discours
précédant "meuf" -
TXM

Pour le contexte droit, le pronom relatif « qui » est susceptible dans la plupart du temps de suivre « meuf » est aussi « que ». Ces **marqueurs de subordination** sont des relatives dont le rôle syntaxique est une **reprise anaphorique** de « meuf ». « DET : REL » pour pronom relatif. La fonctionnalité utilisée ici pour ces deux affichages est l'option d'affichage du contexte de TXM pour voir les frpos c'est-à-dire les parties du discours.

meuf	qui_PRO:REL
meuf	qui_PRO:REL
meuf	qui_PRO:REL
meuf	qui_PRO:REL
meuf	qui_PRO:REL
meuf	qui_PRO:REL
meuf	qui_PRO:REL
meuf	qui_PRO:REL
meuf	qui_PRO:REL
meuf	qui_PRO:REL

Contexte droit de
"meuf" - TXM

Et enfin, le verbe qui apparaît plus d'une fois avant le mot est le verbe « avoir » au présent.

c. Verlan 3 : *goleri* > *rigoler*

Les mots sont toujours triés selon leur pertinence et ayant un indice et une fréquence élevé ; l'indice parce que c'est un élément important. Plus l'indice est haut, plus il permet d'étudier de manière efficiente les contextes du verlan. Moins il est élevé, plus il est loin du contexte « immédiat », tout près du mot. Il peut y avoir une fréquence élevée d'un cooccurrent et un indice bas. Mais tout est jaugé, analysé de sorte que l'analyse puisse se faire.

Pour ce verlan on a donc un nombre d'unités lexicales $t = 38$ avec un plus grand indice = 30 (le verlan même). Les cooccurrents sont :

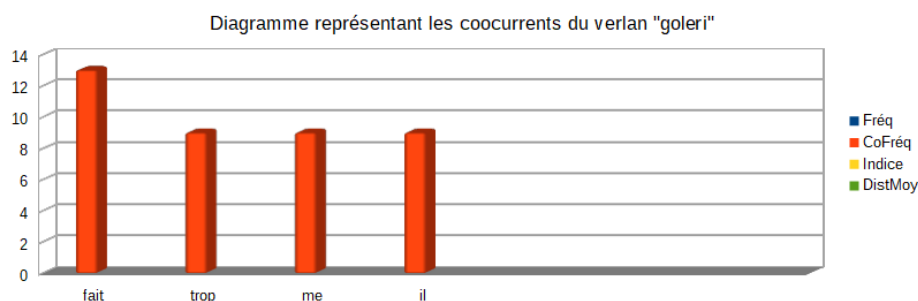
- **1er : fait, fréquence = 490, indice = 10.** [frlemma="faire"][word="goleri"]
- **2e : trop, fréquence = 505, indice = 6.** [frpos="ADV"][word="goleri"]. Là apparaissent l'adverbe « trop » et des hapax « bien goleri » et « pas goleri ».
- **3e : me, fréquence = 745, indice = 4.** [word="me"][word="fait"][word="trop"]
[word="goleri"]
- **4e : il, fréquence = 933, indice = 3.**

Avec le premier cooccurrent « fait », le verlan « goleri » fonctionne toujours comme une locution verbale « faire goleri » pour faire rigoler au présent de l'indicatif. Le verbe « aller » peut être utilisé un peu comme une périphrase verbale⁵ aussi, mais est rare parce qu'il apparaît comme un hapax dans le corpus avec un effectif de 1 même avec ses lemmes « vais » goleri et « vas » goleri. Entre ces deux mots peut s'insérer un adverbe ou un autre type de mot. L'adverbe en question est « trop » comme faire trop goleri.

Le tableau et le diagramme ci-dessous illustrent la prépondérance de « faire » :

	A	B	C	D	E
1	<u>Occ</u>	<u>Fréq</u>	<u>CoFréq</u>	<u>Indice</u>	<u>DistMoy</u>
2	fait		13	10.7144	2.0769231
3	trop		9	6.1607	3.0
4	me		9	4.8037	1.7777777
5	il		9	4.0549	3.5555553

La fréquence n'apparaît pas parce qu'étant donné que certains cooccurrents ont des fréquences élevées et des indices bas, il y a des incohérences lorsque le diagramme est établi. C'est-à-dire qu'au lieu que « faire » supprime les autres cooccurrents, il est tout en bas. Sans la fréquence cela donne donc :



5 Le verbe aller n'a pas son sens plein ici. C'est une subduction ou évidemment sémantique qui vient du linguiste G.Guillaume.

On peut donc voir les cooccurrents qui sont plus dans le contexte du mot que d'autres. C'est en cela que la relève des indices élevés est importante.

Ensuite on a le pronom personnel réflexif « me » qui occupe une place importante parmi les cooccurrents. Comme tout pronom réfléchi, il est parfois en combinaison avec « faire » conjugué > me fait goleri ou me fait trop goleri. Le KWIC Sort d'Antconc est encore utile pour bien visionner :

```
on daron il m'a fait goleri,  
Ça me faisait grave goleri  
rahim il me fait trop goleri  
iipe Elle me fait trop goleri  
dsl elle me fait trop goleri  
vrmt ça me fait bien goleri
```

Plus pour terminer « il » qui est souvent dans le contexte pour parler de quelque chose ou de quelqu'un qui fait rire. Ce qui veut dire que c'est peut être pareil pour le pronom personnel « elle » mais qui n'apparaît pas dans le corpus.

d. Verlan 4 : *keuf* > *policier*

Pour ce verlan, nous avons un nombre d'unités lexicales $t = 46$ pour une fréquence maximale $f_{\max} = 46$. Le plus haut indice est égal à 34, il s'agit du verlan lui-même mais qui ne sera pas pris en compte. Ainsi nous avons :

- 1^{er} : **les**, fréquence = 1396, indice = 9. [word="les"][word="keuf"]
- 2^e : **par**, fréquence = 144, indice = 8. [word="par"][word="les"][word="keuf"]
- 3^e : **courser**, fréquence = 3, indice = 7. Bien que ce cooccurrent ait une fréquence faible, il est néanmoins capital de le relever à cause de son indice. Son usage ici sera bien évidemment expliqué. [word="courser"][word="par"][frpos="DET:ART"][word="keuf"]
- 4^e : **le**, fréquence = 1656, indice = 5. [word="le"][word="keuf"]
- 5^e : **dans**, fréquence = 528, indice = 4. L'indice ne permet pas de donner une requête CQL assez exacte pour retrouver ce cooccurrent. Cela veut dire qu'il est loin du verlan en question. Néanmoins la fréquence a été prise en compte parce que c'est un élément important qui sera expliqué.
- 6^e : **un**, fréquence 1173, indice = 4. [frpos="DET:ART"][word="keuf"]

Avec Antconc, on peut voir que comme le verlan « meuf », « keuf » s'utilise généralement avec un **déterminant**, plus précisément encore un **article défini ou indéfini**. Mais il s'utilise plus au pluriel, *les keuf*. Le contexte droit est un peu hétéroclite parce qu'il mélange un peu de tout des parties du discours qui ne se répètent pas. Il est donc difficile d'observer une **constante**, c'est-à-dire des éléments qui vont au moins se répéter plus d'une fois. Par contre dans le contexte gauche, on peut voir que l'adverbe ou la préposition « **par** » est répété 4 fois. Ce qui permet de dire que

peut-être que **keuf** est parfois précédé de « par » pour signifier qu'un policier fait ou a fait une action envers celui qui s'exprime.

Donc « keuf » est parfois en position **d'agent**. Mais cela reste à vérifier étant donné qu'il n'y a pas assez d'effectif dans le corpus, seulement 3 et parce que l'on sait qu'à la forme passive le verbe utilisé est toujours l'auxiliaire *être*, ce qui n'est pas le cas ici. Cela amène directement au cooccurrent « **courser** » qui a une basse fréquence mais un indice élevé. Ce cooccurrent veut dire qu'on est poursuivi par un policier. Même à l'oral, cela s'entend chez les jeunes, seulement que ce cooccurrent n'apparaît pas énormément dans ce corpus.

Illustration :

se faire courser par les keuf??
 'me fait courser par les keuf Di
 ont sfaire crever par les keuf ya
 l se fait menotté par les keuf Le
 is faites piétiner par les keuf m
 e faisais course par une keuf. Et

Contexte gauche du verlan
 "keuf" - Antconc

Les soulignés en rouge pour préciser que le verbe utilisé n'est pas l'auxiliaire être.

Enfin pour le cooccurrent « dans » qui n'est pas dans le contexte immédiat, il a été relevé parce qu'il va de pair avec *courser*. En effet, le sens qui veut être fait passé dans les tweets est qu'on est *coursé*⁶ ou poursuivi dehors, dans la rue, ou dans un quartier. Donc c'est une **préposition de lieu**.

e. Verlan 5 : *reuss* > *sœur*

Pour ce dernier verlan, nous avons un nombre d'unités lexicales $t = 39$ et une fréquence maximale $f_{\max} = 39$. Le plus indice est 50. Les cooccurrents sont :

- 1^e : *ma*, fréquence = 465, indice = 23. [word="ma"][word="reuss"]
- 2^e : *la*, fréquence = 1875, indice = 12. [word="la"][word="reuss"]
- 3^e : *petite*, fréquence = 47, indice = 5. [word="petite"][word="reuss"]
- 4^e : *sa*, fréquence = 201, indice = 4. [word="sa"][word="reuss"]

Ce que l'on peut constater avec Antconc dans le contexte gauche avec le Kwic Sort toujours, est l'usage de déterminant possessif avec ce verlan = *ma*, *mes*, *sa*. Il y'a l'article indéfini

6 Dictionnaire de la zone : Poursuivre, prendre en chasse.
<https://www.dictionnairedelazone.fr/dictionary/definition/courser>

« la », qui précède souvent et l'adjectif « petite » pour signifier qu'on parle d'une *petite sœur*. Tous ces éléments font savoir qu'il s'agit d'un mot de genre **féminin**.

défend la reuss moi j'valide fr
 o tune ma reuss on dirait Wall
 v cool ma reuss on se connaît
 i bise mes reuss. On dirai ma r
 lit jsuis sa reuss ptdr ma reuss
 s avec ma reuss quand j'étais
 oir c'est la reuss qui régale ! El
 force ma reuss Romeo Elvis e
 e ein mes reuss sachez le J'ai
 ns bvvv la reuss tu gères Si j'è
 e tiktok, la reuss veut commen
 ercher sa reuss à la maternell
 et ta ptite reuss ça va te faire t

Contexte gauche de
 "reuss" - Antconc

L'identification des cooccurrents du contexte droit un peu difficile, mais avec la fonctionnalité **N-gramme** de Antconc. Cette fonctionnalité a été utilisée selon un **tri par la probabilité** qu'à le deuxième mot d'apparaître dans le premier. Le résultat qui apparaît donc est *reuss elle*. Il y a donc peut être plus de chances que « elle » suive ce verlan d'après les résultats :

Concordance		Concordance Plot	File View	Clusters/N-Grams
Total No. of Cluster Types		30	Total No. of Clus	
Rank	Freq	Range	Prob	Cluster
1	4	1	0.103	reuss elle
2	3	1	0.077	reuss hein
3	2	1	0.051	reuss en
4	2	1	0.051	reuss j
5	2	1	0.051	reuss la
6	2	1	0.051	reuss on
7	1	1	0.026	reuss ah
8	1	1	0.026	reuss alors

On peut voir que la probabilité que « elle » vienne après « reuss » est de 0.103 % et la plus haute.

6. Conclusion

Les analyses qui ont pu être menées sur ce corpus de tweets 5 verlanes ne sont peut être pas assez, mais ont néanmoins permis de voir comment s'emploient à l'écrit ces mots. L'analyse des cooccurrents a permis de faire ressortir tout un tas de « règles » qui régissent ces verlanes qu'on ne peut pas déceler sans une analyse avec ce type d'outil en ingénierie des langues. Cela a permis de voir quelles parties du discours sont toujours dans le contexte du verlan *keuf* par exemple, dont on sait maintenant qu'il souvent précédé d'un article indéfini au pluriel. On peut rapprocher certains verlanes de ce corpus comme *teuf*, *meuf* et *reuss* parce que ces trois mots ont un cooccurrent qui est

presque toujours présent : l'article défini ; ***la teuf***, ***la meuf***, ***la reuss*** et ***le keuf***. Le verlan *goleri* quant à lui s'emploie en tant que locution verbale avec le verbe « faire ».

Les analyses auraient pu être enrichies si certaines fonctionnalités de TXM avaient été utilisées comme la partition, le sous-corpus et l'AFC. Cela est dû au fait que le corpus n'est parfait bien que le corpus soit structuré en XML.