“ Nous nous arrachâmes promptement avec ma caisse ” : quels descripteurs linguistiques caractérisent les registres de langue ?

Jade Mekki, Delphine Battistelli, Nicolas Béchet, Gwénolé Lecorvé

▶ **To cite this version:**

Jade Mekki, Delphine Battistelli, Nicolas Béchet, Gwénolé Lecorvé. “ Nous nous arrachâmes promptement avec ma caisse ” : quels descripteurs linguistiques caractérisent les registres de langue ?. [Rapport Technique] IRISA, équipe EXPRESSION; MoDyCo. 2017. hal-01649948

HAL Id: hal-01649948
https://inria.hal.science/hal-01649948

Submitted on 28 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Avril – Juillet 17

« Nous nous arrachâmes promptement avec ma caisse » : quels descripteurs linguistiques caractérisent les registres de langue ?

Jade Mekki[1,2]

Encadrée par Delphine Battistelli[2], Nicolas Béchet[3] et Gwénolé Lecorvé[1]

[1] IRISA/Université de Rennes 1
[2] MODYCO CNRS/Université Paris-Nanterre
[3] IRISA/Université Bretagne Sud

Projet ANR TREMoLo (ANR-16-CE23-0019) : transformation de registres par extraction de motifs langagiers.

INTRODUCTION 3

I – IDENTIFICATION DES DESCRIPTEURS 6

II – TEST DES DESCRIPTEURS 14

A – Réflexion et composition du corpus **14**

B – Préparation du corpus **15**

C – Test des descripteurs **18**
1 – Approche lexicale 19
1.1 – onomatopées : familier 19
1.2 – sur présence de « et » : familier 20
1.3 – Sur présence de « là » - ponctuant : familier 20
2 – Approche grammaticale 22
2.1 – contraction (c'est → cela est) : familier / standard 22
2.2 – contraction (ça → cela) : familier / standard 23
2.3 - contraction de la forme négative : ne/pas, pas : standard / familier 23
2.4 – Sujet « nous » transposé en « nous » 24
2.5 – Termes à redoublement 25
2.6 – Terminaisons en « asse » 27
2.7 – Terminaisons en « ouze » 32
2.8 – Terminaisons en « o » 33
2.9 – Terminaisons en « ou » 34
2.10 – Terminaisons en « iotte » 36
2.11 – Adverbes dérivés d'un nom 37
2.12 – syntagme : c'est + PL 38
2.13 – ça + VB 40
2.14 – C'est… qui… que 46
2.15 – Des ADJ 50
2.16 – « Et puis » en début de phrase 51
2.17 – Verbe du premier groupe majoritaire pour le registre familier 52
2.18 - Absence de l'inversion SUJ/VB lors d'une interrogative 53
2.19 – qu(i) voyelle 59
2.20 – Utilisation de « vraiment » : surreprésenté / familier 60
2.21 – et (qui | que) 61
2.22 – ou quoi 62
2. 23 – et même que 63
2.24 – Des fois / parfois 63
2.25 – (sa | son) x à (elle | lui) 65

III – PROBLEMES RENCONTRES ET PERSPECTIVES DE TRAVAUX ULTERIEURS 68

A – Difficultés soulevés par les travaux **68**

B - Travaux sur les discours rapportés **69**

CONCLUSION 76

Introduction

Le projet TREMoLo travaille sur la notion de registre de langue. Chaque production linguistique est évaluée par l'interlocuteur. Il la caractérise en la situant dans une classe, un registre. Ce dernier permet donc de qualifier une certaine actualisation de la langue, ce qui est à distinguer de la variation linguistique. La variation linguistique est un concept de la sociolinguistique qui va à l'encontre du structuralisme[1] (où le locuteur n'a qu'une seule manière de dire ce qu'il veut dire) puisqu'elle envisage la langue comme fondamentalement hétérogène. Françoise Gadet propose deux typologies de variation[2] :

Selon les usagers :

- Diachronique : temporelle
- Diatopique : spatiale (ex : dialecte)
- Diastratique : sociale (ex : sociolecte, technolecte…)

Selon l'usage :

- Diaphasique : une même personne, quelle que soit son origine sociale, parle différemment selon la situation de communication.
- Diamésique : variation d'usage entre oral et écrit

De plus, la variation agit à tous les niveaux de la langue :

- Phonique
- Morphologique
- Lexicale
- Syntaxique

Ainsi, les registres constituent des étiquettes auxquelles nous associons des productions linguistiques évaluées comme telles. Quant à la variation, c'est un concept dynamique plus large qui permet de saisir les écarts entre différentes manières de dire ce qui veut être exprimé. Des différents types de variations le projet TREMoLo se concentre sur la variation diaphasique et diamésique.

Pour cela, nous avons constitué un corpus :

- *Albertine Disparue* de Proust
  - Registre soutenu
- L'*humanité* – archives
  - Registre courant / standard
- *Kiffe Kiffe Demain* de Guene
  - Registre familier
- Histoires à parler
  - Registre enfantin

La finalité de ce projet est d'identifier automatiquement un registre puis de générer des paraphrases qui le traduisent vers un registre différent. De fait, les « différents registres » classent un énoncé par rapport à une norme implicite : le français « standard ».

[1] SAUSSURE, *Cours de linguistique générale* 1916
[2] GADET, *La Variation Sociale en Français* 2007

Aussi, un locuteur évalue une production linguistique qui, suivant sa variation, s'éloigne plus ou moins de cette norme selon une polarité allant du vulgaire au soutenu. Ce jugement l'amène à attribuer un registre (parfois également appelé « niveau de langue ») à l'énoncé produit. Les enjeux liés à la dénomination même de « registre » ou bien « niveaux de langues » ne seront pas discutés bien que leurs problématiques sociolinguistiques soient avérées. La question éthique associée à la notion de « norme » ou de « bonne formation » ainsi que le jugement axiologique qui les accompagne ne sont pas les sujets de cette étude. En effet, l'objectif est de pouvoir identifier des descripteurs, c'est-à-dire des éléments linguistiques discriminants pour chaque registre, puis de les confirmer en les testant un par un. Mon travail se compose de trois grands temps : un état de l'art théorique, l'identification des descripteurs, test de ces derniers.

La perception d'un registre passe par l'évaluation de la façon de parler d'un locuteur. Si le sujet de notre étude n'est pas de discuter des ressorts qui associent telle production avec telle connotation, il existe de fait une évaluation axiologique de chaque production linguistique : une forme de jugement de valeur assortie d'une hiérarchisation inhérente à toute évaluation de production linguistique. Le registre qui se définit comme « une variété linguistique appropriée à une situation sociale particulière, ou encore un sous-système linguistique caractérisé par un certain nombre de constructions spécifiques et réservés à des situations circonscrites »[3], a donc tendance à assigner le locuteur dans un espace social défini au sein d'une hiérarchie[4].

Cette notion de hiérarchie m'a amenée à questionner le principe normatif. L'évaluation s'opère grâce à quels critères ? Le français dit « standard » semble constituer ce que nous appelons la norme. Toutefois sa définition est problématique car difficile à établir. Cette difficulté peut s'expliquer par une inversion du mouvement normatif ancré dans l'imaginaire commun. En effet, la norme ne produit pas l'évaluation mais elle la renforce grâce à son mécanisme de rationalisation. Autrement dit, elle donne des justifications a posteriori de ce que nous avons perçu comme étant une variation : l'évaluation et la norme s'entretiennent. Ainsi, je propose de définir la norme de deux manières différentes : faite des règles de grammaire d'une langue ou bien constituée de règles qui émergent d'une opposition à la variation. Ces deux définitions ont pour finalité de produire une norme composée des mêmes règles. Toutefois, leurs mouvements sont inversés : la première est une définition en soi tandis que la seconde se construit en réaction et donc en interaction avec les multiples productions linguistiques potentielles. La seconde proposition m'a permis d'établir une première supposition : tout ce qui ne sera ni familier, ni soutenu sera standard. Il est intéressant de noter que le registre « neutre », celui que nous concevons comme la base même de notre apprentissage scolaire et l'échelle 0 des registres de langue ne se laisse pas saisir aussi nettement et facilement que nous l'aurions imaginé.

En outre, cette définition basée sur les règles de grammaire qui produiraient un registre « neutre » / « standard » reflète une pratique vue comme « correcte » car calquée sur l'écrit : à l'école pour apprendre à maîtriser le français nous faisons des dictées non des improvisations orales. Ainsi, c'est à l'écrit qu'intervient la standardisation tandis que la variation se produit à l'oral. Ce paradigme sous-tend toute la notion de « registre » puisque l'évaluation opère selon lui. Plus la production linguistique

[3] FERGUSON, « Simplified registers and linguistic theory», Exceptional language and linguistics, Loraine K. Obler & Lise Menn éds, New York, Academic Press, 1982, p.49-66.
[4] BLOOMFIELD 1927

semble orale plus elle va être perçue comme familière car spontanée et non structurée, en revanche plus elle paraît écrite plus elle va être appréhendée comme soutenue car perçue plus construite et réfléchie (donc plus rationnelle).

Ce fut là ma première difficulté : notre corpus est uniquement composé d'œuvres écrites. La représentation écrite d'un registre dont la classe dépend de son degré d'oralité fausse la variation par son medium. Autrement dit, la norme implicitement indexée sur l'écrit entraîne une construction des registres qui opère sciemment un écart avec ce qui est considéré comme « standard » voire « correct ».

Cette action renvoie une conception de la langue hétérogène puisqu'elle occasionne elle-même un écart afin de créer une variation. Or, une caractéristique intrinsèque de la variation est la non-conscience du locuteur de cette dernière. En effet, les locuteurs réalisent une action d'uniformisation spontanée : ils conçoivent la langue comme quelque chose d'homogène, or comme nous l'avons vu, percevoir la variation passe par l'évaluation de la façon dont les gens parlent. Paradoxalement nous pouvons donc évaluer la variation chez un autre locuteur mais pas dans notre propre production linguistique. Or les textes étudiés proposent une représentation construite d'un certain registre de langue où la variation est sciemment reconstituée.

La frontière entre style et genre devient alors difficile à délimiter. En effet, le registre est « une variété linguistique appropriée à une situation sociale particulière, ou encore un sous-système linguistique caractérisé par un certain nombre de constructions spécifiques et réservées à des situations circonscrites »[5] tandis que le style « n'est pas défini en termes de formalité, mais plutôt en fonction d'une situation spécifique et de l'emploi qu'on y fait, d'un lexique spécialisé et d'un ensemble précis de constructions grammaticales »[6]. Cette difficulté provient de l'utilisation à l'écrit du style (fort et spécifique à chaque auteur) dans le but de mimer un registre.

Ainsi, comment dissocier le style du registre ? En d'autres termes, comment distinguer la partie construite qui se situe en dehors du registre (puisqu'elle tend justement à le représenter avec des outils linguistiques sciemment utilisés) et le registre lui-même (qui échappe à la construction et devient alors symptomatique d'une variation linguistique) ?

Le style nous informerait alors de la conception des registres grâce aux textes rendus mimétiques de ces derniers tandis que le registre nous indiquerait la présence d'une variation opérée de manière non-consciente dans les textes. Par exemple, prenons différentes citations tirées de l'étude comparative « Les registres familier, populaire et vulgaire dans *Le Canard enchaîn*é et *Charlie Hebdo* »[7] :

- « Vous détenez un petit compte bancaire aux Bahamas ou à Jersey ? Comment puiser discrètement, et à distance, dans ce *magot* sans être tracassé ? »[8],

[5] FERGUSON 1982 (travail sur la diglossie, et des langues en contact)
[6] AUGER 1997
[7] ILMOLA 2012
[8] C., J. 'La combine de BNP Paribas pour gruger le fisc', Le Canard enchaîné, le 11 mai 2011, 4.

- « Le conseil général la *saque* non pas parce qu'elle gagne un peu trop, mais au contraire pour ne pas avoir atteint un « revenu minimum » d'au moins 701 euros par mois avec son activité d'autoentrepreneur ! »[9]
- « Aussi dément que cela paraisse, il prend au sérieux le droit à l'alimentation, sur une terre où 1,2 milliard [sic] de *couillons* loin de chez nous souffrent d'une faim chronique. »[10]

Dans ces trois exemples seul un terme par phrase est du registre familier voire vulgaire : « magot », « saque », « couillons ». Il est intéressant de noter que le registre neutre oscille avec le soutenu car nous trouvons la forme « cela » dans le troisième exemple, ainsi que des planificateurs de discours dans le deuxième : « non pas parce qu'… », « mais au contraire pour… ». Il est intéressant de mettre en regard le but de ces citations et leur lexique : ce sont deux journaux satiriques qui tendent à faire rire leurs lecteurs. Or l'effet comique naît d'une rupture de l'horizon d'attente du lecteur. Dès lors le décalage entre le genre neutre voire soutenu de la phrase avec un élément lexical du registre familier voire vulgaire tend à opérer cette rupture afin de créer l'effet comique. Ce mécanisme ne met pas réellement en lumière une variation linguistique mais au contraire une utilisation consciente des registres afin de jouer avec la perception de ces derniers chez le lecteur dont la surprise vient de la capacité à identifier les différents registres mis en œuvre. Ainsi, nous trouvons dans ces exemples un style mimétique de différents registres sciemment utilisés : cela me permet de confirmer la perception que nous avons dans notre imaginaire commun de différents registres mais non de comprendre la logique de la variation dans une production linguistique.

Ainsi, traiter un corpus écrit avec une problématique ou l'oralité était discriminante faussait mon approche des registres et plus précisément de la variation linguistique : c'est cette difficulté que j'ai rencontrée et que j'ai voulu solutionner. Pour cela j'ai procédé en trois grands temps : tout d'abord j'ai identifié les descripteurs et les ai testés sur les différents textes, puis j'ai relevés les différentes difficultés liées à la nature du corpus et enfin en vue de dépasser cette résistance j'ai travaillé sur les discours rapportés afin de me dégager d'une oralité forcément familière.

I – Identification des descripteurs

Comment appréhender les textes afin d'en dégager des éléments discriminants qui les classeraient dans tel ou tel registre ? Suite à un état de l'art théorique, j'ai envisagé deux approches : lexicale et grammaticale.

Instinctivement nous adoptons l'approche lexicale. En effet, les termes lexicaux peuvent être fortement connotés et donc permettent de discerner un registre. Prenons ces mots comme exemple :

[9] BARRE, I. 'Trop pauvre pour le RSA', Le Canard enchaîné, le 18 mai 2011, 4.
[10] NICOLINO, F. 'L'agriculture bio l'arme ultime contre la faim', Charlie Hebdo, le 18 mai 2011, 7.

« Sécu »[11] ou bien « Sécurité sociale »[12], « meuf »[13] ou bien « Mademoiselle »[14], « discrétos »[15] ou bien « discrètement »[16]. Ces couples de termes lexicaux ont pour référent extralinguistique le même objet, toutefois les conations véhiculées par les différentes actualisations lexicales agissent sur le sens et déterminent un registre différent. Ainsi, ce sont des éléments linguistiques avec une forte dimension discriminante. Cependant, la première difficulté pour traiter le niveau lexical est son impossible exhaustivité. En effet, la variation par essence même diffère selon les contextes : il semble donc difficile de faire une liste de tous les termes non standards.

Une mesure de richesse lexicale aurait pu être envisagée : plus il y a de différents termes lexicaux employés plus nous pouvons supposer que le lexique est riche. Toutefois, le registre familier est reconnu pour sa créativité. Par exemple, sont présentes de nombreuses dérivations de termes : « quelqu'un de gauche » « gaucho », « directeur » « dirlo », « vache » vachement ». Ainsi une évaluation de richesse lexicale pourrait être pertinente si nous sommes en mesure de déterminer si les termes sont standards ou non. Là encore, la nécessité d'un dictionnaire exhaustif de tous les termes standards représente un obstacle. En effet, si un inventaire complet peut-être réalisable, en revanche définir si chaque terme est standard ou non relève de l'arbitraire.

Ainsi, le lexique est paradoxalement l'élément sans doute le plus discriminant mais le plus difficile à saisir également. Effectivement, il présente une entrave majeure de par sa nature même : chercher un terme lexicalisé et l'indexer selon sa connotation n'est pas raisonnable en traitement automatique des langues. Cela explique la faible importance de l'approche lexicale dans mon travail.

Voici les différents descripteurs identifiés associés aux références bibliographiques dont ils sont extraits (pour plus de précision cf annexes : « *tableau synthèse* ») :

<u>Approche lexicale</u>

- Liste non-exhaustive de termes discriminants - Branca-Rosoff 1999 - Bilger et Cappeau 2004
  Ex : *néanmoins, à même de, en raison de, moindre, car, compte tenu de, suite à, dans la mesure où*

- Emprunts - Gadet 2003
  Ex : *marav, pourav, graillav, toubib, kawa*

- Onomatopées - Maarit Ilmola 2012
  Ex: *pouah, ah, oh, bah, beh, ben*

- Adverbes spécifiques - characteristics_linguistics_formality.pdf 2015
  Ex : *en tout état de cause → en tout cas, néanmoins, malgré cela → quand même, de plus, en outre → en plus*

[11] GUENE *Kiffe Kiffe demain*
[12] L'*Humanité*
[13] GUENE *Kiffe Kiffe demain*
[14] PROUST *Albertine Disparue*
[15] GUENE *Kiffe Kiffe demain*
[16] L'*Humanité*

- Anglicisme familier
  Ex : *job*

- Verbes - Maarit Ilmola 2012
  Ex : *chier, saquer, foutre, dézinguer*

- Préférence pour les verbes du premier groupe
  Ex : *chuter → choir, solutionner → résoudre*

- Locutions verbales - Maarit Ilmola 2012
  Ex : *trouver une solution*

- Décumul des comparatifs synthétiques- Gadet 1997
  Ex : *plus mauvais, plus pire*

- Sur présence de "là" - Gadet 1997
  Ex : *ça va là, c'est bon là ! tu me soules là*
- Fréquence des "et" – Jade Mekki

- « Et » en début de phrase – Jade Mekki

- Couple de termes lexicalisé « des fois » / « parfois » – Jade Mekki

- « Et puis » en début de phrase – Jade Mekki

<u>Approche grammaticale</u>

- Contraction
  Ex : *ça → cela / c'est → cela est*

- Sujet « on » transposé en « nous » - Bilger et Cappeau 2004
  Ex : *on → nous*

- Terminaisons en -asse - Maarit Ilmola 2012
  Ex : *connasse, chiasse, pétasse, chagasse…*

- Terminaisons en -o - Maarit Ilmola 2012
  Ex : *Gaucho, crado, dirlo, auto, moto, cyclo, chrono, micro, proprio,*

- Terminaisons en -ouze - Maarit Ilmola 2012
  Ex : *partouze, tantouze, flouze*

- Terminaisons en -iotte - Maarit Ilmola 2012
  Ex : *fiotte, chiotte*

- Dérivation d'un nom/adjectif en adverbe - Maarit Ilmola 2012
  Ex : *vachement, sacrément*

- Terminaisons en -ou - Maarit Ilmola 2012
  Ex : *relou, chelou*

- Répétition de voyelles - Branca-Rosoff 1999
  Ex : *Je m'en fooooooooous, ras le boooool*

- Utilisation des MAJ - Branca-Rosoff 1999
  Ex : *RIEN A FOUTRE ! TA GUEULE*

- Répétition des signes de ponctuation - Branca-Rosoff 1999
  Ex : *vas y !!!!!!*

- Passage du tutoiement au vouvoiement - Bilger et Cappeau 2004
  Ex : *Tu vas bien ? / Allez-vous bien ?*

- Augmentation du nombre de syllabes en /oe/ sous l'effet de mots verlanisés - Gadet 2003
  Ex *: keuf, meuf*

- Troncation finale (fréquent) - Gadet 1997
  Ex : *assoce, biz, délib, deg, anar, écolo*

- Troncation initiale (plus rare) - Gadet 1997
  Ex : *blème*

- Redoublement - Gadet 1997
  Ex : *leurleur, zonzon*

- Inversion de l'ordre - verlan - Gadet 1997
  Ex : *noich oinich oinch*

- Elision du "e" - Favart 1966-2006
  Ex : *C'est pas d'ma faute, si j'pleure pas. J'suis un optimiste ! , Et il répond quoi le p'tit ? , nous c'qu'on est ça se voit tout de suite*

- Apocope dans le pronom "tu" - Favart 1966-2006
  Ex : *T'as pas envie d'aller à la plage ? Après ça, t'as plus envie de leur parler.*

- Apocope du "r" - Favart 1966-2006
  Ex : *Eh m'dame, c'est quoi, vot' religion ?*

- Liaisons - "z" parasite - Favart 1966-2006 / Gadet 2003
  Ex : *On passe des zheures à se parler , Et c'est qu'y z'y tenaient de pire qu'à leurs yeux à leur démarcation du diable!*

- Ecriture phonétique avec une utilisation des chiffres par exemple - Sommant 2005
  Ex : 100 = sans, mat1

- Mots plus courts - Sommant 2005
  Ex : *koi*

- Lettre unique - Sommant 2005
  Ex : *C*

- Elision du pronom "qui" devant une voyelle - characteristics_linguistics_formality.pdf 2015

Ex : *qui avait / qu'avait*

- Subjonctif → alignement sur le présent - Gadet 2003
  Ex : *« c'est elles qui t'apprendent » « ils croivent »*

- Utilisation du subjonctif ou non utilisation - Gadet 2003
  Ex : *il faut que tu viennes / il faut venir*

- Type d'interrogative - sans inversion - Gadet 2003
  Ex : *il dit quoi? / que dit-il?, il vient? / vient-il?*

- Syntaxe neutre avec un élément lexical discriminant - Maarit Ilmola 2012
  Ex : *Je souhaiterais obtenir un rendez-vous auprès du dirlo*

- La construction de la négation - Bilger et Cappeau 2004
  Ex : *ma mère elle danse pas ; il a pas fait / il ne l'a pas fait*

- "ça" + verbe
  Ex : *ça craint, ça fait chier/dégueu, ça te bousille les yeux*

- Construction du futur
  Ex : *je vais à la plage, je vais manger, je vais boire, je vais dormir*

- Mauvaise utilisation des prépositions :
  Ex : *aller chez le coiffeur / aller au coiffeur*

- "Relative populaire" - Gadet 2003
  Ex : *c'est ça que j'ai besoin ?*

- Utilisation du passé simple vs passé composé
  Ex : *il fit, il a fait / il fut, il a été / il vient, il est venu*

- Mise en commun du sujet pour plusieurs verbes successifs - Bilger et Cappeau 2004
  Ex : *C'étaient des personnes qui n'avaient pas prévu leur retraite et envisageaient un avenir normalement*

- "c'est" devant un SN pluriel - Favart 1966-2006
  Ex : *les deux sont liés bien sûr mais en réalité c'est deux crises*

- Métaphores animales - Maarit Ilmola 2012
  Ex : *maquereau, chienne, chien*

- Effacement du pronom sujet "il" dans les constructions impersonnelles - il = y - Favart 1966-2006
  Ex : (Particulièrement avec "avoir" et "falloir") *Fallait pas y faire, Y a du grabuge*

- Sujet et Verbe en un mot = phonétique - Sommant 2005
  Ex : *GT, TT, JTM, Tpa, Tki, TKt*

- Transposition de la structure "qu'est-ce que" qu'est-ce" - characteristics_linguistics_formality.pdf 2015
  Ex : *Qu'est ce t'as maigri*

- Surutilisation du présent indicatif - characteristics_linguistics_formality.pdf 2015

- Verbe 'être' au singulier + syntagme pluriel - Bilger et Cappeau 2004
  Ex : *c'est nos amis, c'est mes affaires*

- COD précède le verbe et oublie de l'accord -characteristics_linguistics_formality.pdf 2015
  Ex : *La robe que j'ai mis, la chaussure que j'ai porté*
- Participe passé du verbe "faire" est suivi d'un infinitif → accord en nombre ni en genre - Gadet 2003
  Ex : *La maison qu'elles se sont faite(s) construire est très grande*

- Si → conditionnel - characteristics_linguistics_formality.pdf 2015
  Ex : *Si je serais venu*

- Fréquence (basse) des phrases passives → recours au "c'est" - characteristics_linguistics_formality.pdf 2015
  Ex : *Le chalet a été loué par des vacanciers italiens -> c'est des italiens qui ont loué le chalet*

- Utilisation d'"est-ce que" pour les interrogatives - Maarit Ilmola 2012
  Ex : *Est-ce que tu vas bien? / Comment vas tu? , quand est-ce que c'est qu'il a dit ça?*

- La fréquence des mots de liaisons - Gadet 2003
  Ex : *certes, je ne vais pas …, mais …, enfin peut être.*

- Des éléments "ponctuant" - Gadet 2003
  Ex : *comme quoi, vas-y, grave…*

- Outils de "planification" du discours- (en une phrase) Branca-Rosoff 1999 - Bilger et Cappeau 2004
  Ex : *d'abord, ensuite enfin*

- Utilisation d'adverbe + parataxe : phrase adverbiale
  Ex : *vraiment bien, vraiment chiant*

- La substantivation / nominaliser - Bilger et Cappeau 2004
  Ex : *les façades ont été ravalées → le ravalement des façades*

- Emploi de relatifs - Gadet 2003
  Ex : *dont / lequel, Le lieu dans lequel elle va passer la fin de sa vie, cette nouvelle association là dont je vous parlais*

- Maintien de "des" devant un adjectif - characteristics_linguistics_formality.pdf 2015
  Ex : *Ils font des très bonnes tartes*

- La pronominalisation est faite avec une préposition et non un pronom
  Ex : *Où est la voiture ? Ton père est parti avec*

- Les pronoms "y" et "en" pour faire référence à un être humain - Favart 1966-2006
  Ex : *Il y pense tout le temps*

- Surutilisation de "que" / à la place d'autres pronoms relatifs - Favart 1966-2006

Ex : *le type avec lequel tu es parti = le type que tu es parti avec*

- Inversement de l'ordre des mots - characteristics_linguistics_formality.pdf 2015
  Ex : *Rouge, qu'elle est, la maison de ma mère*

- Phrase impérative : COI devant le COD - characteristics_linguistics_formality.pdf 2015
  Ex : *Donne moi-le*

- Phrase impérative: pronom "en" devant le COI - characteristics_linguistics_formality.pdf 2015
  Ex : *garde moi-s-en*

- Fréquence de "c'est…qui/que…" - characteristics_linguistics_formality.pdf 2015
  Ex : *c'est beau! C'est sympa! c'est lui qui a fait ça*

- Juxtaposition - characteristics_linguistics_formality.pdf 2015
  Ex : *c'est le chien qui court, qui saute*

- Utilisation de "que" plutôt que de "où" - characteristics_linguistics_formality.pdf 2015
  Ex *: Je l'ai acheté dans le magasin qu'on était l'autre jour avec toi*

- Rajout "à lui" "à elle" après les pro "son" "sa" - Gadet 2003
  Ex : *sa chaise à elle, son pull à lui*

Les descripteurs qui émergent de l'état de l'art théoriques opèrent soit une utilisation d'éléments rarement utilisés de la norme ou bien un écart avec cette dernière. De fait, chaque unité a une fréquence relative au sein d'un texte. Des termes que nous appelons « mots outils » tels que « et », « un », « une », « le », « la » … sont très fréquents et ont donc une fréquence relative haute. En revanche, d'autres termes moins souvent utilisés tels que « cependant », « en revanche », « toutefois » ont, une fréquence relative basse.

$$fréquence\ relative = \frac{nombre\ d'occurence\ du\ terme\ dans\ le\ texte}{nombre\ total\ de\ terme\ dans\ le\ texte}$$

Ce calcule malgré sa simplicité met en exergue des termes très utilisés ou bien au contraire peu utilisés ce qui me permet d'employer le couple de notion : fréquents/rares. Ainsi, nous pouvons remarquer que plusieurs descripteurs pour le registre soutenu se base sur l'emploi de termes « rares » : outils de planification du discours et surtout dans l'approche lexicale. En effet, si nous ne pouvons pas faire de liste exhaustive, nous pouvons néanmoins penser les motifs discriminants pour le registre soutenu comme rarement utilisés.

De plus, même si une mesure de richesse lexicale n'a pas été retenue comme pertinente, la fréquence de certains termes tel que « et » permet de discriminer le registre familier et dénote d'une surutilisation d'une même conjonction de coordination : il est possible d'en déduire une certaine pauvreté du lexique.

En outre, le registre familier apparaît à travers les descripteurs comme le résultat d'écarts avec la norme qui peuvent être définis comme des « fautes ». Ce constat est intuitif puisqu'est fautif « tout ce

qui ne se conforme pas à la norme »[17] or le registre familier est un registre qui diffère du registre « courant » car non-standard donc non normé.

Pourtant la notion de faute est intéressante. Non pas pour sa valeur axiologique implicitement associée, mais pour les raisons qui amènent le locuteur à produire un énoncé fautif. En effet, « *On ne fait pas des fautes pour le plaisir de faire des fautes. (…) Dans un grand nombre de cas la faute, qui est passé jusqu'à présent pour un phénomène quasi-pathologique sert à prévenir ou à réparer les déficits du langage correct.* »[18]. Autrement dit, la faute serait le symptôme d'un « déficit » du Français. Il y aurait donc une sorte de régularisation spontanée des irrégularités arbitraires de la langue normée. Frei présente les fautes comme venant palier un « besoin » : besoin de désambiguïser, besoin d'être expressif…

Il propose dans *La Grammaire des Fautes* une typologie de ces « besoins » :

- L'assimilation
- La différenciation pour une recherche de la clarté, pour désambiguïser
    - Il croit / ils croient : sont phonétiquement similaires
    - Ils croivent : permet une différenciation phonétique
- L'économie
    - « Structuraliste / en tout cas / Lacan, / il se le dit » : parataxe, le propos est bref
    - « Les femmes, ça veut toujours plus qu'on leur donne » : « les femmes » sont reprises par « ça » ce qui permet une invariabilité de nombre.
- L'expressivité pour « agir contre l'usure sémantique »[19]
    - « Moi / les enquêtes de terrain / bof »

Ainsi, le locuteur adapterait naturellement son énoncé linguistique selon ses besoins au dépend du respect de la norme grammaticale. Cela diffère un peu de ce que propose Françoise Gadet. En effet, elle ne parle pas de « besoins » mais de « déficits » ou bien « d'irrégularités » que les locuteurs homogénéisent spontanément.

Prenons trois syntagmes donnés comme exemples par F. Gadet dans *Le Français ordinaire* :

1. « Pierre lui la donne » – fautive
2. « Pierre me la donne » – correcte
3. « Pierre la lui donne » – correcte

La règle grammaticale est la suivante : l'ordre des unités lexicales change selon le pronom personnel. Ainsi, à la première personne du singulier arrive tout d'abord le pronom personnel qui est complément d'objet indirect puis le pronom personnel complément d'objet direct tandis qu'à la troisième personne c'est le pronom personnel complément d'objet direct puis le pronom personnel complément d'objet indirect.

1ère personne SG : COI / COD
3ème personne SG : COD / COI

Cette règle complexe et isolée est « palliée » spontanément par le locuteur qui aligne la construction sur le schéma 3. De plus, cette représentation de la faute peut être appuyée par le fait que peu d'erreur, voire aucune, aboutisse à un échec de la communication.

[17] GADET, *Le Français ordinaire* 1997
[18] FREI, *La Grammaire des Fautes* 1929
[19] FREI, *La Grammaire des Fautes* 1929

Ainsi, la « faute » chez Françoise Gadet permet d'identifier un mécanisme linguistique où le locuteur va spontanément « corriger » les règles de grammaire arbitraire du Français. D'un point de vue TAL nous pouvons donc supposer que les règles grammaticales « complexes » ou bien « irrégulières » représentent des « zones syntaxiques » où nous avons une forte probabilité d'y trouver une variation linguistique.

Dès lors, la faute résulte d'une variation causée par deux dynamiques différentes. La première envisagée par Henri Frei met en exergue le manque de ressource linguistique par rapport aux besoins des locuteurs : cela conditionne alors la manière d'exprimer ce qu'ils veulent dire et peut occasionner des « fautes ».

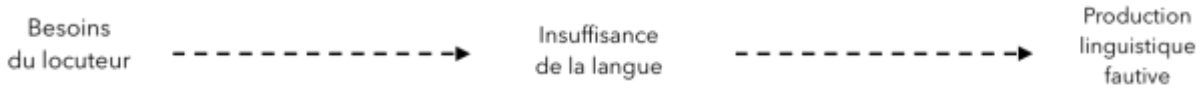


Figure 1 mécanisme de la faute chez Henri Frei

La seconde présentée par Françoise Gadet présente une tendance instinctive à régulariser les irrégularités de la langue française.

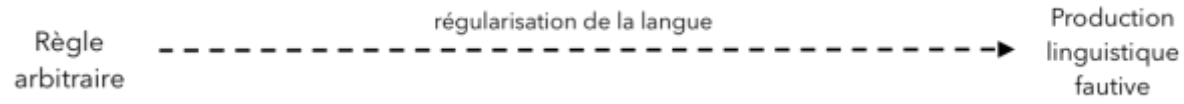


Figure 2 mécanisme de la faute chez Françoise Gadet

Ainsi, la faute est discriminante pour le registre familier et permet d'identifier une dynamique qui déclenche une variation linguistique. Cette rapide analyse des descripteurs me permet de consolider la différence entre registre et style. Il est important de noter ici que le style est toujours motivé, voire contraint, par différents facteurs : irrégularités, besoin de désambiguïser…

II – Test des descripteurs

A – Réflexion et composition du corpus

J'ai rapidement engagé une réflexion sur la constitution du corpus lui-même. La problématique liée au support écrit des textes me poussa à élargir le corpus. En effet, non seulement le corpus était écrit mais surtout chaque texte était représentatif à la fois d'un registre et d'un genre particulier.

- *Albertine Disparue* de Proust
  - Registre soutenu
  - Genre littéraire
- L'*humanité*
  - Registre courant / standard
  - Genre journalistique
- *Kiffe Kiffe Demain* de Guene
  - Registre familier

    - Genre littéraire
- Histoires à parler
    - Registre enfantin
    - Genre littéraire (jeunesse)

Dès lors le style est contraint par deux facteurs : le support écrit mais également le genre du texte. Cette double dépendance crée un contenu textuel très spécifique : les textes devenaient alors insuffisamment représentatifs du registre auquel ils étaient associés.

Le texte pour le genre familier, *Kiffe Kiffe Demain,* était trop spécifique à un « sous-registre » dont le lexique constituait un facteur majeur de son appartenance au registre familier :

- « Les profs, entre deux grèves, se sont dit que j'avais besoin de voir quelqu'un parce qu'ils me trouvaient renfermée. »
- « Il est gros, il est con, quand il ouvre la bouche ça sent le vin de table Leader Price et en plus il fume la pipe. »
- « Tous des cons, franchement Schihont, c'est son responsable. »
- « Cet enfoiré de M. Schihont »

Or comme nous l'avons vu, j'ai essayé d'éviter une approche basée sur le lexique.

Il en va de même pour le texte extrait du journal l'*Humanité* : la forte présence de discours rapporté tendait à classer ce texte dans le familier bien qu'il soit censé être représentatif du standard :

- " Mais on n'a pas encore touché à ces sommes pourtant providentielles. "
- " on est dans une logique désespérante "
- « - Ah bon ! Tom Cruise, maintenant ! il ne manquait plus que ça ! »
- « - (…) Les larmes en russe, ça se dit : sliozy. »

Afin de contrebalancer ces deux obstacles j'ai rajouté plusieurs textes :

- *Le Monde* – archives de l'année 2016 rubrique internationale
    - Contenu : tous les titres et tous les résumés
    - Registre : standard
    - Genre : journalistique
- *L'Assommoir* de Zola
    - Registre : familier
    - Genre : littéraire
- *Voyage au bout de la nuit* de Céline
    - Registre : familier
    - Genre : littéraire

Le Monde équilibre le registre standard dont fait partie l'*Humanité* car je n'ai extrait (cf *extract_lemonde.pl* dans les annexes) que les titres et les résumés de la rubrique international sur toute l'année 2016. Autrement dit, cela devrait éviter la présence de discours rapporté (ex : entretiens, interviews…) et donc contrebalancer le texte extrait de l'*Humanité*. Pour le registre familier qui est le plus difficile à systématiser j'ai dû rajouter des textes de sous-registres différents : des styles et des époques différentes. Pour les deux registres j'ai toutefois respecté les genres afin de garder une cohérence dans les comparaisons inter-registres : journalistique pour le courant et littéraire pour le familier.

B – Préparation du corpus

Tout d'abord j'ai voulu extraire le contenu textuel des archives de l'*Humanité* : les fichiers initiaux du corpus étaient des fichiers XML (cf. Figure 3). Pour cela j'ai utilisé un script perl qui s'appuyait sur la particularité des fichiers XML : chaque phrase/syntagme est encapsulé(e) par les balises <P></P>.

```
<?xml version="1.0" encoding="UTF8"?>
<Documents>
<Document>
  <DocId>2007-11-02_International_Tensions-au-pied-des-monts-Cudi</DocId>
  <Article>
    <Date Annee="2007" Mois="11" Jour="02">
    </Date>
    <Secteur>International</Secteur>
    <Corps>
      <Titres>
        <SurTitre>
          <P>événement</P>
        </SurTitre>
        <Titre>
          <P>Tensions au pied des monts Cudi</P>
        </Titre>
      </Titres>
        <Chapo>
          <P>kurdistan turc . À Ciszré et Silopi, cités frontalières de l'Irak, l'inquiétude grimpe alors que l'armée turque est de plus en
        </Chapo>
      <Texte>
          <P>Frontière turco-irakienne, envoyé spécial.</P>
          <P>Près des localités de Ciszré (50 000 habitants) et de Silopi (25 000 âmes), des hélicoptères de combat turcs ont pilonné lundi
          <P>Tout comme Silopi, Ciszré se trouve au coeur du conflit. La région est quadrillée. À quelques kilomètres de l'entrée de la vil
          <P>quand les civils</P>
          <P>évitent les sujets qui fâchent</P>
          <P>Comme beaucoup de Kurdes, il ne désapprouve pas les actions du PKK. « Les télés et les journaux ne parlent que des soldats tué
          <P>Dans les cafés, les gens regardent, sans commenter, les chaînes d'info télé, qui matraquent à longueur de journée sur la doule
          <P>Dans cette ville dont le maire, Aydin Budak, a été condamné en juin 2006 à un an et trois mois de prison pour propos « séparat
          <P>Jusque-là, Tayyip Erdogan, dont le parti - l'AKP (Parti de la justice et du développement) - est arrivé en deuxième position d
          <P>l'inquiétude des habitants de silopi, cité commerçante</P>
          <P>Une dizaine de kilomètres plus loin, à Silopi, cité commerçante qui jouxte immédiatement l'Irak et qui est traversée par des c
          <P>Comme à Ciszré, les alentours de Silopi sont parsemés de postes militaires sur les hauteurs de la montagne et le long du Tigre
          <P>les turcs davantage pénalisés</P>
          <P>que les irakiens</P>
          <P>Avant de passer le poste frontière de Habur, situé à quelques kilomètres à la sortie de la ville, nombreux sont les camionneur
          <P>Depuis que le Conseil national de sécurité turc (MGK) a décidé de limiter les exportations de la Turquie vers l'Irak, le trafi
          <P>Près de 15 000 Turcs, pas tous d'origine kurde, sont installés dans le Kurdistan irakien et qui opèrent comme agents commercia
          <P>Hassane Zerrouky</P>
      </Texte>
    </Corps>
  </Article>
</Document>
<Document>
  <DocId>2007-11-02_Societe_Arche-de-Zoe-le-Tchad-lache-du-lest</DocId>
  <Article>
    <Date Annee="2007" Mois="11" Jour="02">
    </Date>
    <Secteur>Societe</Secteur>
    <Corps>
      <Titres>
        <SurTitre>
          <P>société</P>
        </SurTitre>
        <Titre>
          <P>Arche de Zoé : le Tchad lâche du lest</P>
        </Titre>
```

Figure 3 fichiers XML - archive de l'Humanité

extraction_texte_XML.pl

```perl
#!/usr/bin/perl

# Script a lancer dans le dossier où se trouvent tous les fichiers que l'on veut traiter
# Si non : modifier le chemin du dossier à donner en argument à &extraction
#-------------------------------------------------------------------------------

my $output="extract_humanite.txt";
open my $out, ">>:encoding(utf8)", $output;


&extraction(".\/");


sub extraction(dossier){
  my $path=shift(@_);
  opendir (DIR, $path);
  my @files = readdir(DIR);
  close (DIR);
  print @files;
  foreach my $file (@files) {
    if (-f $file){
      next if $file =~ /^\.\.?$/;
      my $texte= "";
      open my $input, "<:encoding(utf8)", $file;
      while (my $ligne = <$input>){
        chomp $ligne;
        $texte = $texte . $ligne;
        # print $texte;
      }
      while ($texte =~ /<P>(.+?)<\/P>/g){              #à modifier selon les fichiers XML traités
        # my $titre=$1;
        # my $chapo=$2;
        my $txt=$1;
        print $out "$txt\n\n\n";
      }
    }
  }
}
```

Figure 4 script perl pour extraite le contenu textuel des fichiers XML

```
Tensions au pied des monts Cudi

kurdistan turc . À Ciszré et Silopi, cités frontalières de l'Irak, l'inquiétude grimpe alors que
l'armée turque est de plus en plus visible face à la montagne où se cachent les combattants kurdes.

Frontière turco-irakienne, envoyé spécial.

Près des localités de Ciszré (50 000 habitants) et de Silopi (25 000 âmes), des hélicoptères de combat
turcs ont pilonné lundi et mardi les vallées en contrebas des monts Cudi séparant la Turquie de
l'Irak. On est là à la toute proximité du poste frontière de Habur, à la sortie de Silopi, l'unique
voie de passage vers l'Irak, empruntée chaque jour par des centaines de camions. Le périmètre a été
décrété zone de guerre sécurisée par l'armée turque depuis les affrontements avec le PKK.

Tout comme Silopi, Ciszré se trouve au coeur du conflit. La région est quadrillée. À quelques
kilomètres de l'entrée de la ville, des check-points de la gendarmerie se multiplient. Les identités
des automobilistes et des passagers des bus en provenance de Diyarbakir - la capitale du Kurdistan
turc - sont systématiquement vérifiées (lire également notre reportage en page 4). Tout autour, sur
les collines et les crêtes des monts bordant le Tigre, qui coule tout près de la ville, des postes
militaires surveillent la région. Et les combattants du PKK ne sont pas si loin. « Là-haut, sur les
monts Cudi », indique avec un air amusé Emer, étudiant en médecine.
```

Figure 5 passage du fichier texte brut contenant le texte extrait

La seconde tache fut d'annoter avec Tree Tagger chaque terme des textes de leur « part of speech - POS » afin de pouvoir accéder à leur catégorie morphosyntaxique. Pour cela j'ai, tout d'abord, utilisé un script perl pour « tokeniser » c'est à dire segmenter le texte afin d'obtenir un fichier texte brut avec

un terme par ligne. Ce dernier est donné en argument à la commande qui lance Tree Tagger et qui retourne un fichier à trois colonnes : terme lexical | POS | lemme.

La commande bash utilisée est la suivante :

```
perl tokenise-utf8.pl ./fichier_a_tagger.txt | ./tree-tagger -token -
 lemma -no-unknown french-oral-utf-8.par > fichier_retourne_tagge.txt
```

```
Tout  DET:pre    tout
le    DET:def    le
reste NOM   reste
n'est ADJ   n'est
que   KON   que
déceptions NOM   déception
et    KON   et
fatigues   NOM   fatigue
.     SENT  .
Notre DET:pos    notre
voyage     NOM   voyage
à     PRP   à
nous  PRO:ton    nous
est   VER:pres   être
entièrement     ADV   entièrement
imaginaire NOM   imaginaire
.     SENT  .
```

Figure 6 fichier retourné par Tree Tagger

J'ai donc suite à ces traitements deux sous-corpus différents :

1. Tous les fichiers en texte brut
2. Tous les fichiers annotés morphosyntaxiquement en texte brut

C – Test des descripteurs

La seconde étape de mon travail consistait à tester les différents descripteurs avec des outils de traitement automatique des langues. Le but de ces tests n'était pas de développer une solution outillée pour repérer automatiquement les descripteurs dans un texte mais de confirmer ou bien d'infirmer chaque descripteur. Autrement dit, de vérifier si les hypothèses formulées précédemment se vérifiaient dans les différents textes/registres.

Pour cela j'ai utilisé plusieurs outils :

1. Python
2. Unitex

→ Tous les scripts utilisés se trouvent dans les annexes

1 – Approche lexicale

1.1 – onomatopées : familier

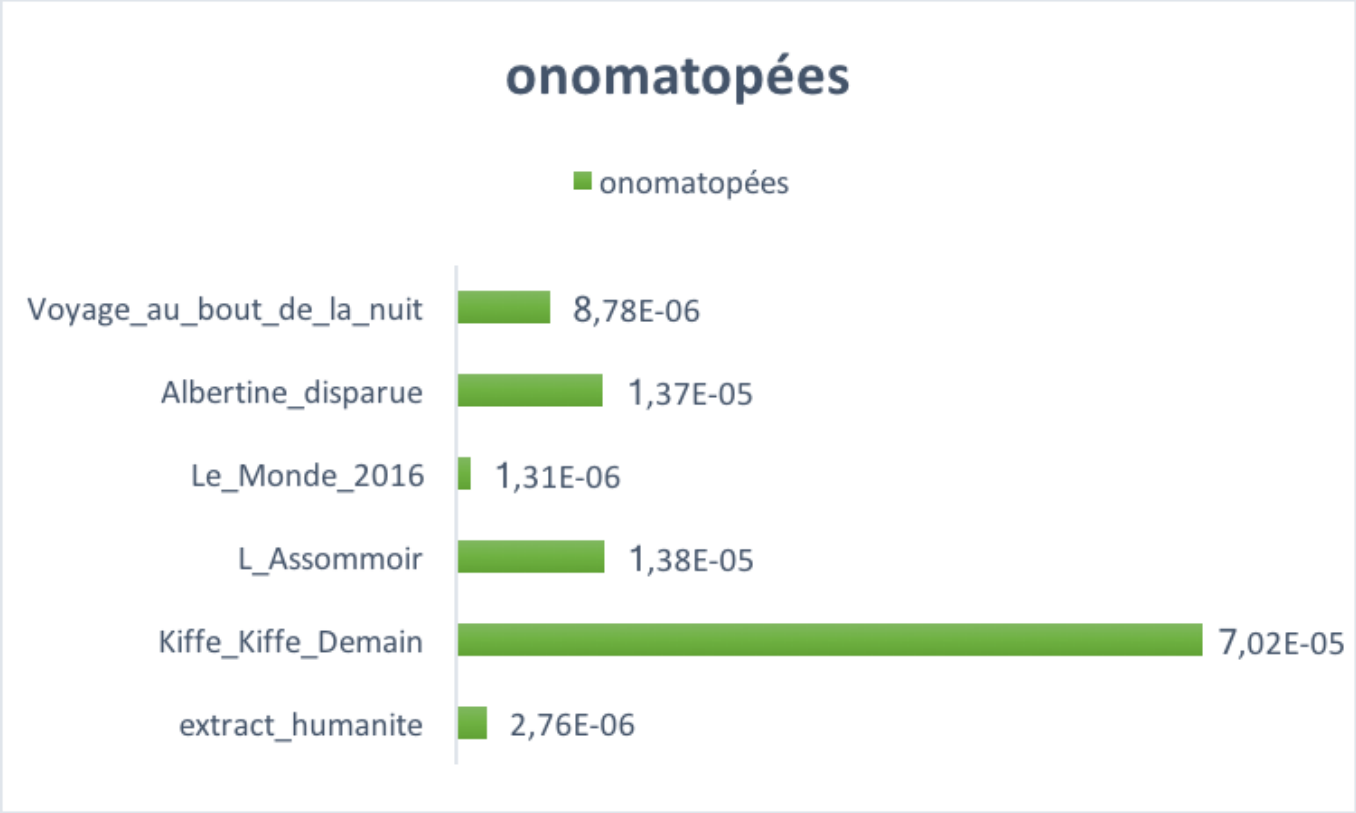


Figure 4 graphique des fréquences relatives des onomatopées dans chaque œuvre/registre

→ Il se dégage clairement une plus grande proportion d'onomatopée dans le registre familier avec *Kiffe Kiffe demain* et L'*Assommoir*. La présence en troisième position d'Albertine Disparue vient nuancer les résultats.

1.2 – sur présence de « et » : familier

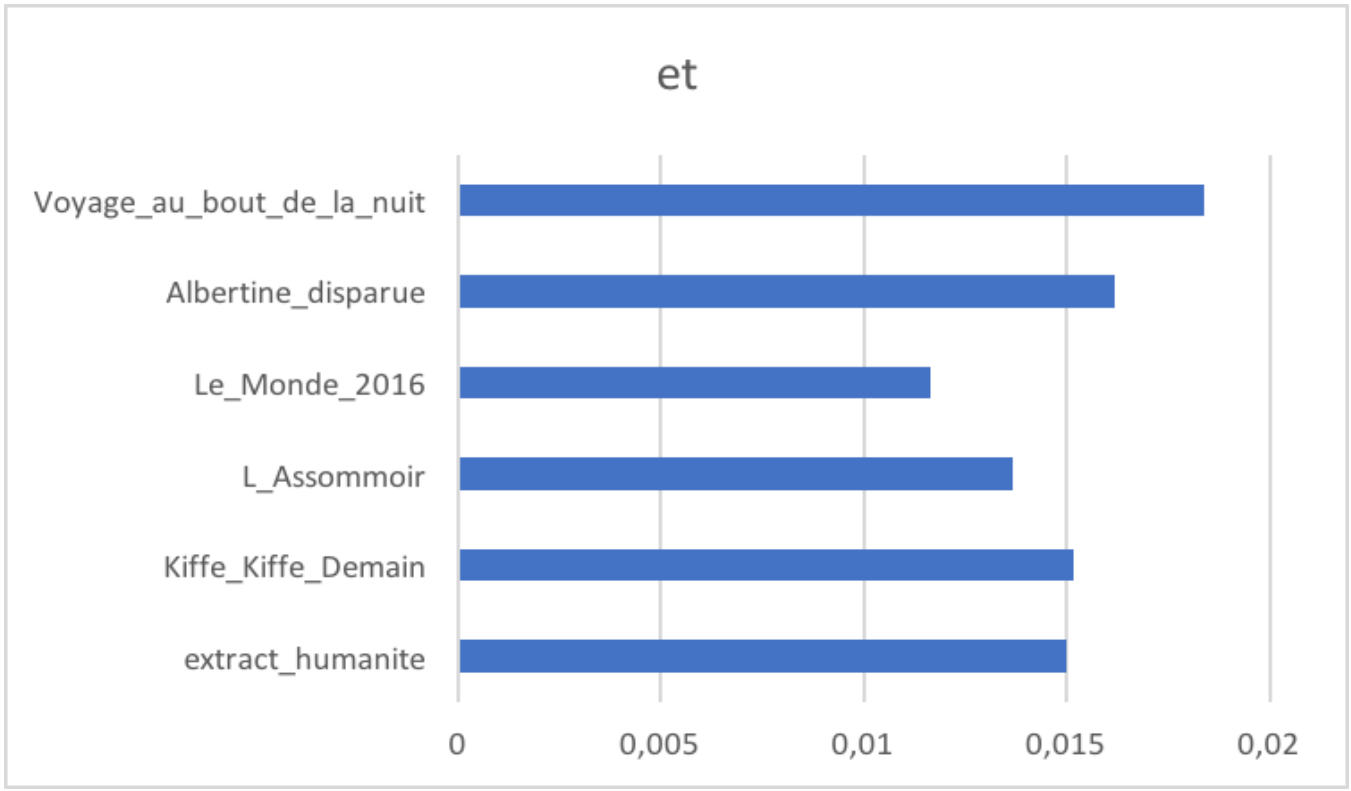


Figure 5 graphique des fréquences relatives des "et" dans chaque œuvre/registre

→ La présence de la conjonction de coordination « et » ne semble pas être discriminante pour un registre de manière très nette bien que ce soit *Voyage au bout de la nuit* donc une œuvre du registre familier qui ressorte. Suite au test de ce descripteur j'ai essayé de tester un patron lexical plus précis et plus discriminant « Et puis » en début de phrase : à voir dans l'approche grammaticale.

1.3 – Sur présence de « là » - ponctuant : familier

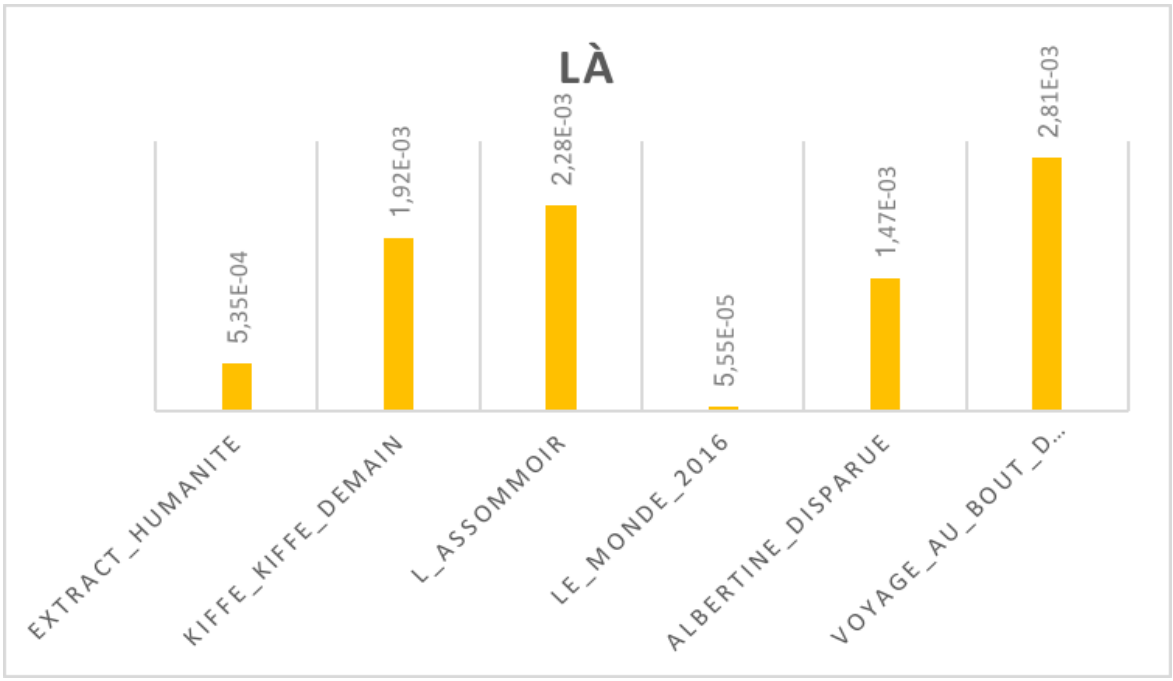


Figure 6 graphique des fréquences relatives des "là" dans chaque œuvre/registre

→ Le descripteur est nettement confirmé puisque les 3 textes avec la plus forte fréquence relative de « là » sont les trois textes du registre soutenu. Nous pourrions préciser la place du motif dans la phrase. Effectivement, une autre dimension discriminante de « là » vient de sa fonction de « ponctuant » : il serait intéressant de chercher là en fin de phrase. Pour cela j'ai simplement segmenter sur les espaces afin de conserver les signes de ponctuation et j'ai cherché le motif « là. »

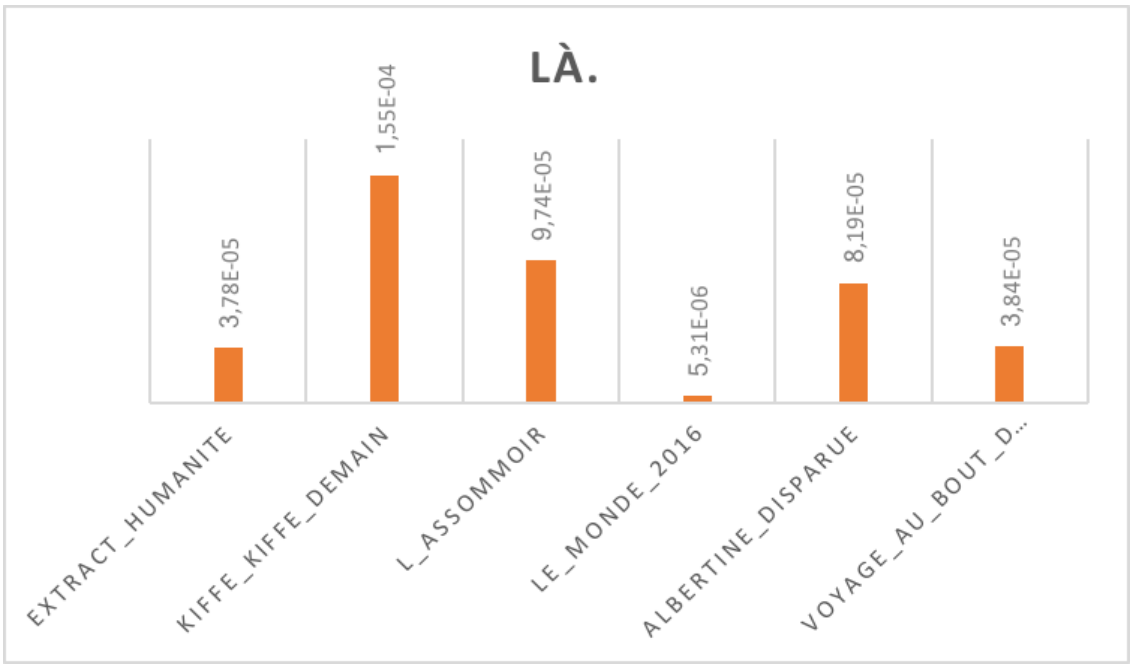


Figure 7 graphique des fréquences relatives des "là." c'est à dire en fin de phrase dans chaque œuvre/registre

→ Avec le motif plus spécifique car envisagé comme un « ponctuant » c'est toujours le registre familier qui ressort avec *Kiffe Kiffe demain* puis *L'Assommoir*. Il est intéressant de noter que la 3ème œuvre qui

ressort est Albertine Disparue ce qui va à l'encontre de notre hypothèse. Toutefois la différence de fréquence est plus grande entre le 1er texte / familier et le 3ème / soutenu qu'avec le motif simple « là ». Cela me permet de confirmer la dimension discriminante du descripteur « là » en fin de phrase pour le registre familier.

2 – Approche grammaticale

2.1 – contraction (c'est → cela est) : familier / standard

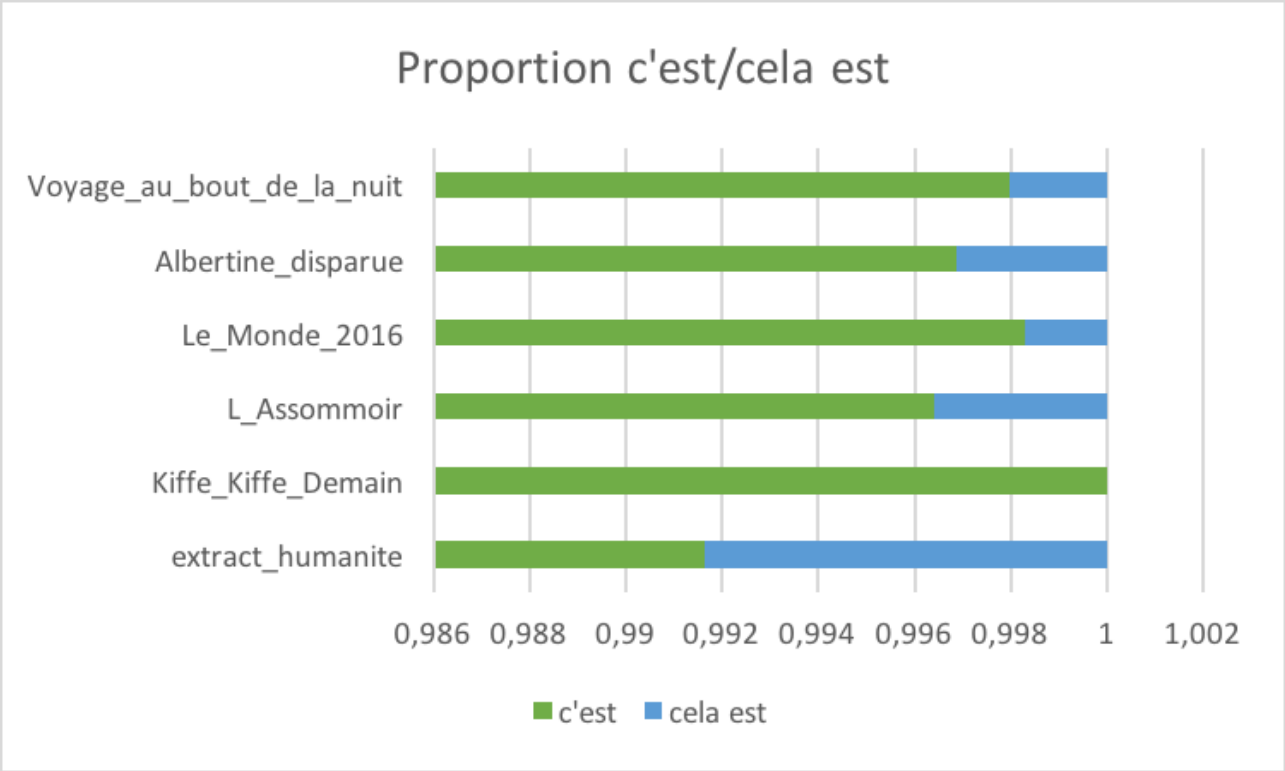


Figure 8 graphique des fréquences relatives des "c'est"/"cela est" dans chaque œuvre/registre

La fonction permet de retourner des fichiers texte brut où sont stockés les phrases extraites car contenant le motif recherché : cela me permettait d'avoir un contexte. Vous pouvez trouver ces fichiers dans les annexes.

→ Descripteur nettement confirmé. *Kiffe Kiffe Demain* n'a aucune occurrence du motif « cela est » discriminant pour le standard (nous pourrions nous interroger sur le registre : courant ou bien soutenu ?). Toutefois les minces différences entre les textes « standard » et « familier » m'ont poussé à renverser le descripteur. Autrement dit, ce descripteur ne permet d'identifier un registre par sa présence donc sa grande fréquence mais par son absence donc sa basse fréquence : ici le texte familier est repéré car n'ayant aucune occurrence de la forme non contractée.

2.2 – contraction (ça → cela) : familier / standard

→ J'ai utilisé le même script avec les motifs modifiés. Vous trouverez également dans les annexes des fichiers texte brut où sont stockés l'extractions des phrases où les motifs occurrents.

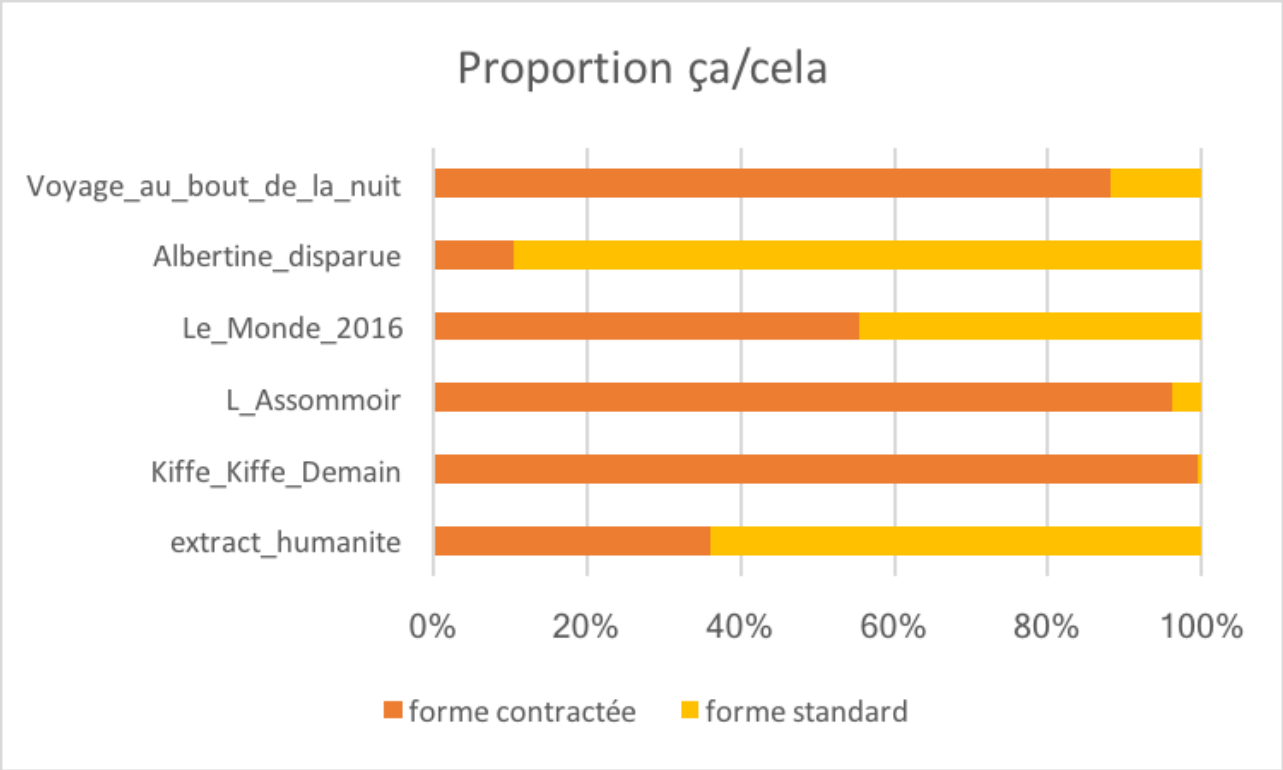
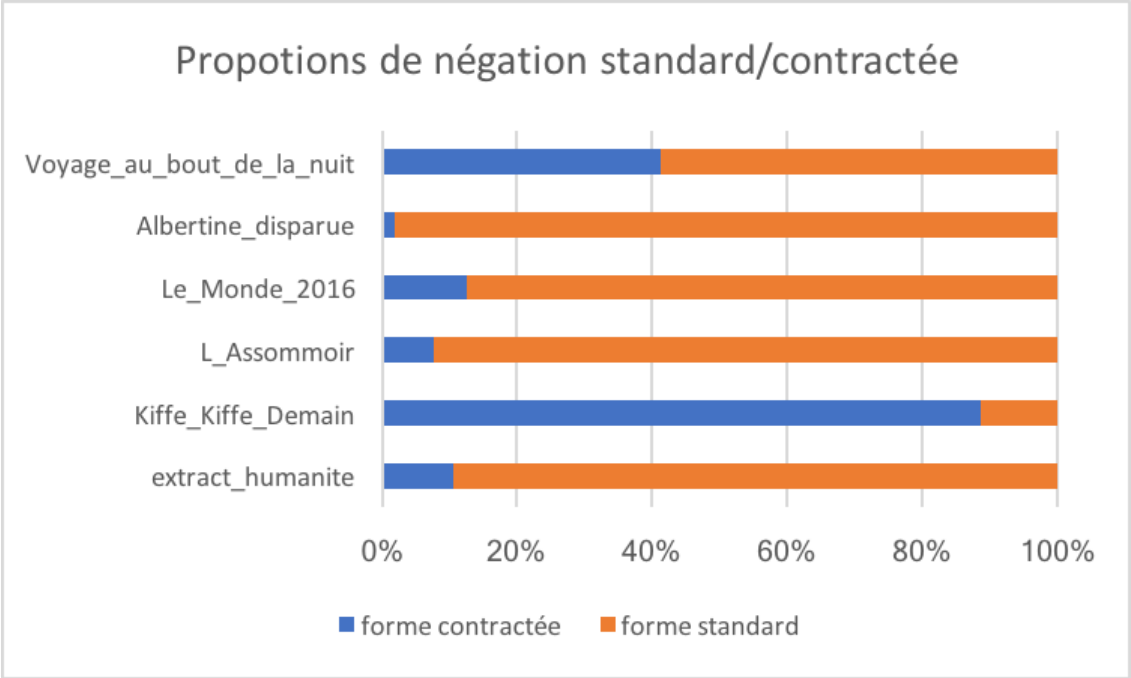


Figure 9 graphique des fréquences relatives des "ça"/"cela" dans chaque œuvre/registre

→ Descripteur nettement confirmé puisque les trois œuvres où la proportion des termes contractés sont les œuvres représentatives du registre familier. De plus, afin de corroborer cette hypothèse nous pouvons remarquer que l'œuvre représentative du soutenu est celle où la forme contractée apparaît le moins souvent.

2.3 - contraction de la forme négative : ne/pas, pas : standard / familier

→ Vous trouverez dans les annexes des fichiers texte brut où sont stockés l'extractions des phrases où les motifs apparaissent.



→ Descripteur confirmé puisque les 3 œuvres où le motif contracté apparaît le plus sont des œuvres associés au registre familier et au contraire l'œuvre où le motif non contracté est le plus présent est l'œuvre représentative du registre soutenu.

2.4 – Sujet « nous » transposé en « nous »

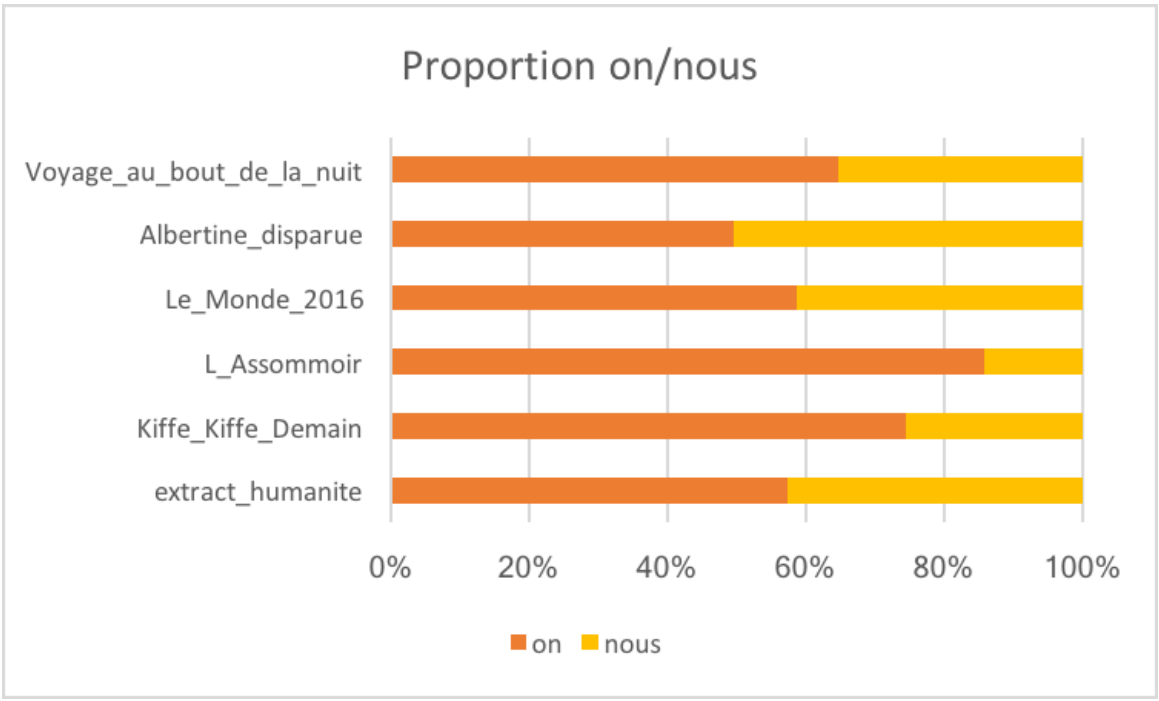


Figure 11 graphique des proportions de "on"/"nous" dans les différents registres

→ Vous trouverez dans les annexes des fichiers texte brut où sont stockés l'extractions des phrases où les motifs apparaissent. Le descripteur est confirmé puisque les trois œuvres avec la plus forte proportion de « on » sont du registre familier et l'œuvre avec la plus grande proportion de « nous »

est du registre soutenu. Nous pourrions regarder la différence de l'utilisation du « on » et du « nous » dans les œuvres associées au genre familier :

L'Assommoir de Zola :

On :

1. 'On la rencontrerait une nuit sur un trottoir, pour sûr.'
2. 'Alors, comme on ne parlait pas toujours de leur mariage, elle voulut s'en aller, elle tira légèrement la veste de coupeau.'
3. 'Tout compte fait, on se trouverait quinze à table.'
4. 'Alors, on revint par les boulevards et par le faubourg poissonnière.'

Nous :

5. 'Je sais où est Lantier… nous avons nos chagrins comme tout le monde, mon dieu ! '
6. '… j'ai quelque chose à laver, je vous garderai une place à côté de moi, et nous causerons.'
7. '— Ca suffit entre nous, madame Gervaise, murmura-t-il.'

Kiffe Kiffe Demain de Guene :

On :

8. "On lui crie après sans arrêt, et on la surveille pour vérifier qu'elle pique rien dans les chambres."
9. "Cette meuf, on dirait qu'elle a besoin d'être heureuse à la place des autres."
10. 'On a reçu des coupons de la caf.'

Nous :

11. 'Une fois, il a dit à ma mère qu'en dix ans de métier, c'était la première fois qu'il voyait " des gens comme nous avec un enfant seulement par famille ".'
12. "Ma mère, elle dit que si mon père nous a abandonnées, c'est parce que c'était écrit."

→ « on » permet au locuteur de rester flou et est souvent associé à des termes qui généralisent le propos. Tandis que « nous » a tendance à identifier des locuteurs précis :

1. « on » + « pour sûr »
2. « on » + « toujours »
8. « on » + « sans arrêt »
9. « on » + maxime au présent de vérité général « elle a besoin d'être heureuse à la place des autres »

5. « Lantier » + « Je » = « nous »
6. « moi » + vous = « nous »
7. « madame Gervaise » + locuteur = « nous »
11. « ma mère » + « ma » (1 personne du SG) = « nous »
12. « ma mère » + « ma » / « mon » (1 personne du SG = « nous »

2.5 – Termes à redoublement

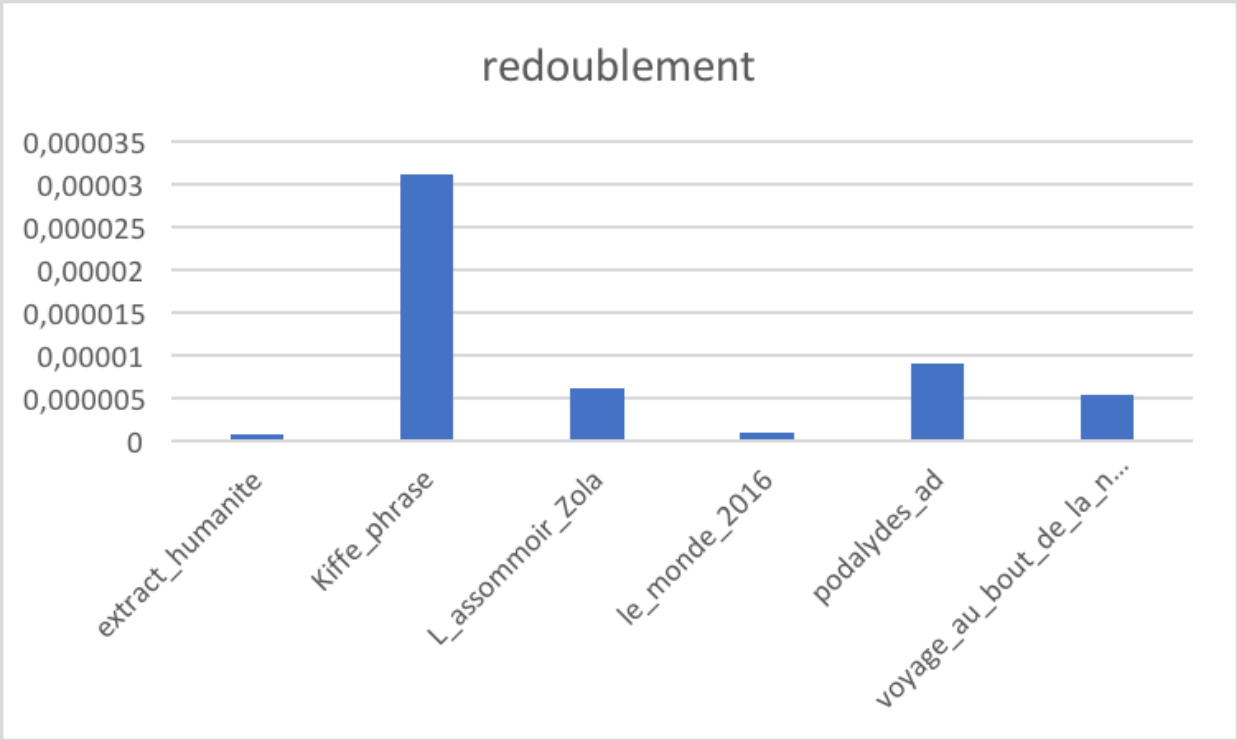


Figure 12 graphe des proportions de terme à redoublement dans chaque œuvre

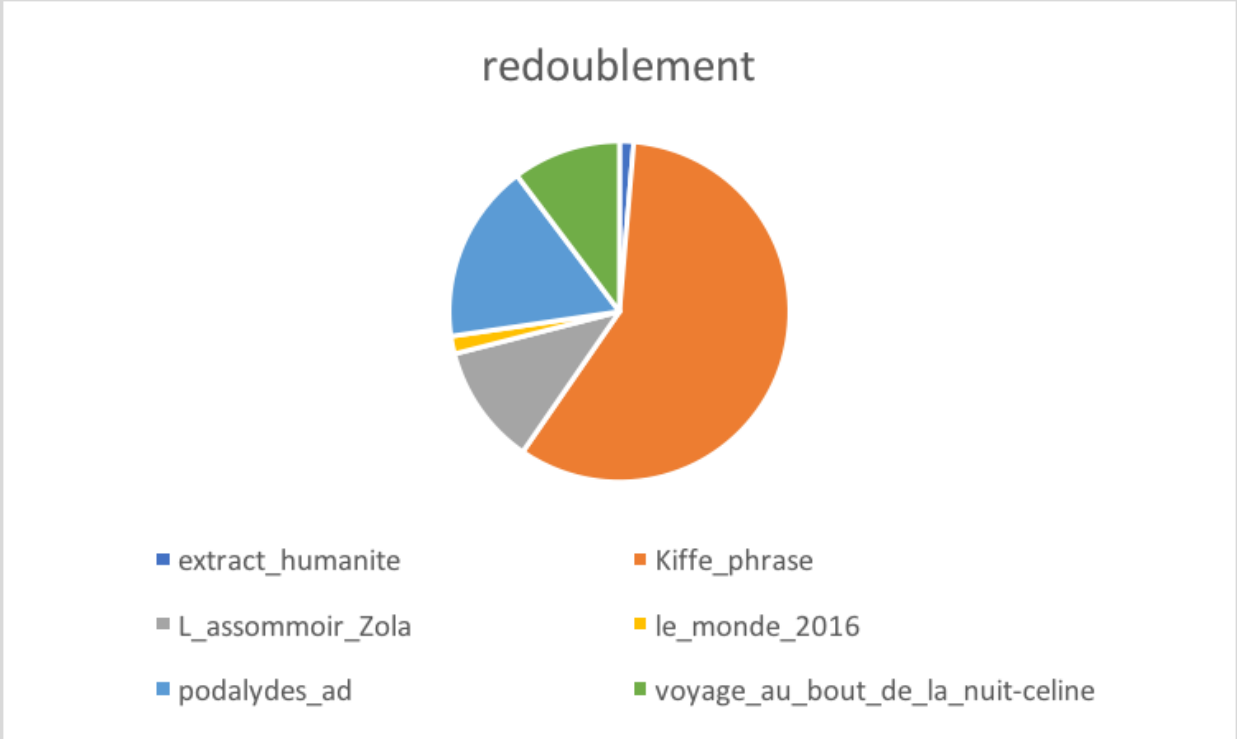


Figure 13 graphe des proportions de terme à redoublement dans chaque œuvre

→ Le script utilisé affiche les différents termes identifiés dans les textes/registre lorsqu'il est lancé.

extract_humanite.txt
['', 'chercher', 'papa', 'II', 'MM', 'PP', 'pp', 'bébé', 'toutou', 'baba', 'nana', 'rentrent', 'mm', 'SS', 'TUTU', 'dodo', 'gogo', 'ss', 'dada', 'kaka', 'BB', 'pépé', 'couscous', 'bonbon', 'chacha', 'TomTom', 'XX', 'MAMA', 'boubou', 'youyou']

kiffe_phrase.txt
['', 'bébé', 'chercher', 'pipi', 'couscous', 'papa', 'caca', 'zizi', 'nana', 'zonzon', 'boubou', 'bonbon', 'dodo']

L_assommoir_Zola.txt
['', 'chercher', 'papa', 'II', 'dodo', 'banban', 'flafla', 'bébé', 'coucou', 'bobo', 'gnangnan', 'joujou', 'pipi', 'coco', 'jeanjean', 'rentrent', 'cancan', 'quiqui', 'dada', 'nounou', 'toutou', 'tata', 'ronron']

le_monde_2016.txt
['', 'MM', 'chercher', 'caca', 'II', 'bébé', 'nounou', 'DimDim', 'mm', 'SS', 'papa', 'tonton', 'ISIS', 'rentrent', 'PP', 'bobo', 'pipi', 'BB', 'LL', 'yéyé', 'gogo', 'll']

podalydes_une_ligne.txt
['', 'chercher', 'papa']

voyage_au_bout_de_la_nuit-celine.txt
['', 'chercher', 'pipi', 'papa', 'zizi', 'coco', 'corocoro', 'caca', 'tamtam', 'gaga', 'bonbon', 'glouglou', 'rentrent', 'ronron', 'bébé', 'mimi']

→ La fréquence des termes à redoublement est bien un descripteur puisque discriminante pour les registres familiers. De plus, les termes trouvés sont pour la plupart des termes familiers voire argotiques. Il est intéressant de noter que les termes peuvent être des entités nommées notamment pour le familier des sortes de surnom « jeanjean », « guigui ». De plus, le détail des extractions me permet « manuellement » de ne pas classer les termes trouvés pour Albertine Disparue dans le registre familier puisque « chercher » est un terme standard, seul « papa » peut être envisagé comme familier/courant.

2.6 – Terminaisons en « asse »

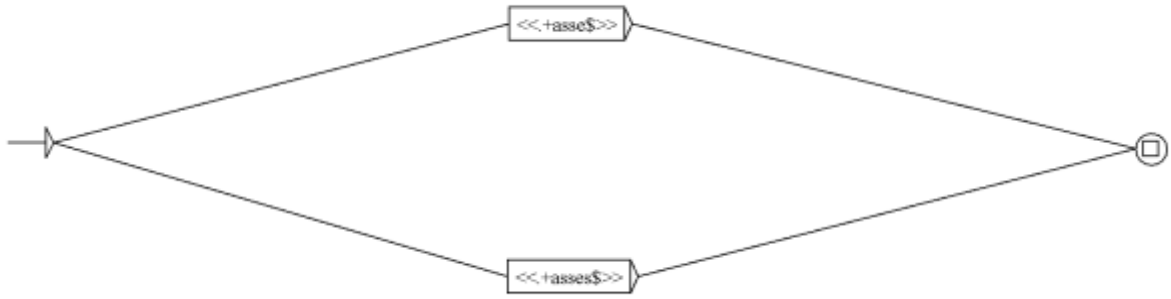


Figure 14 automate pour identifier les terminaisons en "asse"

Le concordancier complet est disponible dans les annexes : 2.6_concord.html.

Contexte dans l'œuvre *Kiffe Kiffe Demain* :

{S}Je suis sûre que Karim la blondasse, elle fréquente des types comme ce M.
{S}Kaba.
c'est pas une fille comme ça qu'il lui faut à Hamoudi parce que maintenant il ressemble à ces mecs trop impeccables avec leur coupe plaquée au gel qui font du porte-à-porte pour vendre des encyclopédies.

- Utilisation d'un article défini, puis reprise du syntagme par le pronom personnel sujet « elle », 1 terme lexical discriminant « type » dans la même phrase.
- Phrase suivante croisement avec 4 descripteurs du familier
    - La forme contractée : « c'est »
    - Une forme négative incomplète
    - La forme contractée : « ça »
    - Un terme lexical discriminant : « mecs »

{S}Un homme qui s'enfuirait pas de L' autre côté de la Méditerranée ou avec une blondasse décolorée sur des talons hauts.

- Dans la même phrase :
    - Une forme négative incomplète

{S}Mais quand elle vient pas, ça reste bien dégueulasse pendant des semaines, comme là ces derniers temps.
{S}Dans l ascenseur, y avait de la pisse et des mol ards, ça sentait mauvais, mais on était quand même contentes que ça marche.
{S}Heureusement qu `on connaît l emplacement des boutons par rapport aux étages, parce que la plaquette est grattée et ça a fondu.
{S}On a dû les brûler au briquet.

- Dans la même phrase :
    - Une forme négative incomplète
    - Forme contractée : « ça »
    - Ponctuant : « là »
- Deuxième phrase :
    - Forme contractée de « il » : « y »
    - Termes lexicaux : « pisse », « mollards »
    - Forte présence de la forme contractée : « ça »
    - Forme contractée : « on »
- Troisième phrase :
    - Forme contractée : « on »
    - Forme contractée : « ça »
- Quatrième phrase :
    - Forme contractée : « on »

{S}Bref, les deux pétasses qui ont dit ça derrière nous, je me suis même pas retournée pour les avaler toutes crues ou leur déchiqueter les narines.

- Dans la même phrase :
    - Forme contractée : « ça »
    - Forme négative incomplète

Contexte dans le contenu extrait de l'*humanité* :

" Nicolas Sarkozy a le mérite de la franchise et d'afficher sa vision partisane de la presse mais ce discours dans la bouche du président de la République est inacceptable et dangereux.{S} Qu'il se satisfasse de la montée en puissance des industriels dans les médias témoigne d'une totale désinvolture sur le sujet. "

- Aucune occurrence d'un descripteur pour le familier
- Le terme lui-même est un verbe discriminant pour le soutenu : subjonctif

---

{S}Valérie Niquet.{S} S'il n'y a pas de réponse aux attentes dans le domaine économique, la déception pourrait déboucher sur une attitude moins favorable à la Chine en ce qui concerne le fameux partenariat stratégique.{S} Exemple : la France souhaite depuis longtemps que la question de la coopération en Afrique en matière de codéveloppement fasse partie du dialogue stratégique.{S} Mais Pékin est réticent.{S} Ce sujet de divergence pourrait être pointé.

- Aucune occurrence d'un descripteur pour le familier
- Le terme lui-même est un verbe discriminant pour le soutenu : subjonctif

Contexte dans *l'Assommoir* :

{S}Elle ne se sentait plus, tant elle était lasse et vide.
{S}La dernière idée nette qui l'occupât, fut que sa garce de fille, au même instant, mangeait peut-être des huîtres.
{S}Ensuite, tout se brouilla, elle resta les yeux ouverts, mais il lui fallait faire un trop grand effort pour penser.
{S}Et la seule sensation qui persistait en elle, au milieu de l'anéantissement de son être, était celle d'un froid de chien, d'un froid aigu et mortel comme jamais elle n'en avait éprouvé.

- Terme lexical seul discriminant pour le soutenu et non le familier
- Dans la même phrase : aucun descripteur pour le familier
- Dans la deuxième phrase :
    - Terme lexicalisé discriminant pour le familier : « garce »
    - Lexie familière : « froid de chien »
- Ce passe est singulier car il mêle un registre soutenu (que nous pourrions également qualifier de littéraire) avec un registre familier. Ce registre familier semble être lié à la présence de discours rapporté (dans le cadre d'un monologue intérieur retranscrit) : le discours indirect libre qui retranscrit les paroles du personnage à travers un style particulier.

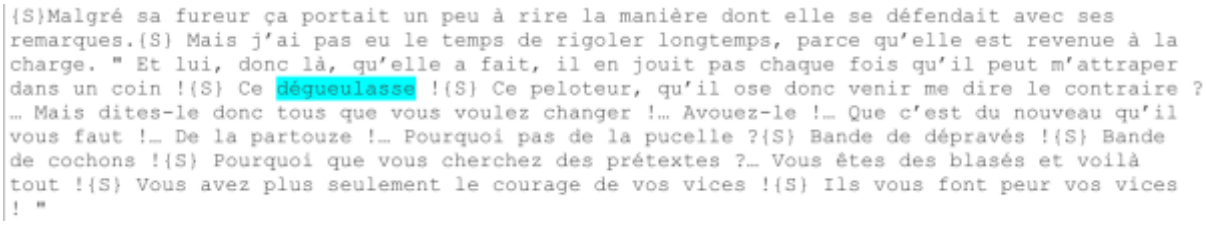
— Moi, je m'abonne, murmura-t-il, pour qu'on me fasse comme ça pipi dans la bouche.

- Terme lexicalisé identifié discriminant pour le registre soutenu : subjonctif bien formé
- Dans la phrase plusieurs autres descripteurs pour le familier :
    - Formes contractées :
        - « on »
        - « ça »
    - Terme lexical discriminant : « pipi »

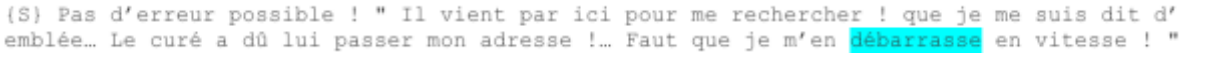
Contexte dans *Voyage au bout de la nuit* :

{S} Mes rares arrivées à table aussi furtives et silencieuses que je m'appliquasse à les rendre prenaient l'ampleur de réels événements.{S} Dès que j'entrais dans la salle à manger, les cent vingt passagers tressautaient, chuchotaient…

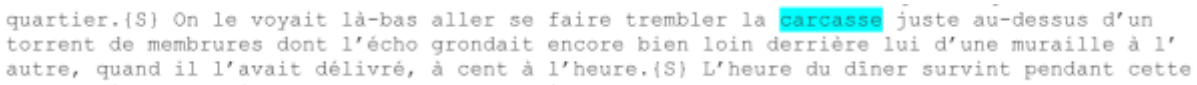
- Terme lexicalisé identifié discriminant pour le registre soutenu : subjonctif bien formé


{S}Malgré sa fureur ça portait un peu à rire la manière dont elle se défendait avec ses remarques.{S} Mais j'ai pas eu le temps de rigoler longtemps, parce qu'elle est revenue à la charge. " Et lui, donc là, qu'elle a fait, il en jouit pas chaque fois qu'il peut m'attraper dans un coin !{S} Ce dégueulasse !{S} Ce peloteur, qu'il ose donc venir me dire le contraire ? … Mais dites-le donc tous que vous voulez changer !… Avouez-le !… Que c'est du nouveau qu'il vous faut !… De la partouze !… Pourquoi pas de la pucelle ?{S} Bande de dépravés !{S} Bande de cochons !{S} Pourquoi que vous cherchez des prétextes ?… Vous êtes des blasés et voilà tout !{S} Vous avez plus seulement le courage de vos vices !{S} Ils vous font peur vos vices ! "

- Terme lexical discriminant conforté par les éléments trouvés en contexte
- Contexte riche en descripteur discriminant pour le familier :
    - Forme contractée : « ça »
    - « Relatives populaires »[20] :
        - « qu'elle a fait » → dans un contexte discours rapporté, plus précisément un discours indirect mal formé dans un discours direct
        - « Pourquoi que vous cherchez des prétextes » → dans un contexte de discours direct
    - Métaphore animale : « Bande de cochons ! »
    - Terminaisons en « ouze » : « partouze »


{S} Pas d'erreur possible ! " Il vient par ici pour me rechercher ! que je me suis dit d' emblée… Le curé a dû lui passer mon adresse !… Faut que je m'en débarrasse en vitesse ! "

- Terme extrait non discriminant pour le registre familier, toutefois le contexte permet de l'identifier comme tel :
    - « Relative populaire » dans un cadre de discours direct où un discours rapporté indirect est mal formé : « que je me suis dit d'emblée… »
    - Absence du pronom impersonnel « il » dans la dernière phrase : Faut que je m'en débarrasse en vitesse ! »


quartier.{S} On le voyait là-bas aller se faire trembler la carcasse juste au-dessus d'un torrent de membrures dont l'écho grondait encore bien loin derrière lui d'une muraille à l' autre, quand il l'avait délivré, à cent à l'heure.{S} L'heure du dîner survint pendant cette

- Terme non discriminant pour le registre familier mais nous avons dans la phrase la forme contractée « on » qui est, elle, discriminante pour le familier.

Aucune forme discriminante pour le texte extrait des archives du *Monde*.

Un examen « manuel » des motifs retournés par Unitex met en exergue des motifs communs qui ne sont pas discriminants autrement dit, du bruit. Afin de dégager les termes discriminant je propose de traiter le corpus de manière comparée : tous les termes communs au 5 textes ne sont pas des termes discriminant malgré leurs terminaisons en « asse », en revanche les termes se terminant par « asse » et qui sont absents de certains textes sont discriminants.

Script :

[20] Françoise Gadet 2003

```
#!/usr/bin/pythin

import re, os, nltk
from nltk import sent_tokenize, word_tokenize

def idTerminaisonsAsse(texte):
    """
    input: un texte
    output: une liste avec les termes se temrinants par "asse(s)"
    """
    texte=open(texte).read()
    texte=texte.lower()
    texteMots=re.split("\W", texte)
    extractTermes=[mot for mot in texteMots if mot[-4:]=="asse"]
    return extractTermes


def idDiscriminant(texte1, texte2, texte3, texte4, texte5, texte6):
    dicoDiscrminant={}
    listeNonDiscriminant=[]
    listeDiscriminant=[]
    mots1=idTerminaisonsAsse(texte1)
    mots2=idTerminaisonsAsse(texte2)
    mots3=idTerminaisonsAsse(texte3)
    mots4=idTerminaisonsAsse(texte4)
    mots5=idTerminaisonsAsse(texte5)
    mots6=idTerminaisonsAsse(texte6)
    listeNonDiscriminant=[m for m in set(mots1) & set(mots2) & set(mots3) & set(mots4) & set(mots5) & set(mots6)]
    listeTous=[ele for ele in set(mots1) | set(mots2) | set(mots3)| set(mots4) | set(mots5) | set(mots6)]
    listeDiscriminant=[set(listeNonDiscriminant) ^ set(listeTous)]
    dicoDiscrminant["non discriminant"]=listeNonDiscriminant
    dicoDiscrminant["discriminant"]=listeDiscriminant
    return dicoDiscrminant
```

Figure 15 script pour identifier les termes discriminant

Résultats :

→ Tous les textes de tous les registres afin d'avoir les termes communs à tous ou bien très spécifiques car présents dans seulement un texte.

{'non discriminant': ['fasse', 'passe', 'basse', 'chasse', 'classe'], 'discriminant': [{'rencontrasse', 'paillasse', 'tirasse', 'refasse', 'dépasse', 'trouvasse', 'terrasse', 'éprouvasse', 'embrasse', 'satisfasse', 'contrebasse', 'liasse', 'hasse', 'repasse', 'pourchasse', 'surpasse', 'brasse', 'carcasse', 'oubliasse', 'blondasse', 'charasse', 'cessasse', 'bécasse', 'rapasse', 'aimasse', 'gardasse', 'calebasse', 'montparnasse', 'détournasse', 'renonçasse', 'mélasse', 'débarrasse', 'impasse', 'ducasse', 'rêvasse', 'jacasse', 'ramenasse', 'annemasse', 'relevasse', 'espinasse', 'fadasse', 'tétasse', 'retrouvasse', 'fracasse', 'degrasse', 'entasse', 'allasse', 'pigasse', 'outrepasse', 'ramasse', 'tracasse', 'chiasse', 'ragouillasse', 'bavasse', 'cuirasse', 'grasse', 'crasse', 'légasse', 'tasse', 'appliquasse', 'embarrasse', 'surclasse', 'montpernasse', 'nasse', 'asse', 'masse', 'cybercasse', 'dégueulasse', 'potasse', 'priasse', 'lavasse', 'cocasse', 'lasse', 'amasse', 'crevasse', 'poufiasse', 'hommasse', 'casse', 'tignasse'}]}

→ nous pouvons croiser moins de texte afin d'avoir des termes discriminants par registre. Si nous utilisons le même script dans le but de comparer les termes uniquement pour le registre familier : cela nous retourne donc les termes avec une très forte valeur discriminante pour un registre spécifique.

{'non discriminant': ['casse', 'classe', 'basse', 'passe', 'chasse', 'fasse'], 'discriminant': [{'tignasse', 'embrasse', 'embarrasse', 'repasse', 'appliquasse', 'lasse', 'carcasse', 'cocasse', 'impasse', 'paillasse', 'tasse', 'chiasse', 'rêvasse', 'dépasse', 'crevasse', 'cuirasse', 'mélasse', 'fadasse', 'terrasse', 'grasse', 'dégueulasse', 'potasse', 'poufiasse', 'priasse', 'tracasse', 'débarrasse', 'ramasse', 'jacasse', 'lavasse', 'montpernasse', 'masse', 'liasse', 'crasse', 'tétasse', 'ragouillasse', 'blondasse'}]}

→ ou bien croiser un texte de chaque registre ("podalydes_balise.txt", "extract_humanite.txt", "kiffe_phrase.txt") afin d'avoir des termes discriminant pour plusieurs registres.

{'non discriminant': ['basse', 'passe', 'classe', 'fasse', 'chasse'], 'discriminant': [{'calebasse', 'nasse', 'entasse', 'bécasse', 'cocasse', 'grasse', 'ramenasse', 'espinasse', 'satisfasse', 'tasse', 'tirasse', 'bavasse', 'hasse', 'embarrasse', 'carcasse', 'pigasse', 'éprouvasse', 'charasse', 'annemasse', 'détournasse', 'débarrasse', 'oubliasse', 'liasse', 'fracasse', 'lasse', 'rencontrasse', 'cuirasse', 'dépasse', 'trouvasse', 'outrepasse', 'blondasse', 'légasse', 'relevasse', 'retrouvasse', 'aimasse', 'contrebasse', 'casse', 'terrasse', 'repasse', 'renonçasse', 'brasse', 'gardasse', 'pourchasse', 'dégueulasse', 'embrasse', 'crasse', 'masse', 'impasse', 'refasse', 'chiasse', 'poufiasse', 'allasse', 'cessasse', 'hommasse', 'montparnasse', 'ramasse', 'amasse'}]}

- Les termes à forte valeurs discriminante sont
  - Les verbes : (soutenu)
    - Aimasse, gardasse, retrouvasse, renonçasse, oubliasse, détournasse, satisfasse, éprouvasse, cessasse, hommasse
      - Imparfait du subjonctif
  - Les noms communs : (familier)
    - Chiasse, poufiasse, déguelasse, blondasse, ragouillasse

2.7 – Terminaisons en « ouze »

→ Utilisation du même script : modification du de la terminaison recherchée.

Résultats :

→ Pour les 6 textes :

{'non discriminant': [], 'discriminant': [{'partouze', 'mirouze', 'gouze', 'douze', 'debbouze'}]}

→ Pour les 3 texte du registre familier :

{'non discriminant': ['douze'], 'discriminant': [{'partouze'}]}

→ Pour 3 textes de 3 registres différents : ("podalydes_balise.txt", "extract_humanite.txt", "voyage_au_bout_de_la_nuit-celine.txt")

```
{'non discriminant': [], 'discriminant': [{'douze', 'gouze', 'partouze', 'mirouze'}]}
```

- Beaucoup moins de bruit : selon les textes donnés en argument de 0 à 1 terme non discriminant.
    - Il est intéressant de noter que le terme non discriminant est trouvé lorsque nous comparons les différentes extractions d'un même registre (cela paraît logique puisque nous croisons des textes du même registre : le degré de discrimination est donc plus élevé).
    - Terme discriminant identifié : « partouze »
    - Les autres termes semblent être des entités nommées : « Mirouze », « Debbouze »

2.8 – Terminaisons en « o »

→ Utilisation du même script : modification du de la terminaison recherchée avec quelques modifications. En effet, je n'ai pas exactement mis en place la même chaîne de traitement : je n'ai pas segmenté le texte sur tous ce qui n'étaient pas des caractères Unicode, et je n'ai pas remplacé toutes les majuscules par des minuscules. Ces changements étaient motivés par plusieurs raisons. Tout d'abord, la terminaison recherchée peut être extraite d'un élément appartenant à un terme tel que « judéo-chrétien » : j'ai donc modifié le split afin qu'il ne segmente pas le texte sur '-' et me retourne « judéo » comme un motif extrait. Ensuite, la première fois que j'ai examiné les résultats retournés par l'algorithme il y avait beaucoup de terme dont je n'arrivais pas à comprendre le sens, puis j'ai compris qu'ils étaient des entités nommées. J'ai donc laissé les majuscules afin de les identifier comme telles.

Résultats :

→ Pour les 6 textes : les motifs retournés étant très nombreux j'ai créé un fichier de sortie « 2.8.1_extractionsMotifsTerminaisonsO.txt » qui stocke les extractions et qui se trouve dans les annexes du rapport.

J'énumère ici quelques motifs discriminant que j'ai relevés « à la main » : métro-boulot-dodo, histoire-géo, l'auto, l'electro, sono, Métro, Yolo (you only live once), conso, psycho, credo, L'ex-numéro, Bobo, toxico, sculpture-auto, socio, frigo, phono, coco, Stylo, incognito[21], porno, Fiasco, philo, d'aristo, vidéo, Secundo[22]…

Nous trouvons également beaucoup de termes italiens et d'entités nommées.

[21] Egalement un emprunt Françoise GADET 2003
[22] Ibidem

→ Sur les 3 textes du registre familier :

{'non discriminant': ['numéro'], 'discriminant': [{'studio', 'dingo', 'L'écho', 'Bernardo', 'Toto', 'Bambola-Fort-Gono', 'Corocoro', 'fluo', 'Cluedo', 'mytho', 'Francisco', 'rigolo', 'd\x92Antonio', 'halo', 'Rio', 'Ho', 'maso', 'écho', 'photo', 'Fort-Gono', 'Momo', 'dodo', 'géo', 'do', 'Brasero', 'psycho', 'stylo', 'boléro', 'bravo', 'polo', 'DiCaprio', 'Congo', 'l'apéro', 'No', 'coco', 'd'auto', 'Leonardo', 'Hebdo', 'cacao', 'bobo', 'cato', 'cargo', 'vidéo', 'scénario', 'Casino', 'Enrico', 'euro', 'Bikomimbo', 'phono', 'l\x92apéro', 'moto', 'l'adagio', 'aristo', 'kimono', 'Ornano', 'proprio', 'parano', 'l'écho', 'frigo', 'Zozo', 'corocoro', 'Bravo', 'micro', 'Stabilo', 'brasero', 'Topo', '"Bravo', 'auto', 'Togo', 'judo', 'caraco', 'd'aristo', 'ado', 'piano', 'Twingo', 'Niro', 'vélo', 'kilo', 'porto', 'Zorro', 'métro'}]}

Les résultats sont les mêmes avec quelques termes très discriminants tels que « mytho », « dodo », « rigolo », « l'apéro », « maso », « parano »… puis beaucoup d'entités nommées.

→ 3 textes de 3 registres différents : tous les motifs retournés sont consultables dans les annexes sous le fichier « 2.8.2_extractionsMotifsTerminaisonsO.txt ».

Voici les termes discriminant après une sélection manuelle :

- Tous non discriminant sont : ['numéro', 'piano', 'photo']
- Les discriminant relevés : gogo, Mélo, hebdo, Véto, info, grosso, Pro, banco[23], intello, prolo, d'Agro, bistro, éco, contrario[24], collabo, gigolo, topo, l'expo, promo…

Le descripteur permet donc plusieurs choses :

1. Identifier des termes discriminants pour le registre familier
2. Identifier des entités nommées
3. Identifier des emprunts de langue également discriminant pour le familier

→ Croisement de 2 descripteurs : terminaisons en « o » et emprunt à une langue étrangère

2.9 – Terminaisons en « ou »

→ Réutilisation du script en modifiant la terminaison et en gardant seulement les majuscules (je segmente à nouveau sur les caractères non Unicode).

→ Sur les 6 textes des 3 registres :

{'non discriminant': ['fou', 'ou'], 'discriminant': [{'Baldeyrou', 'Camou', 'Fou', 'chou', 'Sardou', 'Hangzhou', 'Kontomanou', 'Cassou', 'Zouzou', 'tourlourou', 'Abdou', 'Guillou', 'Doudou', 'Danielou', 'cou', 'tabou',

[23] Egalement un emprunt Françoise GADET 2003
[24] Ibidem

'zoulou', 'You', 'Amadou', 'Bénichou', 'assou', 'bijou', 'Rotrou', 'Babou', 'Benamou', 'Sanou', 'Zabou', 'nou', 'Mariétou', 'Manaudou', 'Cheffou', 'Herrou', 'Pirou', 'Bou', 'Oumou', 'Nou', 'Azou', 'Férou', 'prou', 'Beitenou', 'Cariou', 'Zou', 'toutou', 'mou', 'Bayrou', 'marlou', 'Dahou', 'Boussou', 'Kitsou', 'Ahoyansou', 'Genou', 'andalou', 'Galledou', 'Beidou', 'Hibou', 'Berrou', 'Mou', 'Larrouturou', 'doulou', 'Cotonou', 'Neimatou', 'Kindorou', 'caribou', 'Tatihou', 'Corfou', 'flou', 'Amsatou', 'Treilhou', 'Loulou', 'vou', 'Netanyahou', 'Remanjou', 'matou', 'Guigou', 'Lanaudou', 'moscou', 'grisou', 'Agbegnenou', 'Nouadhibou', 'genou', 'Abeilhou', 'Crocolou', 'Pamboutzoglou', 'joujou', 'Lou', 'Cachou', 'toubou', 'Kathmandou', 'Bintou', 'patrou', 'nounou', 'Trou', 'Hamidou', 'acajou', 'Isou', 'Kiriakou', 'verrou', 'pou', 'relou', 'Boutou', 'Bijou', 'Lancou', 'Issoufou', 'Kourou', 'Papandréou', 'Ouagadougou', 'Cheyrou', 'Yangassou', 'Zinsou', 'Ghizou', 'Badiou', 'Carriou', 'Suzhou', 'Mabanckou', 'gou', 'Salou', 'écrou', 'Poitou', 'Mamoudzou', 'Kalidou', 'boubou', 'Ndoumou', 'guilledou', 'you', 'walou', 'Miou', 'Daoudou', 'Jacqou', 'Picaudou', 'hindou', 'Youssou', 'Piriou', 'Chikirou', 'Bouzou', 'Pokou', 'Anjou', 'Follorou', 'abou', 'Hou', 'Chamayou', 'Sémou', 'Gammou', 'Chou', 'Moscou', 'cachou', 'hou', 'Massou', 'Badarou', 'tou', 'Kabou', 'Anbou', 'bambou', 'Kirikou', 'Badrou', 'Vishnou', 'lou', 'Hukou', 'Quiniou', 'manitou', 'Troudou', 'Abou', 'Nétanyahou', 'Peyniblou', 'Tayou', 'Zenabou', 'Vassou', 'Azzemou', 'caillou', 'Sassou', 'Carquefou', 'Guangzou', 'Marcou', 'Douzou', 'gourou', 'Montesquiou', 'Riou', 'Xenou', 'bagou', 'trou', 'Guerdjou', 'Seydou', 'glouglou', 'Picsou', 'Mamadou', 'Spirou', 'Bakou', 'Duigou', 'voyou', 'sadhou', 'Mamoudou', 'Barou', 'Afflelou', 'Daou', 'Flou', 'Thanou', 'Tombouctou', 'coucou', 'Semou', 'Changzhou', 'Katmandou', 'Benhamou', 'Poutou', 'Cou', 'Chatou', 'Peyrou', 'Wenzhou', 'rou', 'dou', 'Garikou', 'grigou', 'Bascou', 'Pompidou', 'Moguerou', 'Pérou', 'Benchicou', 'Priou', 'Govou', 'youyou', 'sou', 'Milou', 'Sekou', 'Ahmédou', 'Oubrou', 'Mansoibou', 'amadou', 'bisou', 'clou', 'Datou', 'Caillou', 'Issindou', 'chelou', 'Zizou'}]}

→ Avec les 3 textes du registre familier :

{'non discriminant': ['voyou', 'trou', 'ou', 'fou'], 'discriminant': [{'walou', 'Remanjou', 'acajou', 'grigou', 'tourlourou', 'Caillou', 'clou', 'Mamadou', 'nounou', 'Pérou', 'sou', 'amadou', 'manitou', 'cou', 'marlou', 'coucou', 'guilledou', 'mou', 'dou', 'glouglou', 'Trou', 'bagou', 'chelou', 'cachou', 'Sardou', 'Moscou', 'toutou', 'relou', 'caillou', 'matou', 'Fou', 'boubou', 'Mou', 'verrou', 'bambou', 'chou', 'patrou', 'Zouzou', 'you', 'genou', 'joujou', 'Vassou'}]}

→ 3 textes de 3 registres différents :

{'non discriminant': ['fou', 'ou', 'cou'], 'discriminant': [{'doulou', 'grisou', 'Datou', 'nou', 'Pompidou', 'Dahou', 'Cassou', 'Nouadhibou', 'Pamboutzoglou', 'Picsou', 'Baldeyrou', 'hindou', 'Kontomanou', 'Tatihou', 'Benhamou', 'Hou', 'Zouzou', 'Carquefou', 'Poitou', 'Ouagadougou', 'Badarou', 'abou', 'Boussou', 'Sardou', 'Bou', 'chou', 'Zizou', 'Kindorou', 'Genou', 'mou', 'clou', 'Lanaudou', 'écrou', 'Cariou', 'Danielou', 'caillou', 'Azzemou', 'vou', 'rou', 'Manaudou', 'zoulou', 'voyou', 'Benamou', 'Zabou', 'Amsatou', 'Anbou',

'Kathmandou', 'verrou', 'Crocolou', 'Guangzou', 'Hibou', 'prou', 'glouglou', 'Lou', 'Cotonou', 'Peyrou', 'lou', 'Amadou', 'Montesquiou', 'acajou', 'Pérou', 'boubou', 'Yangassou', 'sou', 'you', 'Troudou', 'Gammou', 'youyou', 'Vishnou', 'Govou', 'Chikirou', 'Boutou', 'dou', 'Riou', 'moscou', 'Bintou', 'Kiriakou', 'toutou', 'cachou', 'Lancou', 'Miou', 'Guillou', 'Jacqou', 'Badiou', 'Flou', 'Mamadou', 'Picaudou', 'Cou', 'Sassou', 'bisou', 'trou', 'Isou', 'tabou', 'flou', 'Anjou', 'Xenou', 'Netanyahou', 'Vassou', 'Babou', 'Treilhou', 'Oumou', 'Bénichou', 'Guigou', 'Tayou', 'tou', 'gourou', 'Sanou', 'genou', 'marlou', 'manitou', 'Férou', 'Cheyrou', 'Piriou', 'gou', 'Bijou', 'andalou', 'Galledou', 'Berrou', 'Kirikou', 'Sémou', 'pou', 'Bayrou', 'Quiniou', 'Massou', 'Thanou', 'Issindou', 'Spirou', 'Abou', 'sadhou', 'Ahoyansou', 'Duigou', 'Kalidou', 'Mabanckou', 'Youssou', 'bambou', 'assou', 'Seydou', 'Carriou', 'Semou', 'Mariétou', 'Mansoibou', 'Doudou', 'bijou', 'Priou', 'Mou', 'Moguerou', 'Ahmédou', 'hou', 'You', 'Beidou', 'Milou', 'Benchicou', 'Daou', 'Moscou', 'Fou', 'Zou', 'Douzou'}]}

- Il est intéressant de remarquer que les termes discriminants sont également des termes à redoublement : « zouzou », « glouglou », « youyou », « coucou », « boubou », « nounou », « toutou », « joujou », « doudou ».
    - Croisement ici de 2 descripteurs pour le familier : redoublement // terminaisons en « ou ».
- Et des termes avec une inversion de l'ordre[25]: « chelou », « relou »
    - Croisement ici de 2 descripteurs pour le familier : une inversion de l'ordre // terminaisons en « ou »

2.10 – Terminaisons en « iotte »

→ utilisation du même script avec une modification de la terminaison à identifier.

Résultats :

→ Avec les 6 textes des 3 registres :

{'non discriminant': [], 'discriminant': [{'griotte', 'chiotte', 'fiotte'}]}

→ 3 textes du registre familier :

{'non discriminant': [], 'discriminant': [{'chiotte'}]}

→ 3 textes de 3 registres différents : "podalydes_balise.txt", "extract_humanite.txt", "kiffe_phrase.txt"

[25] Verlan - Gadet 1997

```
{'non discriminant': [], 'discriminant': [{'chiotte', 'griotte'}]}
```

- Beaucoup moins de termes extraits qu'avec les autres terminaisons
- Tous de même nature : nom commun
- Nous pouvons confirmer ce descripteur puisque les 6 textes incrémentés donnent 2 termes discriminants sur 3 retournés.

2.11 – Adverbes dérivés d'un nom

Ce descripteur représente une difficulté semblable à celle posée par l'approche lexicale. Effectivement, n'étant pas des termes « standards » ils varient selon le contexte et le locuteur : nous ne pouvons pas faire une liste exhaustive des termes à chercher. A titre indicatif, j'ai testé une identification du terme « vachement » souvent utilisé dans le registre familier afin de comparer sa fréquence relative selon les différents textes/registres.

```
def idVachement(texte):
    texte=open(texte).read()
    texte=texte.lower()
    texte=re.split("\W", texte)
    extract=[mot for mot in texte if mot == "vachement"]
    return len(extract)/len(texte)

def tousFichier(dossier):
    dicoOut={}
    for fichier in os.listdir(dossier):
        if fichier[0]==".":continue
        texte=os.path.join(dossier, fichier)
        dicoOut[texte]=idVachement(texte)
    return dicoOut

print (tousFichier("corpus"))
```

Figure 16 script pour calculer la fréquence relative de l'adverbe "vachement"

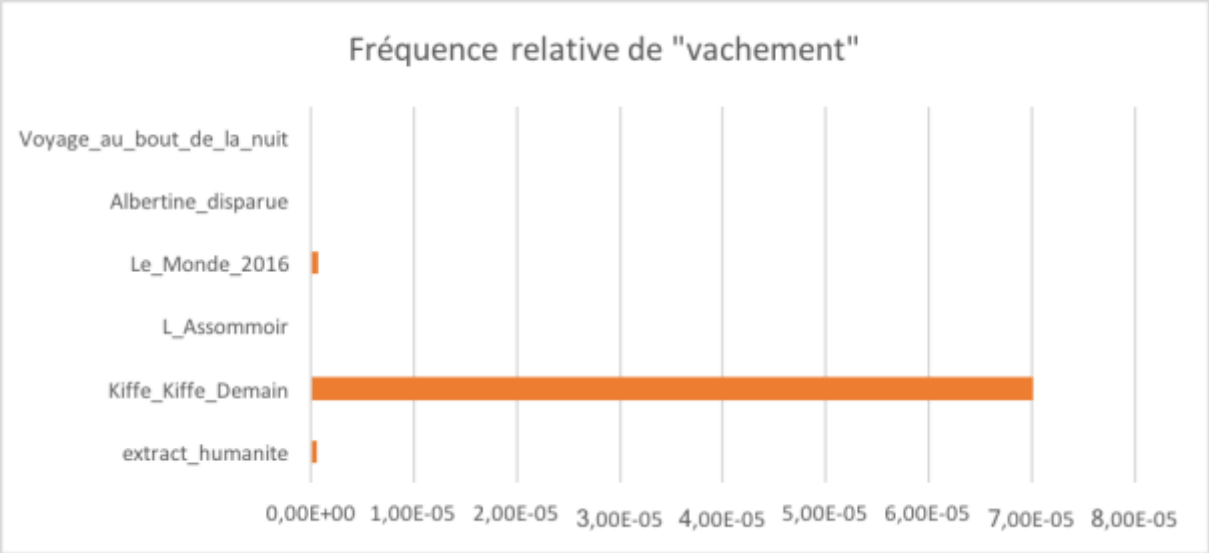


Figure 17 graphe de la présence du terme "vachement" dans chaque texte/registre

- Le descripteur se confirme nettement par sa très forte présence dans l'œuvre *Kiffe Kiffe Demain* associée au registre familier.
- Ce descripteur lexicalisé (« vachement ») peut être envisagé comme un terme « récent » car il n'a aucune occurrence dans les œuvres plus datées : *Albertine disparue* 1925, *L'Assommoir* 1877, *Voyage au bout de la nuit* 1932. En revanche, il et présent (même faiblement) dans les archives du *Monde* et du journal l'*Humanité* : 2016.

2.12 – syntagme : c'est + PL

→ Les motifs extraits sont disponible dans les annexes dans le fichier « 2.11_extractCestPL.txt ».

```
def cestPL(dossier):
    for fichier in os.listdir(dossier):
        if fichier[0]==".": continue
        print ("\n\n_____"+fichier)
        # if fichier[3:]!="txt":continue
        texte=open(os.path.join(dossier,fichier)).read()
        texte=texte.lower()
        texte=re.split("\W", texte)
        nbr_mots=len(texte)
        liste=[]
        for i in range(len(texte)-4+1):
            quatregrammes=texte[i:i+4]
            if quatregrammes[0]=="c":
                milieu=quatregrammes[1]
                if milieu=="est":
                    milieu_fin=quatregrammes[2]
                    if milieu_fin[-1:]=="s":
                        fin=quatregrammes[3]
                        if fin[-1:]=="s":
                            # print (quatregrammes)
                            liste+=[quatregrammes]
        print ("\n",liste, "\n\n",len(liste)/nbr_mots)
```

Figure 18 script pour identifier le syntagme "c'est + PL" et calculer sa fréquence relative dans chaque texte/registre

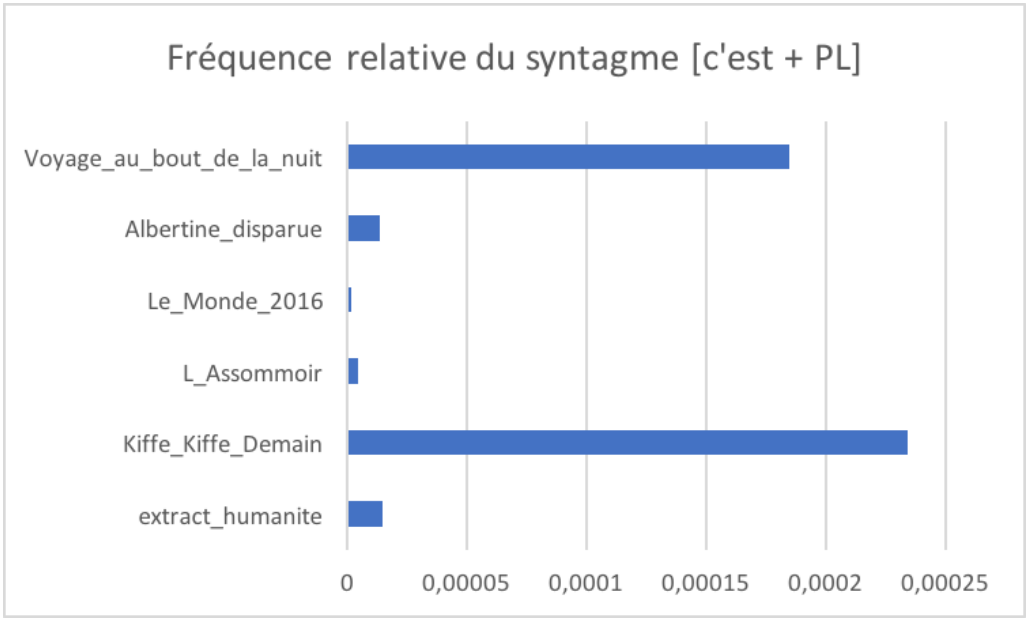


Figure 19 graphe de la présence du syntagme "c'est + PL" par texte/registre

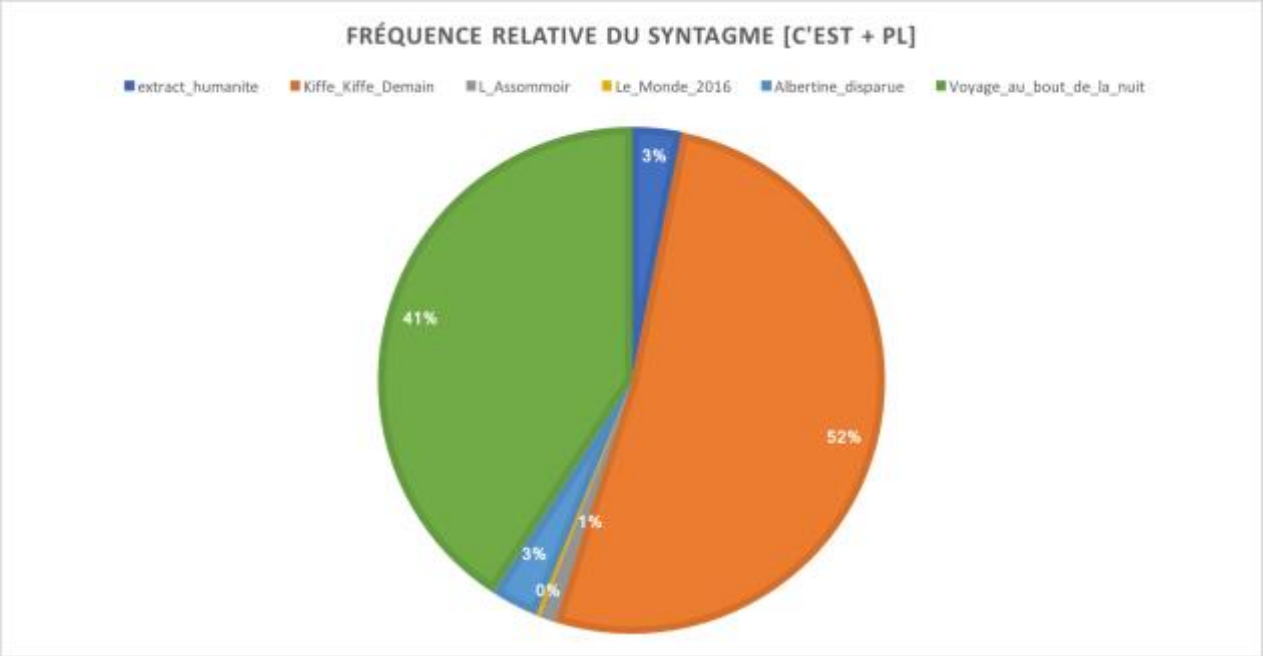


Figure 20 comparaison des fréquences relatives du syntagme entre les différents textes/registres

- Le descripteur est nettement confirmé avec un net écart de fréquence entre les deux œuvres familières et les autres registres.

2.13 – ça + VB

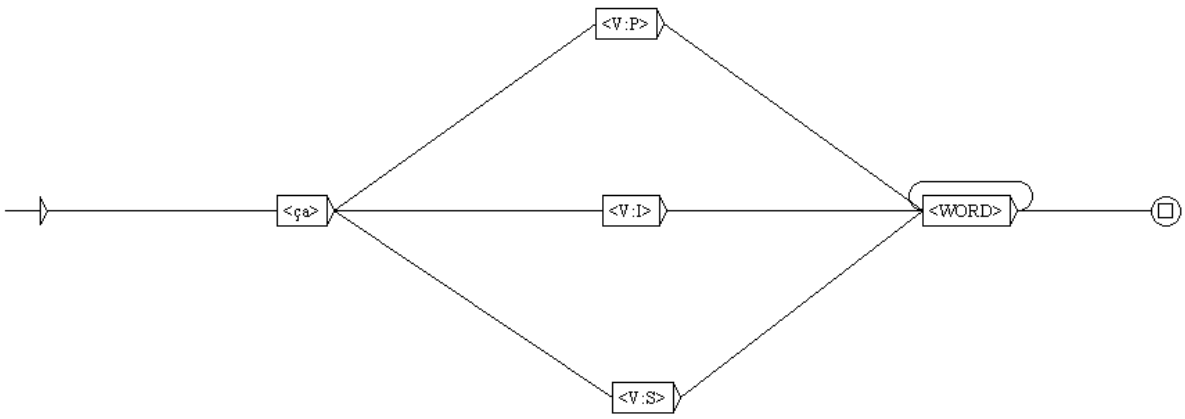


Figure 21 automate pour identifier le syntagme "ça" + VB

Albertine disparue

→ Tous les concordanciers sont dans les annexes.

$ je ne me mêle pas à tout ça;(S) si ça amuse le pauvre Swann de faire des bêtises et de ruiner son existence
igner sa pitié. $ (S)Et elle ajouta: «Ça devait arriver, elle était trop heureuse, la pauvre,
$ (S)Ça ne fait mal à personne, et si ça peut leur faire plaisir à ces bonnes gens, ce n'est
me prend pas avec ces choses-là, tout ça peut très mal finir, je les laisse se débrouiller.»
faisait très chaud déjà là-bas et que ça tapait dur même sous les arbres, restaient dans l'he

je ne me mêle pas à tout ça;(S) si ça amuse le pauvre Swann de faire des bêtises et de ruiner son existence, c'est son affaire, mais on ne me prend pas avec ces choses-là, tout ça peut très mal finir, je les laisse se débrouiller.» C'est le _Suave mari magno_ que Swann lui-même me conseillait à l'égard des Verdurin, quand il avait depuis longtemps cessé d'être amoureux d'Odette et ne tenait plus au petit clan. $

- Contexte de prise de parole explicite :
  - Plusieurs marqueurs d'oralité :
    - Première personne du singulier « je » et les pronoms personnels « me »
    - Formes contractées :
      - « ça »
      - « on »

(S)Et elle ajouta: «Ça devait arriver, elle était trop heureuse, la pauvre, elle n'a pas su connaître son bonheur.» Que le jour est lent à mourir par ces soirs démesurés de l'été! $
(S)Un pâle fantôme de la maison d'en face continuait indéfiniment à aquareller sur le ciel sa blancheur persistante. $

- Contexte de prise de parole explicite :
  - verbe de parole : « ajouta »
  - ponctuation : « : »
  - marqueurs typographiques : guillemets

(S)Ça ne fait mal à personne, et si ça peut leur faire plaisir à ces bonnes gens, ce n'est pas moi qui y trouverai à redire! $
en quoi ça peut-il me gêner? $
(S)Comme je ne fréquenterai jamais la fille d'une femme qui a fait parler d'elle, elle peut bien être marquise long comme le bras pour ses domestiques. $
(S)Mais dans les actes de l'état civil ce n'est pas la même chose. $
(S)Ah! $

- Contexte de prise de parole implicite :
  - Première personne du singulier : « moi », « me », « je »
  - D'autres formes contractées : « ça »
  - Onomatopées : « Ah ! »

je ne me mêle pas à tout ça;(S) si ça amuse le pauvre Swann de faire des bêtises et de ruiner son existence, c'est son affaire, mais on ne me prend pas avec ces choses-là, tout ça peut très mal finir, je les laisse se débrouiller.» C'est le _Suave mari magno_ que Swann lui-même me conseillait à l'égard des Verdurin, quand il avait depuis longtemps cessé d'être amoureux d'Odette et ne tenait plus au petit clan. $

- Contexte de prise de parole explicite :
  - Marqueurs typographiques : ‘ » ‘
  - Première personne du singulier : « je », « me »
  - Frome contractée : « on »

(S)Puis la blanchisseuse amenait ses petites amies et elles se baignaient et après, comme il faisait très chaud déjà là-bas et que ça tapait dur même sous les arbres, restaient dans l'herbe à se sécher, à se caresser, à se chatouiller, à jouer. $
(S)La petite blanchisseuse m'a avoué qu'elle aimait beaucoup à s'amuser avec ses petites amies, et que voyant Mlle Albertine qui se frottait toujours contre elle dans son peignoir, elle le lui avait fait enlever et lui faisait des caresses avec sa langue le long du cou et des bras, même sur la plante des pieds que Mlle Albertine lui tendait. $
(S)La blanchisseuse se déshabillait aussi, et elles jouaient à se pousser dans l'eau;(S) ce soir-là elle ne m'a rien dit de plus, mais, tout dévoué à vos ordres et voulant faire n'importe quoi pour vous faire plaisir, j'ai emmené coucher avec moi la petite blanchisseuse. $

- Contexte de narration non de prise de parole

→ très peu d'occurrence du motif sur l'œuvre et lorsqu'il apparaît cela est majoritairement dans un contexte de prise de parole avec plusieurs indices explicites : verbe de parole, marqueurs typographiques, personne.

L'Humanité

{S}Juste en face de la mosquée, la librairie « Les millefeuilles », un des lieux prisés par les intellectuels de la capitale.{S} Ici, le 5 décembre, Wassila Tamzali, a dédicacé son dernier livre Une éducation algérienne (éditions Gallimard). « Ça a bien marché, il y avait beaucoup de monde et même des attachés culturels des ambassades de France et d'Italie », explique le libraire.{S} Les gens y viennent sans donner l'impression de prêter attention à tous ces SDF installés dans la rue. « C'est grave, on en est bien conscient, mais on ne fait plus attention.{S} On vit dans une société à deux vitesses », explique Djamal Mati, écrivain. {S}À moins d'une centaine de mètres, la rue Didouche, c'est un autre monde.{S} Une circulation dense, des jeunes en rollers zigzaguant au milieu des véhicules conduits par des automobilistes pressés, des boutiques de mode, de luxe, des cafés, des bars et des restaurants, des librairies.{S}.. Des femmes vêtues à la mode ou portant le voile islamique « branché ».{S} Ici, pas de SDF, ni de mendiants, ni de marchands à la sauvette.{S} Mais des policiers des unités antiterroristes, armes au ceinturon.

- Contexte de prise de parole explicite :
    - Marqueurs typographiques : guillemets

{S}Niemeyer est de ces rares architectes pour qui la beauté de l'oeuvre n'est pas dans le seul emballage !{S} Certes l'apparence extérieure de ses oeuvres est importante, mais c'est de l'intérieur, dans le ventre, qu'on peut sans doute y sentir que c'est une conception de la vie, de la circulation, de la communication entre les vivants qu'il s'agit dans son architecture.{S} Les immeubles de Niemeyer, ce ne sont pas des boîtiers empilés d'appartements derrière une belle façade. « Quand un architecte projette un édifice et regarde ses dessins sur la planche, il voit le plan projeté comme une oeuvre déjà construite.{S} En transe, s'il a l'amour du projet, il y pénètre avec curiosité, examinant formes et espaces libres et considère les lieux en pensant là à un tableau, une sculpture ou simplement un dessin noir et blanc... » Niemeyer est ingénieur architecte.{S} Il pense à « comment ça doit fonctionner ».{S} Ce qui doit fonctionner, c'est la vie humaine collective. « La vie, c'est un souffle », aime-t-il dire. « Quand je vois un colibri par exemple, c'est tellement joli, tellement bien fait, on dirait un truc pour un concours, ce mouvement des ailes.{S} C'est un mystère. »

- Contexte de prise de parole explicite :
    - Marqueurs typographiques : guillemets
    - Verbe de parole : « pense »

{S}Alors que le Figaro hier consacrait une page de pub au site Internet de Kadhafi, pour Sarkozy, celui-là même qui a vu l'ensemble des patrons des médias se réjouir de sa victoire au Fouquet's, qui a pu s'en remettre sur le yacht de Bolloré, qui a rapidement placé ses hommes et qui s'est payé le luxe d'annoncer aux journalistes des Échos l'arrivée de Nicolas Beytout à leur tête, « la presse est globalement de gauche, non pas socialiste, non pas partisane, mais culturellement à gauche ».{S} Et, dans une interview aujourd'hui au Nouvel Obs, d'oser, en citant le « soutien » de l'ancien patron du Monde à Ségolène Royal : « La liberté de la presse, c'est le soutien de la gauche.{S} Quand on me soutient, ça devient la mainmise des médias. » Au sommet de sa forme, le locataire de l'Élysée qui durant sa campagne fournissait lui-même les images aux télés et avait donné comme mot d'ordre à son staff de communicants d'inonder les rédactions, assure que « la presse, globalement, a été opposée à ma candidature.{S} C'est son droit.{S} Dire qu'elle m'a aidé, c'est à exploser de rire ».{S} Et comme plus c'est gros, plus ça passe, il ajoute : « Quant à TF1 que dirige Martin Bouygues, l'un de mes meilleurs amis, qui peut dire que cette chaîne a fait mon élection ? » Échange de bons procédés, il reprend l'argumentaire de son « frère » Arnaud Lagardère : « Réjouissez-vous que des industriels investissent dans la presse. » C'est vrai que ça peut toujours être pire.{S} Et de conclure : « Si je suis le seul risque pour l'indépendance de la presse, vous pouvez dormir tranquille. » Parle-t-il là seulement aux journalistes ?

- Contexte de prise de parole explicite :
    - Marqueurs typographiques : les guillemets
    - Verbe de parole : « citant »

et des situations.{S} Exemple : la rencontre du héros, loser trentenaire, avec l'héroïne, jeune bourgeoise casée, au rayon papier toilette d'un supermarché.{S} Mais malgré de beaux efforts pour se hisser à la hauteur des équivalents américains, qui ménagent séduction et humour trash, ça manque encore de culot.{S} Ce film n'est pas une pipe.
{S}Vincent Ostria
culture
{S}Petit écran

- Contexte de prise de parole implicite

{S}Abîme.{S} L'autre soir, un ancien conseiller en communication _ de Lionel Jospin à Matignon et de Catherine Tasca à la Culture _ nous disait : « Sarkozy nous impose une véritable révolution de la pratique politique, une rupture consommée.{S} Tout est pensé pour une spectacularisation d'une seule personne, quoi qu'il fasse.{S} C'est un soap opera plus politique qu'on ne le croit, à épisodes successifs.{S} Et ça va vite, trop vite pour nous.{S} Nous avons à peine le temps d'analyser qu'il est déjà passé au feuilleton suivant avec la vélocité de celui qui ose tout et n'importe quoi ! » N'importe quoi ?{S} Mais que nous disent les faits ?{S} Qu'il est prêt à tout pour qu'on parle de lui, ce qui nous prive parfois de parler du reste.{S} Quitte à se trouver au bord du précipice en permanence.{S} Qu'importe.{S} Lui s'emploie, adoubé par Bush et le pape, à jeter dans l'abîme un modèle (certes passé et défaillant lui aussi) qui plaçait quand même l'horizon social au coeur des relations humaines.

- Contexte de prise de parole explicite :
  - Verbe de parole : « disait »
  - Marqueurs typographiques : les guillemets

→ Grand nombre d'occurrence et une forte tendance à apparaître avec du discours rapporté

Le Monde

maintenu la « journée d'action » prévue mardi.{S} Droit du travail : l'ex-conseiller de Myriam El Khomri explique pourquoi il claque la porte. « D'ordinaire, un conseiller ministériel, petite main de l'ombre, ça ferme sa gueule. » Pierre Jacquemain explique au « Monde » son départ du ministère, face à une réforme qu'il qualifie de « trahison historique »{S} Marine Le Pen parle aux agriculteurs... en patates.{S} Par Martin Vidberg.{S} En Centre-Val-de-Loire, le président PS promet moins mais plus ambitieux... et coûteux.{S} Alors qu'il a tenu 57 % des 40 promesses du mandat 2010_2015, le président PS de la région s'est contenté d'en formuler seulement une trentaine, ont constaté les étudiants en journalisme de l'EPJT.{S} Réforme du droit du travail : quelles concessions le gouvernement est-il prêt à faire ?{S}. Face à la contestation qui s'étend sur le front politique aussi bien que syndical, Matignon se donne quinze jours pour opérer des réajustements.{S} Droit du travail : Hollande en opération déminage d'une réforme décriée.{S} En visite dans une start-up, le président de

- Contexte de prise de parole explicite :
  - Verbe de parole : « explique »
  - Marqueurs typographiques : guillemets
- Descripteur doublé d'un second descripteur :
  - Terme lexical familier voire vulgaire : gueule

».{S} Sortons de l'économisme pour penser notre « commun ».{S} Interrogé pendant une heure sur Europe 1, le président de la République a défendu ses réformes économiques et martelé que « ça va effectivement mieux pour la France ».{S} Nicolas Sarkozy : « Le débat sur le "Brexit" est une opportunité pour refonder l'Europe ».{S} Les routiers inaugurent mardi une semaine de grèves et de manifestations après le passage en force du projet de loi travail.{S} Ce que propose la réforme pour une « justice du XXIe siècle ».{S} Chômage, croissance, pouvoir d'achat… Si François Hollande se félicite du retour au vert de nombreux indicateurs économiques, ceux-ci n'augurent pas encore une reprise solide de l'économie.{S} Un proche de Serge Dassault condamné à quinze ans de prison pour tentative d'assassinat.{S} Le ministre de l'intérieur a apporté mercredi son « soutien total » aux policiers, qui manifestaient au même moment contre la « haine anti-flics ».{S} Enquête pour tentative d'homicide volontaire après l'incendie d'une voiture de police à Paris.{S} Après deux mois de violences en marge des manifestations contre la loi travail, les policiers se sont rassemblés, mercredi 18 mai, dans une soixantaine de villes.{S} Voici les témoignages de certains d'entre eux.{S} A Mayotte, la

- Contexte de prise de parole explicite :
  - Verbe de parole : « a défendu (…) et a martelé »

- Marqueurs typographiques : guillemets

Voyage au bout de la nuit

{S}Elle était seule à présent la fille pour l'habiter le pavillon, que je me pensais… Ils étaient tous morts, tous… Elle avait dû savoir, ou du moins elle s'était doutée de la façon dont elle avait fini sa vieille à Toulouse… Quel effet que ça avait bien pu lui faire ?

- Pas de contexte explicite de prise de parole mais d'autres descripteurs du familier sont là :
  - Relative populaire : « quel effet que ça avait bien pu lui faire ? »
  - Nous pouvons aussi relever la ponctuation : « … » et « ? »
    - → indice d'une certaine oralité ?

me dégonfle… J'en ai trop vu… Et puis en plus j'ai tout le temps sommeil… Forcément, dormir le jour, c'est pas dormir… Sans compter la poussière des "bureaux" qu'on s'en remue plein les poumons… Tu te rends compte ?… Ça crève un homme… "

- Contexte de prise de parole :
  - Dialogue :
    - Formule interrogative
    - Présence de la seconde personne du singulier : « tu », « te »
- D'autres descripteurs du familier :
  - Non inversion lors de la formule interrogative
  - Forme contractée : « on »
  - Syntagme « Et puis » en début de phrase
  - Construction incomplète de la négation
  - Terme lexical familier voire vulgaire : « crève »

- Tu trouves pas que ça sent la pisse les malades ? qu'il m'a demandé alors.
- Oui, et la sueur aussi…
- Tout de même, fit-il lentement après avoir bien réfléchi, j'aurais bien aimé moi à être infirmier.
- Pourquoi ?

- Contexte de prise de parole explicite :
  - Marqueurs typographiques : les tirets
  - Verbe de parole : « a demandé », « fit-il »
  - Relative populaire : « qu'il m'a demandé »
  - Lexique familier : « pisse »
  - Construction de la négation incomplète : « Tu trouves pas que… »
  - Non inversion sujet/verbe dans la construction de l'interrogative

Kiffe Kiffe demain

{S}Ça va rien coûter à ta mère si c'est ça qui te préoccupe De toute façon, le ski ça pue la merde.
c'est comme si tu faisais du toboggan debout avec un bonnet et une combinaison boudinante et fluo.

- Descripteur croisé avec d'autres descripteurs pour le registre familier :
  - Plusieurs occurrences de formes contractées : « ça », « c'est »
  - Construction de la négation incomplète
  - Lexique vulgaire : « pue la merde »

{S}Et Maman, tout émue, interviewée par des chaînes de télé : " Ça fait lantemps je rêve ma fille monter dans les escaliers de Cannes, alors c'est fourmidable, merci boucoup " Pas les escaliers, Maman, les marches Alors pendant que je les monterais, j espérerais secrètement que la cérémonie soit retransmise sur la télévision marocaine et que mon barbu de père tombe dessus par hasard.

- Contexte de prise de parole explicite avec :
  - Marqueurs typographiques : les guillemets
  - Contexte d'interview
  - Première personne du singulier : « je » « ma »
- Descripteur croisé avec d'autres descripteurs :
  - Ecriture phonétique[26] : « lantemps », « fourmidable », « boucoup »

{S}Ça marche bien.
{S}Quand il y a des mariages dans le quartier, tout le monde l `appelle.
{S}Les filles se font faire des brushings, se font architirer les cheveux pour qu `on croie qu' elles les ont raides naturellement.

- Pas de contexte de parole explicite.
- Descripteurs familiers présent en contexte :
  - Lexique familier avec un néologisme : « architirer »
  - Relative populaire : « pour qu'on croie que »
  - Forme contractée : « on »

{S}Je me suis trompée c'est tout Et puis merde !{S} On est pas obligé d aimer les bébés !{S} Un bébé, ça pleure sans arrêt, ça pue, ça bave, ça fait caca Surtout que le bébé de la photo, il était trop vilain, il était en forme de croissant au beurre.
{S}Et puis, L' autre mouflet, c'est pas mon frère.
c'est juste le fils de mon barbu de père.
c'est pas pareil.

- Pas de contexte de parole explicite
- Plusieurs descripteurs discriminant pour le registre familier :
  - Forme contractée : « on », « ça », « c'est »
  - Le syntagme « Et puis » en début de phrase
  - Lexique familier : « mouflet », « merde », « caca »
  - Construction incomplète de la négative

L'Assommoir

{S}Chez elle, ça entrait et ça sortait.
{S}Elle restait très bien des huit jours sans songer à sa gourgandine ;{S} et, brusquement, une tendresse ou une colère l'empoignait, des fois à jeun, des fois le sac plein, un besoin furieux de pincer Nana dans un petit endroit, où elle l'aurait peut-être embrassée, peut-être rouée de coups, selon son envie du moment.
{S}Elle finissait par n'avoir plus une idée bien nette de l'honnêteté.
{S}Seulement, Nana était à elle, n'est-ce pas ?{S} Eh bien ! lorsqu'on a une propriété, on ne veut pas la voir s'évaporer.

- Pas de contexte de parole explicite
- Plusieurs descripteurs associés au registre familier se rajoutent :
  - Forme contractée : « ça », « on »,
  - Onomatopée : « Eh »

{S}Vrai, ça faisait un fameux débarras.
{S}Mais elle n'avait bien sûr pas laissé que maman Coupeau au fond du trou, dans le petit jardin de la rue Marcadet.
{S}Il lui manquait trop de choses, ça devait être un morceau de sa vie à elle, et sa boutique, et son orgueil de patronne, et d'autres sentiments encore, qu'elle avait enterrés ce jour-là.
{S}Oui, les murs étaient nus, son coeur aussi, c'était un déménagement complet, une dégringolade dans le fossé.

[26] Sommant 2005

- Pas de contexte de prise de parole

{S}Elle resta là un instant, l'échine tendue, l'oeil appliqué contre la vitre, entre deux bouteilles de l'étalage, à guigner Coupeau, dans le fond de la salle ;{S} il était assis avec des camarades, autour d'une petite table de zinc, tous vagues et bleuis par la fumée des pipes ;{S} et, comme on ne les entendait pas gueuler, ça faisait un drôle d'effet de les voir se démancher, le menton en avant, les yeux sortis de la figure.

- Pas de contexte de prise de parole explicite
- D'autres descripteurs en contexte pour le registre familier :
    - Forme contractée : « on »

→ Descripteur confirmé. Sa présence dans des œuvres dont le registre n'est pas familier est justifié par la mise en œuvre d'un contexte de parole. Autrement dit, ce descripteur est non seulement un élément discriminant pour le registre familier mais également utilisé comme tel afin de construire des discours rapportés. La différence avec sa présence dans les œuvres associées au registre familier est justement l'absence d'un contexte de prise de parole : le descripteur se fond dans la narration.

2.14 – C'est… qui… que

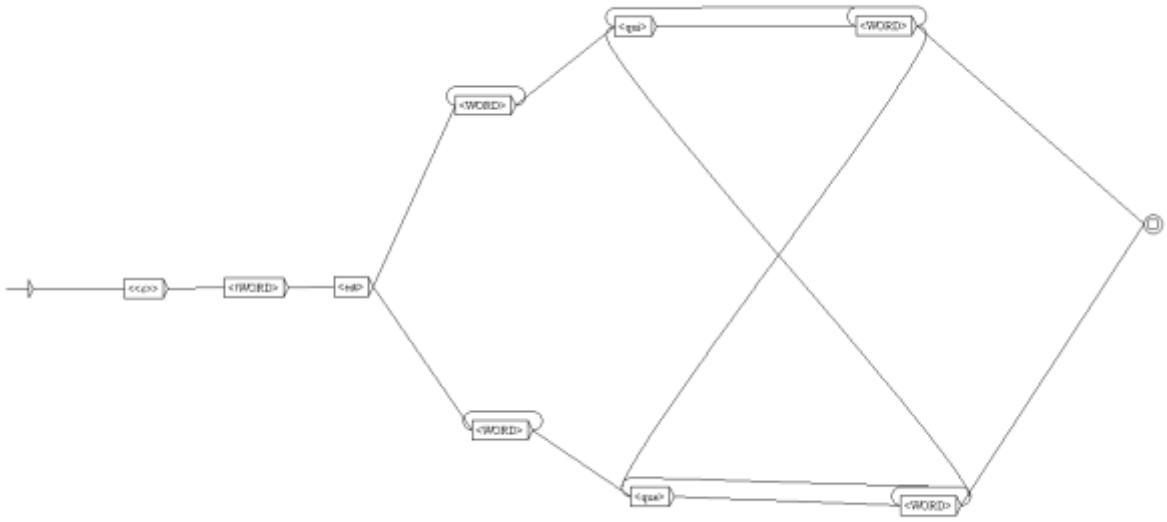


Figure 22 automate utilisé pour identifier le syntagme discriminant pour le registre familier

Kiffe Kiffe demain

c'est comme au Secours populaire avec Maman quand la vieille à qui on demande un sac en plastique pour mettre les pul s qu'`on a choisis nous regarde avec ses yeux mouil és.
{S}À chaque fois, on a envie de lui rendre ses pul s et de se tirer.
{S}M.
{S}Werbert, c'est pareil.
{S}Il se la joue prophète social.
{S}Il me dit que si J'`ai besoin, je peux prendre rendez-vous avec lui… Tout ça pour se donner bonne conscience et raconter à ses potes dans un bar parisien branché comme c'est difficile d'`enseigner en banlieue.
{S}Beurk.

- Pas de contexte de parole

- D'autres descripteurs se croisent :
    - Présence de la forme contractée « on » deux fois dans la même phrase
    - Lexique familier[27] : « se tirer » « se la joue »
    - Onomatopée : « Beurk »

{S}Si maintenant les hommes commencent à se lancer dans l'import-export… c'est dommage que chez nous les mariages ne se passent pas comme aux États-Unis avec le prêtre qui dit la fameuse phrase : « Si quelqu'un s'oppose à cette union, qu'il parle maintenant ou qu'il se taise à jamais.
» Là, il y a toujours un mec hyper courageux qui ose interrompre la cérémonie parce que ça fait huit ans qu'il est secrètement amoureux de la mariée.
{S}Alors, il lui annonce, et elle, la larme à l'œil, lui dit que c'est réciproque.
{S}Le marié qui est bon perdant - même s''il a un peu les boules - serre la main du mec hyper courageux et fait : « Sans rancune, vieux ! » Puis il lui prête le smoking qu'il a loué très cher pour l'occasion et le type courageux épouse la fille à la place du marié bon perdant.
{S}Pour le mariage d'Aziz, Maman pourrait faire la même chose.
elle dirait à Aziz que c'est le mec le plus romantique du quartier et qu''elle a des sentiments sincères envers lui depuis des années malgré son crâne chauve et ses ongles un peu crados.

- Pas de contexte de prise de parole explicite là où nous trouvons le descripteur.
- D'autres descripteurs familiers en contexte :
    - Lexique familier : « mec » « hyper »
    - Forme contractée : « c'est », « ça »

c'est pas un prénom pour faire rêver les petites filles, ça !{S} Françoise, c'est la poupée des petites filles qui rêvent pas.
{S}Quand j'étais plus jeune, je rêvais d'épouser le type qui ferait passer tous les autres pour des gros nazes.
{S}Les mecs normaux, ceux qui mettent deux mois à monter une étagère en kit ou à faire un puzzle vingt-cinq pièces marqué « dès cinq ans » sur la boîte, j'`en voulais pas.
{S}Je me voyais plutôt avec MacGyver.

- Pas de contexte de prise de parole explicite.
- D'autres éléments discriminants pour le registre familier sont présents :
    - Formes contractées : « c'est », « ça »
    - Construction de la négation incomplète
    - Lexique familier : « mecs », « type », « gros nazes »

{S}Le problème, c'est que celui qui doit me conduire à l'autel, c'est censé être mon connard de paternel.
{S}Mais comme il sera pas là, on sera obligés de tout annuler.
{S}Les invités reprendront tous leurs cadeaux de mariage et piqueront dans le buffet pour rapporter chez eux.
{S}Rien à foutre, de toute façon, avant de penser au mariage, faut d'abord trouver le mari.
{S}La chance de notre génération, c'est qu'`on peut choisir qui on va aimer toute sa vie.
{S}Ou toute l'année.
{S}Ça dépend des couples.

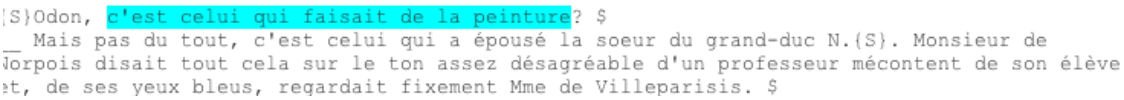
- Pas de contexte de prise de parole explicite.
- Deuxième syntagme plus loin : « c'est qu'on peut choisir qui on va aimer toute sa vie. »
- D'autres éléments discriminants pour le registre familier :
    - Forme contractée : « on »
    - Un lexique familier voire vulgaire : « connard », « rien à foutre »[28]
    - Relative populaire : « le problème, c'est que… »
    - Construction incomplète de phrase négative
    - Absence du sujet impersonnel : « faut d'abord »

Albertine disparue

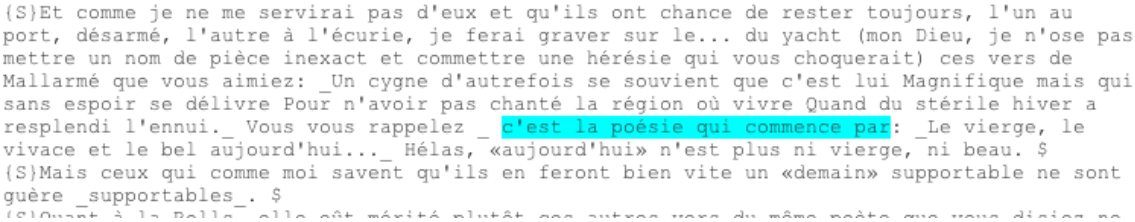
[27] Verbes - Maarit Ilmola 2012

[28] Verbes - Maarit Ilmola 2012

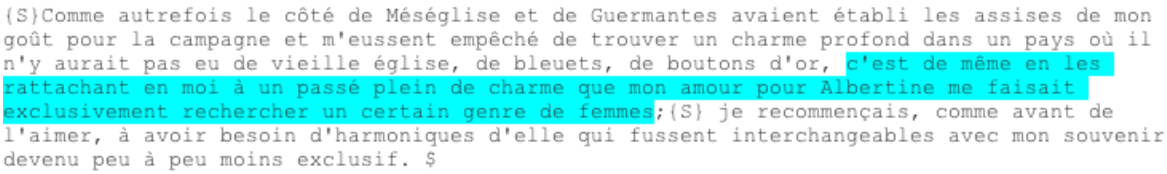
→ Après l'examen du concordancier (que vous pouvez retrouver dans les annexes) il apparaît assez nettement que plusieurs motifs syntaxiques figés, tels que "C'est ainsi que", "c'est ce que", "c'est celui qui", "c'est elle qui", "C'est parce que", reviennent de manière régulière : cela n'était pas le cas pour *Kiffe Kiffe demain.*



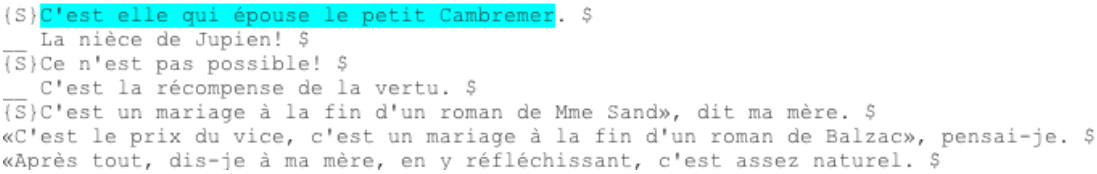
- Forme interrogative + contexte de dialogue : contexte oral



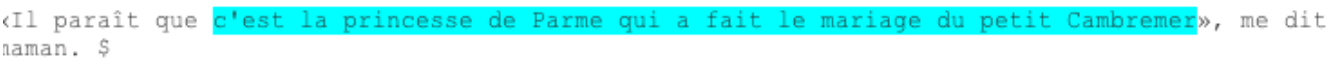
- Contexte de dialogue : contexte oral
    - Marqueurs typographiques : les tirets



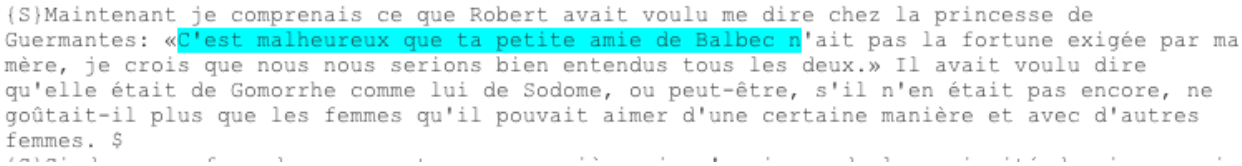
- Contexte de monologue intérieur : fréquence de la première personne du singulier par exemple (pronoms personnels, possessifs, formes verbales...)
    - Un monologue : lien avec l'oral



- Contexte de dialogue : oral
    - Marqueurs explicites tels que les tirets
    - Ponctuation « vive » avec des points d'exclamation



- Contexte de prise de parole : oral
    - Marqueurs explicites : les guillemets
    - Verbe de parole : « dit »



- Contexte de prise de parole : oral
    - Marqueurs explicites avec les guillemets
    - Verbe de parole : « avait voulu me dire »

→ forte dimension d'une monstration de l'objet dont le texte parle.
→ le contexte met en exergue un contexte de prise de parole orale externe (ex : dialogue) ou bien interne (ex : monologue)
→ le discours rapporté est très marqué avec notamment :

- Des marques de ponctuation mimétique de l'oral : "!", "...", "?"
- Des marques typographiques : les guillemets, les tirets bas
- Des verbes de parole : "avait-elle dit", "avait voulu me dire", "me dit", ""

→ Pour *Kiffe kiffe demain* le descripteur n'apparaît pas dans des situations explicitement marquées comme des prises de parole / situations orales.

L'Assommoir

- Le concordancier fait tout de suite apparaître la même régularité des motifs que celle observée pour *Albertine Dipsarue* : "c'est ça qui", "c'est ce que", "c'est lui qui", "c'est moi qui".
- Toutes les captures d'écrans et le concordancier sont dans les annexes

— Maman ! maman ! s'écria brusquement Nana, c'est Augustine qui laisse tomber son pain dans la rôtissoire !
{S}La blanchisseuse accourut et surprit le louchon en train de se brûler le gosier, pour avaler plus vite une tartine toute trempée de graisse d'oie bouillante.{S} Elle la calotta, parce que cette satanée gamine criait que ce n'était pas vrai.
{S}Après le boeuf, quand la blanquette apparut, servie dans un saladier, le ménage n'ayant pas de plat assez grand, un rire courut parmi les convives.

— Ah ! c'est ça qui te chiffonne ! dit le zingueur, sans cesser de ricaner.{S} Tu veux ta part.{S} Alors, grande cruche, pourquoi refuses-tu une consommation ?… Prends donc, c'est

— Dis donc, le petit veau, regarde ça ;{S} c'est un cadeau pour toi.{S} Oui, c'est encore cinquante sous que tu me coûtes… Avec ce joujou-là, je ne serai plus obligé de courir, et tu auras beau te fourrer dans les coins.{S} Veux-tu essayer ?… Ah ! tu casses les tasses !… Allons, houp ! danse donc, fais donc des révérences à monsieur Hardi !

— C'est papa qui vous envoie ? demanda Gervaise.
{S}Mais comme elle se baissait pour rattacher les cordons des souliers d'Étienne, elle vit, à un doigt de Claude, la clef de la chambre avec son numéro de cuivre, qu'il balançait.
— Tiens ! tu m'apportes la clef ! dit-elle, très-surprise.{S} Pourquoi donc ?
{S}L'enfant, en apercevant la clef qu'il avait oubliée à son doigt, parut se souvenir et cria de sa voix claire :
— Papa est parti.

- Pour l'*Assommoir* le descripteur est automatiquement utilisé dans un contexte de dialogue ou de prise de dialogue.
- Nous retrouvons les mêmes marques que pour *Albertine disparue*

→ Une forme encore très fortement liée à la parole orale/aux dialogues

L'Humanité

continuent de réclamer du concret.{S} Liliane Carriou, représentante du comité de Lorient, se dit après coup dubitative : « C'est vrai que les propositions n'ont pas été assez concrètes, j'espère que demain nous y verrons plus clair. »
travailler plus... pour gagner moins

- Contexte de prise de parole mis en exergue par les marqueurs typologiques, les guillemets.

{S}L'invasion de l'Irak que la France continue de considérer comme « une erreur » pourrait être de ceux-là.{S} Tout comme la lutte contre le réchauffement climatique.{S} Le président français devrait aussi mettre ces thèmes en exergue pour donner l'illusion d'un rapprochement qui ne s'apparenterait pas à un alignement total. « On est d'une même famille.{S} C'est ça qui importe.{S} Même si parfois, comme dans toute famille, il y a des désaccords », avait-il déclaré lors de ses vacances dans le New Hampshire.{S} Autrement dit : la France a l'ambition de jouer un grand rôle de... vassal dans le monde.
{S}Bruno Odent
société

- Les guillemets indiquent un discours rapporté direct

rendant, gratuitement, les prestations attendues, de retraites qui ne baissent pas, et d'échanges qui ne soient pas exclusivement marchands.{S} Les travailleurs et leurs syndicats ont plus que raison d'exiger tout cela.{S} L'augmentation générale des salaires, dont la part dans la valeur ajoutée, en France et dans le monde, a baissé ces dernières années ;{S} plus de 200 milliards d'euros dans le total des richesses produites dans notre pays.{S} C'est bien là que se trouve la véritable recette pour répondre à la valse des étiquettes !{S} Quoi, ce serait les salaires, qui dans un produit arrivant sur une console, seraient coupables de son prix trop élevé ?{S} Mais comment donc se forme le prix de vente de la chaussure, dont les matières premières pour la fabriquer, et la chaussure elle-même, auront parcouru des milliers de kilomètres avant d'arriver au pied du « consommateur » ?{S} Peuvent-ils nous faire croire, tous ces docteurs Folamour, que leurs profits (qui, dans les pays les plus riches, ont accaparé plus de 1 350 milliards d'euros supplémentaires ces dernières années) ne seraient pour rien dans le prix sur nos étals ?{S} Nous feront-ils avaler que ces « boucliers fiscaux » et autres baisses d'impôts pour les riches ne conduisent pas à l'augmentation, par exemple, de « l'étiquette » de la TVA qui pèse si lourdement sur le budget des plus défavorisés ?{S} Nous persuaderont-ils que c'est pour faire baisser les prix qu'il faut une franchise médicale

- Pas de prise de parole explicitée.
    - Formule qui permet la monstration

{S}Depuis, Tayeb, le père de Mounir, ne décolère pas. « Il leur fallait des coupables à tout prix, peste-t-il, c'est n'importe quoi.{S} Je condamne les violences faites aux policiers, mais je veux aussi que mon fils soit jugé de manière équitable et que l'on trouve les vrais responsables.{S} Là, Mounir a été condamné sur la base de simples rumeurs, de ragots... » La famille du jeune homme a été atterrée par la décision du tribunal d'Évry.{S} Et sous la houlette d'Anne Coulon, un professeur d'EPS à l'IUFM d'Étiolles, un comité de soutien a été monté courant octobre.{S} Samedi, c'est elle qui a pris le micro en premier, sur ce petit terrain de foot du quartier où les manifestants se sont rassemblés avant de défiler jusqu'à la mairie. « Nous ne défendons pas Mounir et Medhi, précise-t-elle.{S} Leurs avocats sont payés pour cela.{S} Nous, nous défendons le droit et la justice.{S} Une condamnation si exemplaire sur la base d'un témoignage sous X nous semble hautement dangereuse. »

- Pas de prise de parole explicitée.
    - Formule qui permet la monstration

{S}Le proces de l'ancien ministre de l'Interieur Charles Pasqua s'est ouvert hier devant le tribunal correctionnel de Paris.{S} Charles Pasqua, quatre-vingts ans, est poursuivi pour financement illégal présumé de sa campagne des européennes de 1999, où il menait la liste du Rassemblement pour la France en compagnie de Philippe de Villiers.{S} C'est la première fois que le sénateur UMP des Hauts-de-Seine, mis en examen dans plusieurs affaires financières, comparaît devant un tribunal.{S} La justice soupçonne un éventuel lien entre le financement de sa campagne et les profits tirés de la vente du casino d'Annemasse (Haute-Savoie) en 1995.
{S}Institutions Moscovici invite Balladur

- Pas de contexte de prise de parole, la formulation semble vouloir insister sur la situation :
    - Valeurs d'insistance

→ Ce n'est pas aussi marqué que pour *L'Assommoir* ou bien *Albertine Disparue* mais la majorité des descripteurs apparaissent dans un contexte de prise de parole.

- Il est intéressant de noter que le motif peut prendre 2 valeurs différentes (qui reviennent) :
    - La monstration
    - L'insistance

2.15 – Des ADJ

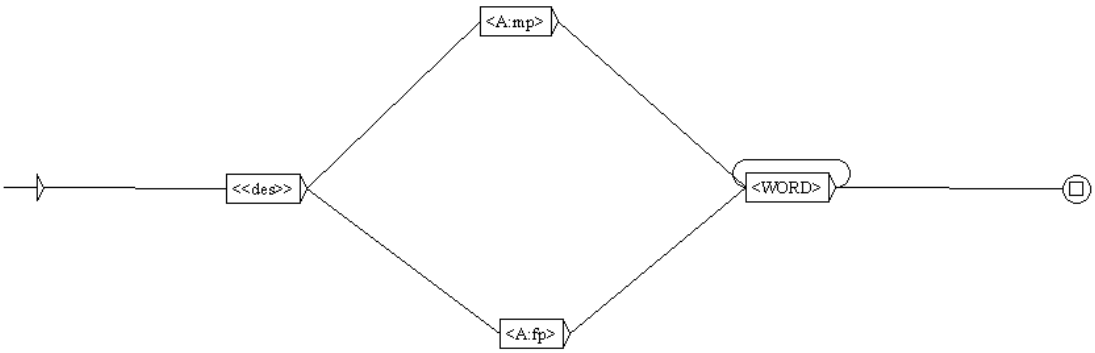


Figure 23 automate qui permet d'identifier le syntagme recherché : "des" + ADJ

Remarque : l'automate génère du bruit puisqu'il repère également les formes commençant par "des". Toutefois, il ne donnait rien lorsque je spécifiais "des" uniquement avec des chevrons simples.

→ Les concordanciers sont disponibles dans les annexes : descripteur non confirmé car présent dans tous les registres.

2.16 – « Et puis » en début de phrase

→ le script se trouve dans les annexes. Ils s'y trouvent étalement des fichiers qui stockent les phrases extraites des différents textes/registres.

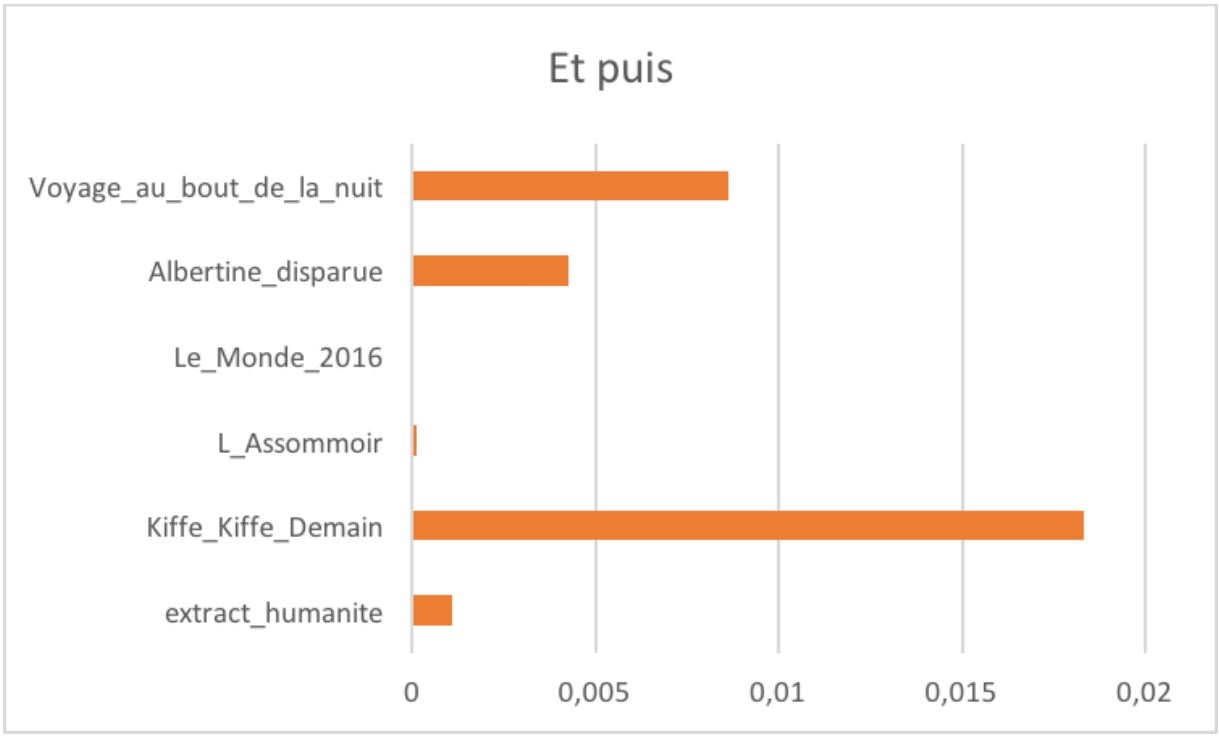


Figure 24 graphe des fréquences relatives du syntagme "Et puis" dans chaque texte/registre

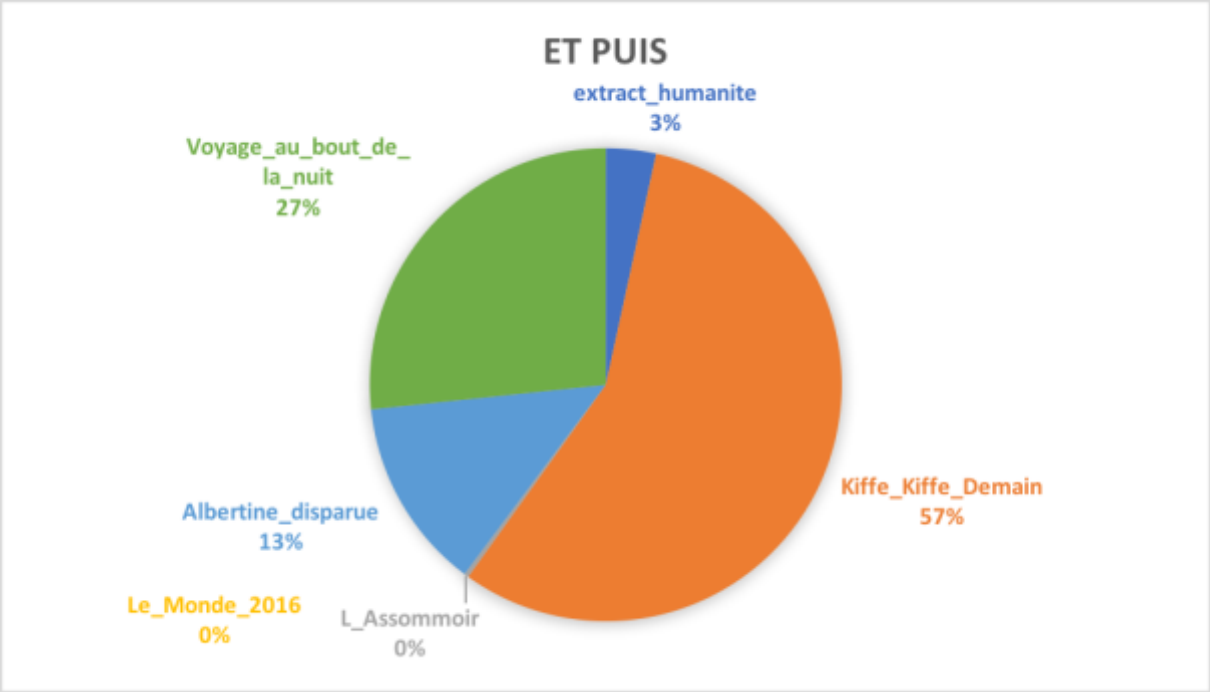


Figure 25 comparaison entre les registre de la présence du motif

→ Le descripteur est confirmé car nous avons une nette différence de présence du motif entre les œuvres du registre familier et les autres.

2.17 – Verbe du premier groupe majoritaire pour le registre familier

→ le script est accessible dans les annexes

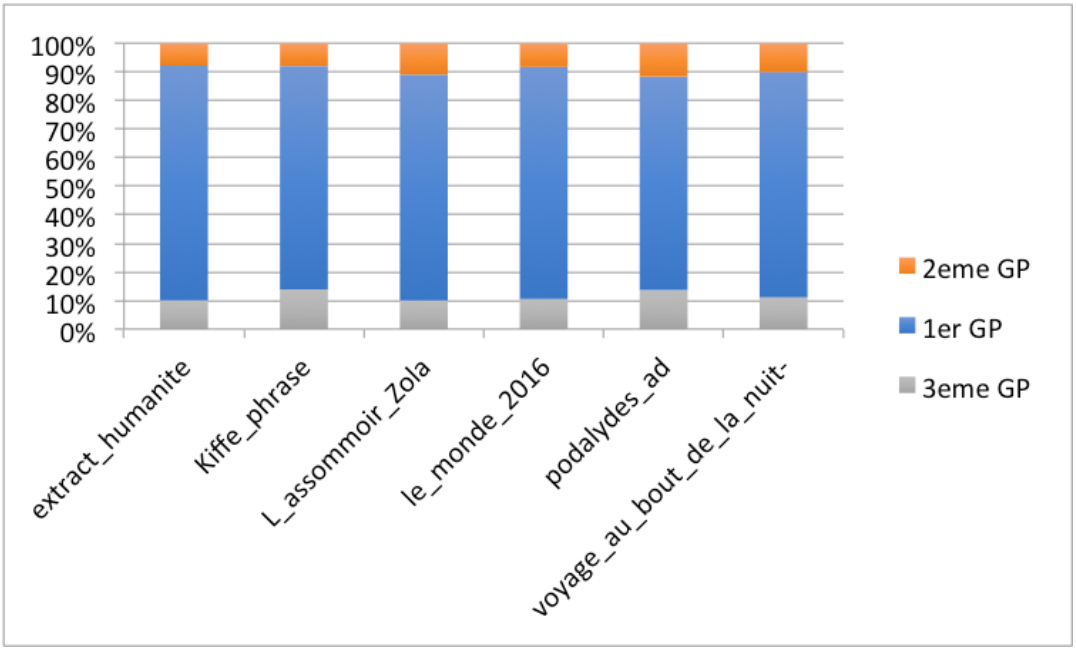


Figure 26 Proportion des différents groupes verbaux dans les textes/registres

→ Le descripteur ne peut pas être confirmé puisque nous avons une large majorité de verbe du premier groupe pour tous les registres. Toutefois, nous pouvons remarquer que le texte où il y a légèrement plus de verbes du deuxième et troisième groupe (et donc moins du premier) est le texte d'*Albertine disparue* (podalydes_Ad).

2.18 - Absence de l'inversion SUJ/VB lors d'une interrogative

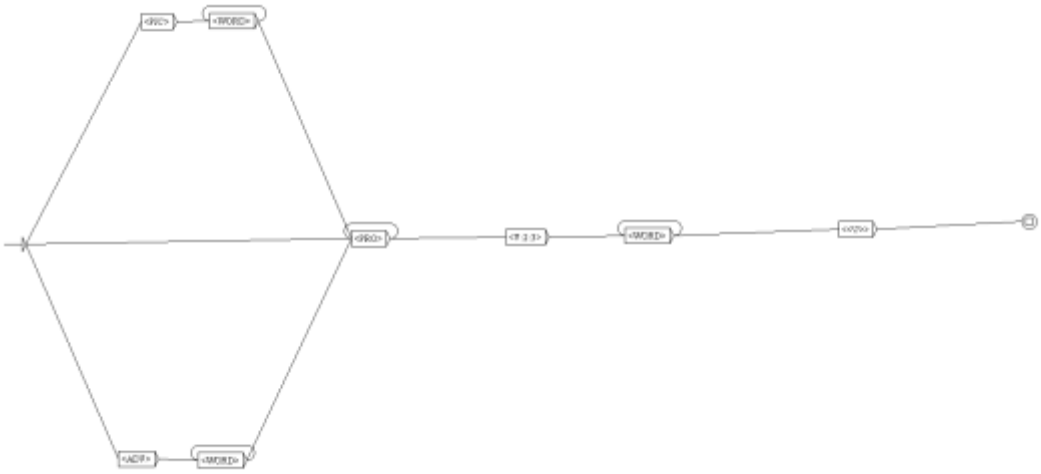


Figure 27 automate pour identifier le motif discriminant

→ Tous les concordanciers sont dans les annexes

L'Humanité

{S}C'était le 1er mai dernier, jour de la Fête du travail.{S} Le candidat d'alors à la présidence de la République allait se rendre en Bretagne.{S} La symbolique de la journée pour le candidat du « travailler plus pour gagner plus » allait être effacée au profit de premières annonces sur la maladie d'Alzheimer, qui se traduiront dans les faits par les franchises médicales pour tous à l'automne. « Qu'est-ce qu'on va foutre dans un centre opérationnel sinistre à regarder un radar ?{S} Qui a eu cette idée de demeuré ?{S}... Je me fous des Bretons.{S} Je vais être au milieu de dix connards en train de regarder une carte ! (...) Derniers jours de campagne dans une salle à voir une carte !{S} Grand sens politique, vraiment ! »
{S}Ces propos, disons « fleuris », émanant d'un prétendant à la magistrature suprême, pouvaient étonner.{S} Ils font l'effet d'une détonation en plein mois d'août, quand le

- Monstration de la prise de parole avec des marqueurs explicites tels que les guillemets - discours direct
- Seconde monstration avec "ces paroles" qui caractérise le discours rapporté
- Longueur courte de la phrase

{S}Jean Hatzfeld.{S} L'un d'eux a tant parlé, dénonçant tout le monde autour de lui, qu'il s'est fait tuer.{S} Ils commencent à me dire des choses sans se rendre compte exactement de ce qui se passe.{S} Ils ne savent pas très bien ce qu'est un livre.{S} Là-bas, il n'y a ni librairie ni bibliothèque.{S} La plupart sont des cultivateurs.{S} Ils ont certes peut-être lu des livres à l'école ainsi que l'Ancien et le Nouveau Testament mais ce ne sont pas des lecteurs aguerris.{S} Ils voient ce Blanc qui leur pose des questions.{S} Ils se disent pourquoi pas ?{S} Ça les sort un peu de l'ordinaire de la prison.{S} Et puis ils ont un souci, ils sont quand même embêtés, pas beaucoup plus que ça, mais tout de même.{S} Ils ont tué tous les jours de la semaine, « à s'en casser les bras », selon leur expression.{S} Ils ont massacré leurs voisins, des enfants, des femmes, sans haine, sans peur, comme un boulot ;{S} comme on taille des bananes.{S} Ils sont quand même très embêtés avec ça.{S} Ils ont du mal avec ces deux personnages en eux.{S} Ils sont à la fois celui qui ne peut pas nier qu'il allait tous les jours dans les marais avec sa machette.{S} Et cet autre qu'ils croient, naïvement sans doute, être redevenu.{S} Ils ont le souci de me donner accroire, grosso modo,

- Entretien > discours rapporté
- Descripteur "ça" présent
- Longueur courte de la phrase

« Nous ne pensons pas que ce genre de mesures d'urgence soit compatible avec le type d'environnement nécessaire à la tenue d'élections libres et honnêtes. » l Et il a trouvé ça tout seul ?
{S}DESMOND TUTU, ARCHEVÊQUE D'AFRIQUE DU SUD :
« Notre monde affronte des problèmes, pauvreté, séropositivité, sida _ une pandémie dévastatrice _, conflits et, face à tout cela, notre Église, en particulier l'Église anglicane, en ce moment, est presque obsédée par les problèmes de la sexualité humaine. » l Là, il fallait le dire.

- « Et » qui commence la phrase
- Présence de ça
- Longueur courte de la phrase

« Alors que les transferts de détenus duraient environ cinq minutes à Moulins, il va falloir maintenant presque une heure pour rejoindre, à travers la campagne bourbonnaise, les TGI de Montluçon ou de Cusset-Vichy, explique, furieux, un surveillant de la prison.{S} En termes de sécurité et de coût, c'est de la folie ! » Ironie du sort, le ministère de la Justice vient juste de débloquer 700 000 euros pour rénover et sécuriser le tribunal de Moulins, installé dans un ancien collège de jésuites du XVIIIe siècle. « La veille de la décision de suppression, deux portiques de sécurité ont été installés au rez-de-chaussée.{S}.. Ça rime à quoi ? »

- Présence de « ça »
- Discours rapporté
- Longueur courte de la phrase

{S}Jean-Pierre Améris.{S} Au moment de la fermeture du centre de Sangatte, en 2003, des bénévoles ont été mis en examen, pour avoir hébergé des migrants chez eux.{S} J'ai ressenti de la surprise, et une sincère indignation.{S} Ça signifie quoi ?{S} Que donner l'hospitalité est criminel ?{S} Olivier Adam avait déjà travaillé, à Calais, sur ce thème, avec des lycéens.{S} Il racontait certaines réactions violentes des lycéens, hostiles aux bénévoles.{S} Et nous avions le projet de raconter l'histoire d'une mère de famille.{S} À l'automne 2003, Olivier a écrit une nouvelle, et moi je suis parti à Calais, auprès des bénévoles.
{S}Et vous êtes devenu bénévole vous-même ?
{S}Jean-PierreAméris.{S} J'avais eu cette démarche avec mon film C'est la vie, sur les soins palliatifs.{S} Au départ, les bénévoles étaient réticents devant mon projet.{S} Ce qui a

- Contexte d'entretien donc discours rapporté
- Présence de « ça »
- Longueur de la phrase : courte

→ nous remarquons que la longueur des interrogatives sont courtes et les descripteurs se croisent pour se consolider.

Albertine disparue

{S}Odon, c'est celui qui faisait de la peinture? $
__ Mais pas du tout, c'est celui qui a épousé la soeur du grand-duc N.{S}. Monsieur de Norpois disait tout cela sur le ton assez désagréable d'un professeur mécontent de son élève et, de ses yeux bleus, regardait fixement Mme de Villeparisis. $
{S}Quand le prince eut fini son café et quitta sa table, Monsieur de Norpois se leva, marcha avec empressement vers lui et, d'un geste majestueux, il s'écarta, et, s'effaçant lui-même, le présenta à Mme de Villeparisis. $
{S}Et pendant les quelques minutes que le prince demeura assis avec eux, Monsieur de Norpois ne cessa un instant de surveiller Mme de Villeparisis de sa pupille bleue, par complaisance ou sévérité de vieil amant, plutôt dans la crainte qu'elle ne se livrât à un des écarts de langage qu'il avait goûtés, mais qu'il redoutait. $
{S}Dès qu'elle disait au prince quelque chose d'inexact il rectifiait le propos et fixait des yeux la marquise accablée et docile, avec l'intensité continue d'un magnétiseur. $
{S}Un garçon vint me dire que ma mère m'attendait, je la rejoignis et m'excusai auprès de Mme Sazerat en disant que cela m'avait amusé de voir Mme de Villeparisis. $
{S}À ce nom, Mme Sazerat pâlit et sembla près de s'évanouir. $
{S}Cherchant à se dominer: __ Mme de Villeparisis, Mlle de Bouillon? $
me dit-elle. $
elle est ici? $
__ Oui, il est même temps que je vous quitte, car je dois aller la voir tout à l'heure.» Et comme, depuis qu'Albertine était morte, je ne la tenais plus prisonnière chez moi comme dans les derniers temps de sa vie, sa visite à Mlle Vinteuil m'inquiétait. $
{S}Je ne voulais pas le laisser voir. $

__ Tu es des Laumes? $
{S}Hé bien, je suis ton prince.» Alors le paysan regarda la figure toute glabre de Basin et lui répondit: «Pas vrai. $
{S}Vous, vous êtes un _english_.» On voyait ainsi dans ces petits récits de la duchesse ces grands titres éminents, comme celui de prince des Laumes, surgir à leur place vraie, dans leur état ancien et leur couleur locale, comme dans certains livres d'heures on reconnaît, au milieu de la foule de l'époque, la flèche de Bourges. $
{S}On apporta des cartes qu'un valet de pied venait de déposer. $
xJe ne sais pas ce qui lui prend, je ne la connais pas. $
{S}C'est à vous que je dois ça, Basin. $

- Contexte de dialogue : marqueurs typographiques tels que les tirets
- Discours direct

Kiffe Kiffe Demain

{S}Il croit que de l'argent ça compense un manque ou quoi ?{S} Faut qu'il arrête de lire les dossiers psycho des magazines féminins posés sur la table basse chez Lila.
{S}Même ce qu'il a écrit c''était nul : « Si t'as besoin de moi, tu sais où me trouver… »
Ouais eh ben ce que je sais, Hamoudi, c'est que t'es plus dans le hal 32.
{S}Tu nous as laissés tomber Rimbaud et moi.
{S}Lâcheur.
{S}Tous pareils de toute façon.
{S}Des lâcheurs.
{S}Même Mme Burlaud, si elle était pas payée pour me voir à heure fixe une fois par semaine, je suis sûre qu''elle m'aurait lâchée.
{S}En passant devant le bar-tabac du centre-vil e, J'`ai remarqué un papier col é sur la vitrine.

- Présence d'un second descripteur : ça
- Ponctuation à la fin "ou quoi ?" → motif familier à explorer

- Ça te dit pas de gagner un peu de sous ?{S} J'`ai trouvé ça gentil qu''elle pense à moi, c'est vrai ça, elle aurait pu proposer ça à n'importe quelle fille du quartier mais non, elle a pensé à moi.
{S}Je retire tout ce que J'`ai dit sur elle, sa permanente et tout le reste… - Comme ça tu pourras t'habiller comme les autres jeunes de ton âge, hein ?{S} Sur le moment, J'`ai pas vraiment su comment le prendre.

- Présence d'un second descripteur : ça (forte présence dans le contexte)
- Dans un contexte plus éloigné (dernière phrase) : construction incomplète de la négation

j'`en ai marre, j'suis à bout… Et puis j'`en profite puisque j'`en ai l'occasion aujourd'hui, je vends ma caisse si ça intéresse quelqu'un, c'est une Twingo verte en bon état, elle est pratiquement neuve, elle a sept ans… » En plus d'apprendre ça par Sarah… Non mais après, ça va être quoi ?{S} Pourquoi Hamoudi ne m'a rien dit ?{S} Il me prend encore pour une gamine ?{S} Il croit peut-être que je comprends pas ce genre de trucs ?{S} J'`ai été capable de comprendre des trucs plus compliqués.

- Présence d'un second descripteur : ça (forte présence dans le contexte)
- Pas de construction complète de la négation dans le contexte.
- Lexique familier : « gamine », « trucs », « caisse »

elle crie alors de sa voix grave - elle avait pas compris que ça servait à rien de crier puisque c'est fait pour ça un micro à la base : - RAYMOND !!! c'est combien le paquet de serviettes hygiéniques de 24 + 2 gratuites, modèle normal avec coussinets absorbants et ailettes latérales protectrices ?! elle hésite un peu et reprend : - Ho !{S} Raymond !{S}! Tu dors ou quoi ?{S} Là, on a entendu une voix diabolique s'`exclamer que ce putain de paquet valait deux euros trente-huit.
{S}Le pire, c'est que j'`avais même pas assez pour le payer et qu''elle m'a fait crédit.
{S}Si j''aurais su, j''aurais même pas eu mes règles… Quand je suis rentrée à la maison, Maman était encore au téléphone avec Tante Zohra.
{S}Youssef va bientôt passer en jugement, c'est pour ça qu''elle est aussi inquiète.
{S}Alors Tante Zohra appelle souvent Maman, même tard le soir, parce qu''elle a du mal à s'endormir.
{S}Toutes les deux, elles ont de longues conversations pleines de silences.

- Présence d'un second descripteur ? "ou quoi" qui ponctue la fin de phrase
- Longueur de la phrase (courte)
- Présence de forme contractée : « ça », « on »
- Descripteur à explorer pour le familier : « même pas »

```
Descripteurs qui se croisent :
    Longueur des phrases
    "ça" / "ou quoi" / « on »
    Absence de la double négation
    -> lexique
    -> syntaxique
```

Voyage au bout de la nuit

{S}En bas, la tante de Bébert m'attendait avec sa tête de circonstance. « Ça ne va pas alors ? » qu'elle s'enquérait.{S} Je compris qu'elle m'avait attendu là, en bas, pendant une demi-heure déjà pour toucher sa commission d'usage : deux francs.{S} Que je l'échappe pas. « Et chez les Henrouille alors, ça a marché ? » voulut-elle savoir.{S} Elle espérait toucher un pourboire pour ceux-là aussi. « Ils ne m'ont pas payé », ai-je répondu.{S} C'était vrai aussi. {S} Son sourire préparé, tourna en moue à la tante.{S} Elle me suspectait.
« C'est pas malheureux tout de même Docteur, de pas savoir se faire payer !{S} Comment voulez-vous que les gens vous respectent ?… On paye comptant au jour d'aujourd'hui ou jamais ! » C'était exact aussi.{S} Je filai.{S} J'avais mis mes haricots à cuire avant de partir.{S} C'était le moment, la nuit tombée, d'aller acheter mon lait.{S} Pendant la journée, les gens avaient le sourire quand ils me croisaient avec ma bouteille.{S} Forcément.{S} Pas de bonne. {S}Et puis l'hiver a traîné, s'est étalé pendant des mois et des semaines encore.{S} On n'en sortait plus de la brume et de la pluie, au fond de tout.
{S}Les malades ne manquaient pas, mais il n'y en avait pas beaucoup qui pouvaient ou qui

- Présence de "ça"
- Inséré clairement dans du discours rapporté avec les marqueurs typographiques : les guillemets

{S}Elle d'abord : « Le Docteur, ton ami, il est gentil n'est-ce pas ? » Elle revenait à la charge, comme si je lui étais resté sur l'estomac. « Il est gentil !… Je ne veux rien dire contre lui, puisque c'est un ami à toi… Mais c'est un homme qu'on dirait brutal tout de même avec les femmes… Je veux pas en dire du mal puisque je crois c'est vrai qu'il t'aime bien… Mais enfin ça serait pas mon genre… J' vais te dire… Ça va pas te vexer au moins ? » Non, rien ne le vexait Léon. « Eh bien, il me semble, le Docteur, qu'il les aime comme trop les femmes… Comme les chiens un peu, tu me comprends ?… Tu trouves pas toi ?… C'est comme s'il sautait dessus qu'on dirait toujours !{S} Il fait du mal et il s'en va… Tu trouves pas toi ? qu'il est comme ça ? »
{S}Il trouvait le saligaud, il trouvait tout ce qu'elle voulait, il trouvait même que ce qu'

- Descripteurs qui viennent redoubler la non inversion
  - "ça"
  - Élision du e[29] : "J'"
- Phrase courte
- Contexte de dialogue avec marqueurs explicites du discours rapporté

« Tu crois toi, qu'elle boufferait le morceau ? que je lui demandai encore pour m'assurer… Mais elle était quand même un peu ta complice ?… Ça devrait la faire réfléchir un moment avant de se mettre à baver ?
- Réfléchir ?… qu'il ressaute lui alors en m'entendant.{S} On voit bien que tu la connais pas… » Ça le faisait rigoler de m'entendre. « Mais elle n'hésiterait pas une seconde !… Comme je te le dis !{S} Si tu l'avais fréquentée comme moi, tu n'en douterais pas !{S} C'est une amoureuse que je te répète !… T'en as donc jamais fréquenté toi des amoureuses ?{S} Quand elle est amoureuse, elle est folle, c'est bien simple !{S} Folle !{S} Et c'est de moi qu'elle est amoureuse et qu'elle est folle !… Tu te rends compte ?{S} Tu comprends ?{S} Alors tout ce qui est fou ça l'excite !{S} C'est bien simple !{S} Ça l'arrête pas !{S} Au contraire !… »
{S}Je ne pouvais pas lui dire que ça m'étonnait quand même un peu, qu'elle en soit arrivée en quelques mois à ce degré de frénésie Madelon, parce que tout de même, je l'avais connue un petit peu moi-même, Madelon… J'avais mon idée à son sujet, mais je ne pouvais pas la dire.
{S}D'après la façon dont elle se débrouillait à Toulouse et telle que je l'avais entendue quand j'étais derrière le peuplier le jour de la péniche, c'était difficile de me figurer qu'elle avait pu changer de dispositions à ce point en si peu de temps… Elle m'avait semblé plus débrouillarde que tragique, gentiment affranchie et bien contente de se caser avec des petites histoires et son petit chiqué partout où ça pouvait prendre.{S} Mais pour le moment, où nous en étions, je n'avais plus rien à dire.{S} J'avais qu'à laisser passer. « Bon !{S} Bien !{S} Ça va ! que je conclus.{S} Et sa mère alors ?{S} Elle a dû faire un peu de bruit aussi la mère, quand elle a compris que tu te débinais pour de bon ?…

- Motif dans une relative (rôle de la relative ici pour permettre la non inversion)
- Lexique familier qui s'ajoute au descripteur : « boufferait »[30]
- Formes contractées dans le contexte : « ça »
- Le début de la phrase suivante commence par un « que »
  - Cela indiquerait que l'interrogative où nous avons trouvé le descripteur est mimétique de la prosodie du locuteur

[29] Elision du "e" - Favart 1966-2006

[30] Verbes - Maarit Ilmola 2012

- Elle était tout de même pas mal la petite Madelon ?{S} Tu peux pas dire le contraire ?
- C'est pas là la question…
- Mais bien sûr que si que c'est la question.{S} Puisque vous étiez libres que tu me dis… Si vous teniez absolument à quitter Toulouse, vous pouviez bien laisser le caveau en gérance à sa mère pendant un temps… Vous seriez revenus plus tard…
- Pour ce qui est du physique, reprit-il, tu peux le dire, elle était vraiment gentille, je l'admets, tu m'avais bien tuyauté en somme, surtout imagine que comme un fait exprès quand j'ai revu pour la première fois, c'est pour ainsi dire elle que j'ai revue en premier, dans une glace… Tu imagines ?… À la lumière !… Y avait bien à peu près deux mois que la vieille était tombée… La vue m'est revenue comme d'un coup sur elle Madelon, en essayant de lui regarder la figure… Un coup de lumière en somme… Tu me comprends ?
- C'était pas agréable ?

- Phrase courte
- Absence de la double négation
- Forme contractée : « c'est »

L'Assommoir

— Eh bien ! il n'y a plus personne ? reprit Bazouge dans le silence.{S} Attendez, on est complaisant pour les dames.
— Rien, ce n'est rien, dit enfin la blanchisseuse d'une voix étranglée.{S} Je n'ai besoin de rien.{S} Merci.

- Verbe de parole qui explicite le discours direct
- Phrase courte
- Ponctuant tel que "eh bien" qui redouble le premier descripteur

— Tu as donc visité et demandé le prix ? dit Coupeau.
— Oh ! tu sais, par curiosité ! répondit-elle, en affectant un air d'indifférence.{S} On cherche, on entre à tous les écriteaux, ça n'engage à rien… Mais celle-là est trop chère, décidément.{S} Puis, ce serait peut-être une bêtise de m'établir.
{S}Cependant, après le dîner, elle revint à la boutique du mercier.{S} Elle dessina les lieux, sur la marge d'un journal.{S} Et, peu à peu, elle en causait, mesurait les coins, arrangeait les pièces, comme si elle avait dû, dès le lendemain, y caser ses meubles.{S} Alors, Coupeau la poussa à louer, en voyant sa grande envie ;{S} pour sûr, elle ne trouverait rien de propre, à moins de cinq cents francs ;{S} d'ailleurs, on obtiendrait peut-être une diminution.{S} La seule chose ennuyeuse, c'était d'aller habiter la maison des Lorilleux, qu'elle ne pouvait pas souffrir.{S} Mais elle se fâcha, elle ne détestait personne ;{S} dans le feu de son désir, elle défendit même les Lorilleux ;{S} ils n'étaient pas méchants au fond, on s'entendrait très bien.{S} Et, quand ils furent couchés, Coupeau dormait déjà qu'elle continuait ses aménagements intérieurs, sans avoir pourtant, d'une façon nette, consenti à louer.

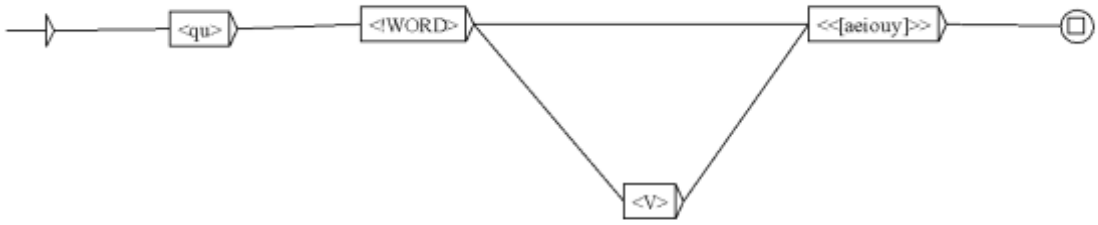
— Juste ! reprit le zingueur, blaguant pour la remettre, j'avais commandé une fille !{S} Hein ! me voilà servi !{S} Tu fais donc tout ce que je veux ?
{S}Et, prenant l'enfant, il continua :
— Qu'on vous voie un peu, mademoiselle Souillon !… Vous avez une petite frimousse bien noire.{S} Ça blanchira, n'ayez pas peur.{S} Il faudra être sage, ne pas faire la gourgandine, grandir raisonnable, comme papa et maman.
{S}Gervaise, très sérieuse, regardait sa fille, les yeux grands ouverts, lentement assombris d'une tristesse.{S} Elle hocha la tête ;{S} elle aurait voulu un garçon, parce que les garçons se débrouillent toujours et ne courent pas tant de risques, dans ce Paris.{S} La sage-femme dut enlever le poupon des mains de Coupeau.{S} Elle défendit aussi à Gervaise de parler ;{S} c'était déjà mauvais qu'on fît tant de bruit autour d'elle.{S} Alors, le zingueur dit qu'il fallait prévenir maman Coupeau et les Lorilleux ;{S} mais il crevait de faim, il voulait dîner auparavant.{S} Ce fut un gros ennui pour l'accouchée de le voir se servir lui-même, courir à la cuisine chercher le ragoût, manger dans une assiette creuse, ne pas trouver le pain.{S} Malgré la défense, elle se lamentait, se tournait entre les draps.{S} Aussi, c'était bien bête de n'avoir pas pu mettre la table ;{S} la colique l'avait assise par terre comme un coup de bâton.{S} Son pauvre homme lui en voudrait, d'être là à se dorloter, quand il mangeait si mal.{S} Les pommes de terre étaient-elles assez cuites, au moins ?{S} Elle ne se rappelait plus si elle les avait salées.

```
– Tu veux donc que je vole ? murmura-t-elle d'une voix sourde.
{S}Mes-Bottes se caressait le menton d'un air conciliant.
– Non, ça, c'est défendu, dit-il.{S} Mais quand une femme sait se retourner…
{S}Et Coupeau l'interrompit pour crier bravo !{S} Oui, une femme devait savoir se retourner.
{S} Mais la sienne avait toujours été une guimbarde, un tas.{S} Ce serait sa faute, s'ils
crevaient sur la paille.{S} Puis, il retomba dans son admiration devant Mes-Bottes.{S} Était-
il assez suiffard, l'animal !{S} Un vrai propriétaire ;{S} du linge blanc et des escarpins un
peu chouettes !{S} Fichtre ! ce n'était pas de la ripopée !{S} En voilà un au moins dont la
bourgeoise menait bien la barque !
{S}Les deux hommes descendaient vers le boulevard extérieur.{S} Gervaise les suivait.{S} Au
bout d'un silence, elle reprit, derrière Coupeau :
– J'ai faim, tu sais… J'ai compté sur toi.{S} Faut me trouver quelque chose à claquer.
```

- Contexte de discours direct avec des marqueurs typographiques explicites et des verbes de parole
- Phrase courte
- Dans chaque phrase interrogative nous avons la présence de « donc » : mais est-ce un descripteur qui s'ajoute ou bien un élément lié au style de l'auteur ?

```
→ la longueur des phrases interrogative est courte
→ Beaucoup de descripteurs secondaires viennent corréler le premier
descripteur dans le contexte
→ Forte présence de "que" et "qu'est-ce que"
```

2.19 – qu(i) voyelle



- Le second chemin qui passe par le verbe est plus précis et donc plus concluant.
  - En effet, l'élision a pour but de repérer le pronom relatif 'qui' : or lorsqu'il est suivi d'un autre pronom cela n'est plus un "écart" avec la norme puisque le pronom devient sujet.
  - En revanche lorsque le token qui suit est un VB alors il y a écart puisqu'il n'y a plus de pronom sujet : 'qu' ne remplit pas la fonction de sujet puisque tronqué.
- A noter : intéressant de voir la forme des VB "illait" qui forment une unité non standard et surtout qui est phonétiquement mimétique de la forme standarde "qui allait".
  - Formes répétées plusieurs fois dans le texte

```
{S}Samra, c'est la prisonnière qu'habitait dans mon immeuble et que le frère et le père ont
poussée à bout jusqu'à ce qu''elle se tire.
{S}On l''aurait aperçue, il y a quelques jours, pas trop loin, ou alors très loin, je sais
plus.
{S}En tout cas, on dit qu''elle s'est enfuie de chez elle pour un garçon.
{S}Je me disais bien qu''elle devait avoir une excellente raison de s'être tirée du
pénitencier.
```

- Descripteur croisé une sur utilisation de que[31] : « et que le frère et le père »
- Forme contractée : « c'est »

{S}Par exemple ce môme qu'était avec moi à l'`école primaire et qui se faisait tout le temps taper dessus.
{S}Un petit blond à lunettes, abonné au premier rang et qui avait tout le temps des bonnes notes en classe, apportait des crêpes à la maîtresse pour mardi gras et mangeait du porc à la cantine.

- Lexique familier : « môme »
- Motif qui revient « et » + (« que » | « qui »)

{S}Enfin, c'est bien pour lui qu'illait rencontré cette fille.
{S}Au moins, il se passe des choses dans sa vie.
{S}Alors que pour moi c'est kif-kif demain.

{S}D'après ce que tout le monde dit, c'est un toubab, enfin un Blanc, un camembert, une aspirine quoi… Alors le frère de Samra, celui qui a un gant de boxe à la place du cerveau, il veut sa peau au pauvre mec, alors que le seul crime qu'illait jamais commis c'est d'`avoir donné un peu d'amour à sa sœur.

c'est trop dommage qu'illait pas continué l'`école.
c'est à cause de la prison.
{S}Il m'a raconté qu'il s'était fait embarquer dans une histoire par des potes mais il veut pas me dire quoi - « c'est pas de ton âge ».

- Forme phonétique[32] : « illait »

2.20 – Utilisation de « vraiment » : surreprésenté / familier

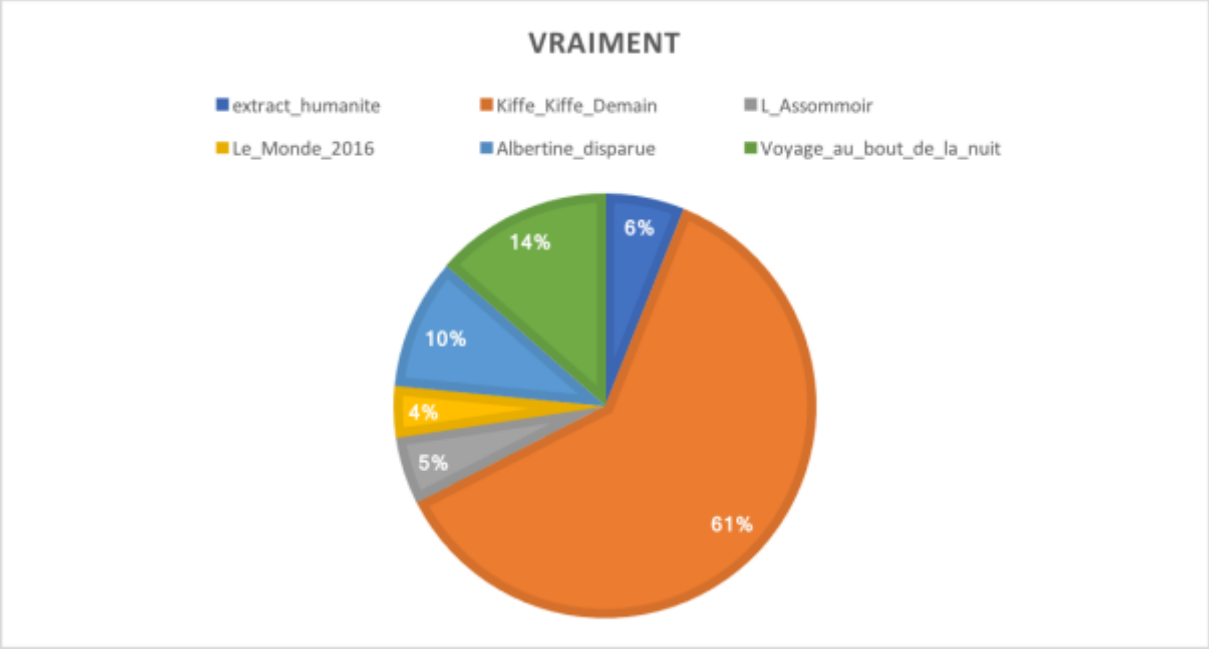


Figure 28 comparaison entre les textes/registres de la présence de "vraiment" dans ces derniers

[31] Surutilisation de "que" / à la place d'autres pro relatifs - Favart 1966-2006
[32] Sommant 2005

→ Fréquence bien discriminante pour le registre familier puisque *Kiffe Kiffe demain* se dégage nettement.

Figure 29 automate pour identifier le motif "vraiment" + ADV

→ descripteur « vraiment bien » « vraiment beau » non validé car présent dans tous les textes de tous les registres : voir annexes.

2.21 – et (qui | que)

→ descripteur qui se retrouve parfois en contexte lors des tests.

```
import re, os, nltk
from nltk import word_tokenize, sent_tokenize


def calculFreqEtque(texte):
    # print("_________"+texte)# test à décommenter
    texte=open(texte).read()
    motif=re.compile("et qui | et que")
    extract=re.findall(motif, texte)
    nbrMots=len(re.split("\W", texte))
    # print (extract) # test à décommenter
    return len(extract)/nbrMots

def toutRep(dossier):
    dicoFreq={}
    for fichier in os.listdir(dossier):
        if fichier[0]==".":continue
        fichier=os.path.join(dossier, fichier)
        dicoFreq[fichier]=calculFreqEtque(fichier)
    return dicoFreq

print(toutRep("corpus"))
```

Figure 30 script pour identifier le motif et (qui | que)

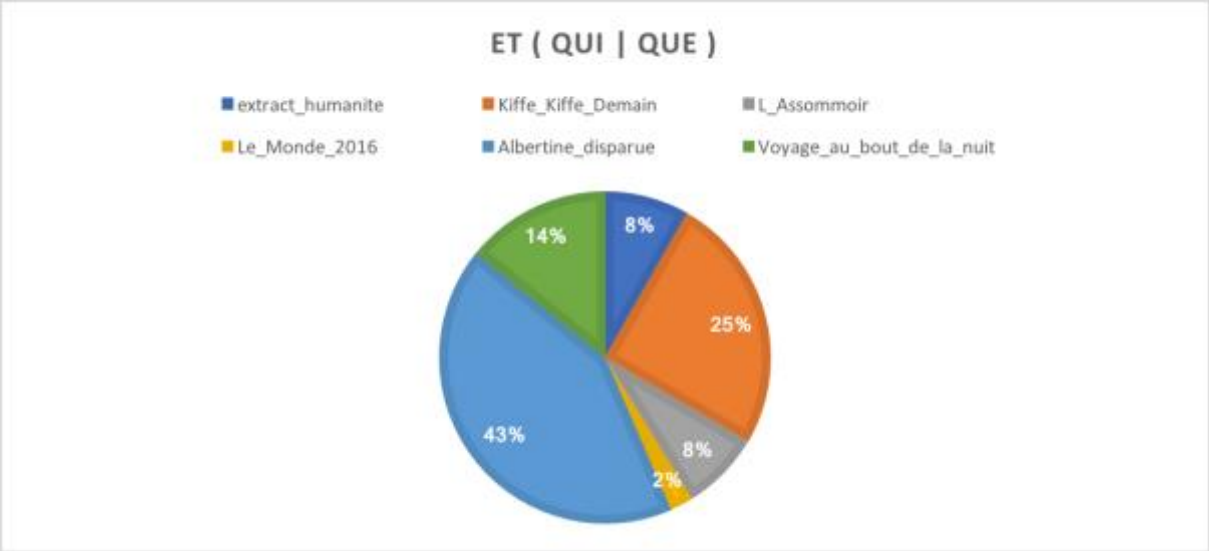


Figure 31 comparaison entre les textes/registre de la présence du motif dans ces derniers

→ descripteur non confirmé puisque l'œuvre avec la plus forte fréquence relative du motif est associée au registre soutenu. Toutefois, nous remarquons également que les deux œuvres suivantes sont celles du registre familier : *Kiffe Kiffe Demain* et *Voyage au bout de la nuit*.

2.22 – ou quoi

→ réutilisation du même script avec une modification du motif à rechercher

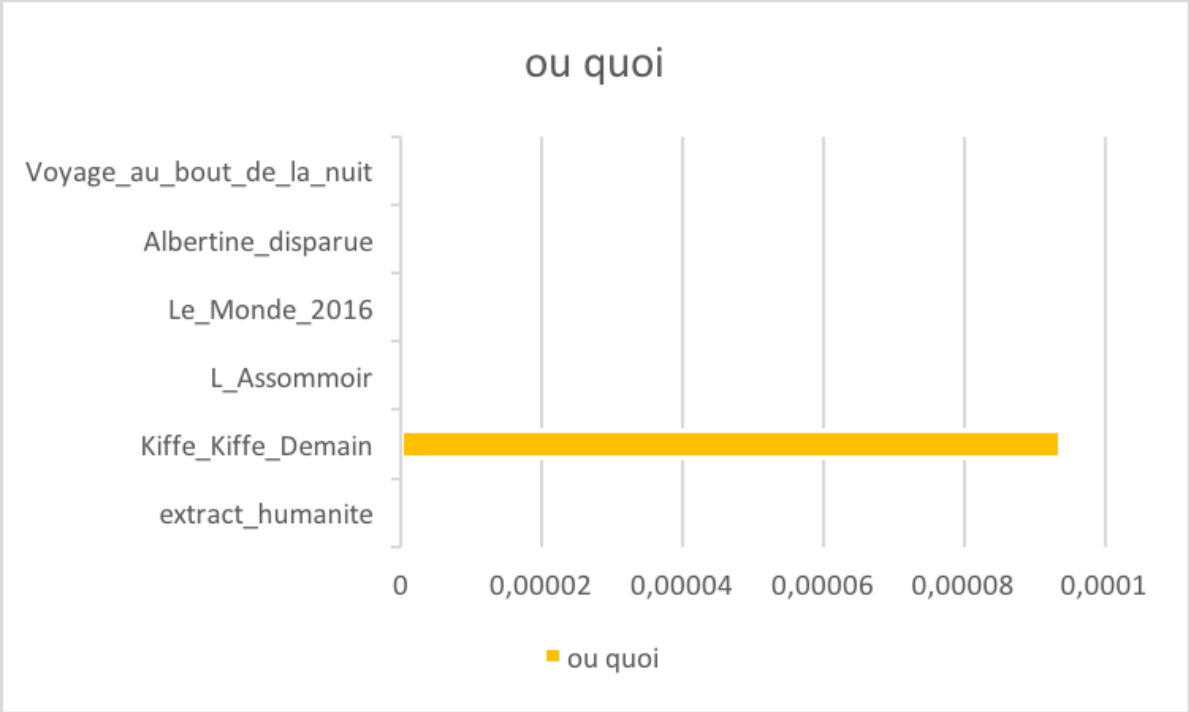


Figure 32 graphe des fréquences relatives du motif

→ Le descripteur est largement validé par sa seule présence dans l'œuvre *Kiffe Kiffe demain* : nous pouvons aussi penser que c'est un descripteur « contemporain » car absent des deux autres œuvres associés au registre familier mais plus anciennes.

2. 23 – et même que

→ réutilisation du même script avec une modification du motif à rechercher

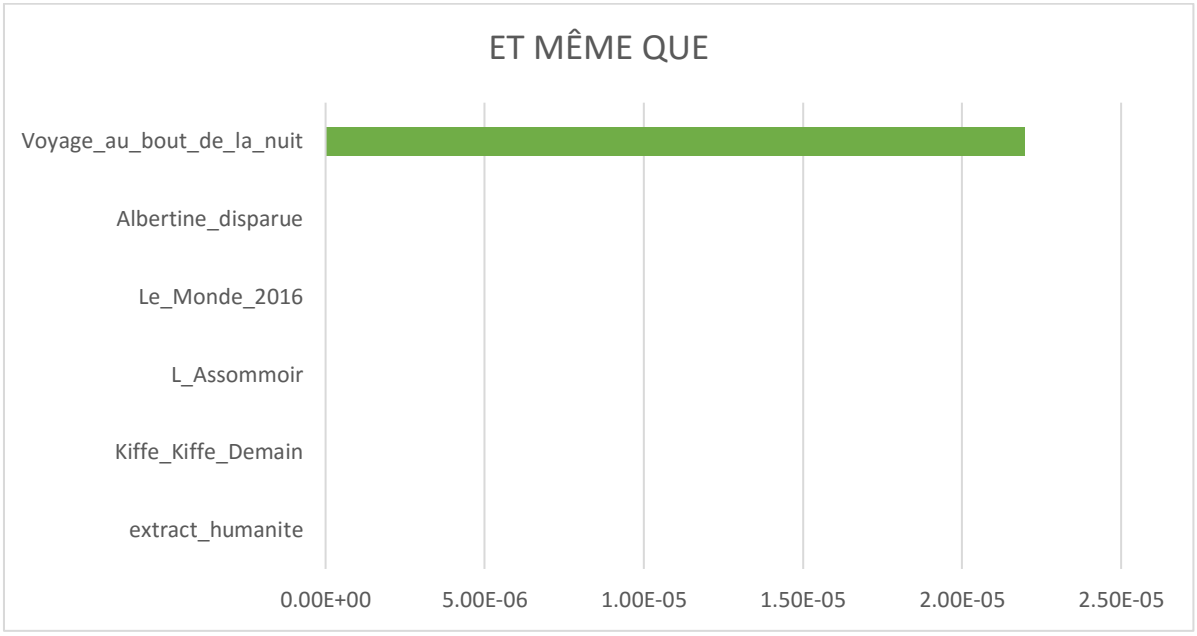


Figure 33 graphe des fréquences relatives du motif

→ Même conclusion que pour « ou quoi » descripteur largement confirmé car uniquement présent dans une œuvre associée au genre familier. Nous pouvons cependant nuancer ces deux constatations par le genre littéraire des deux textes : serait-ce un trait stylistique de l'auteur ou bien un descripteur discriminant pour le registre ou bien un descripteur envisagé comme discriminant par l'auteur qui tend à mimer ce registre ?

2.24 – Des fois / parfois

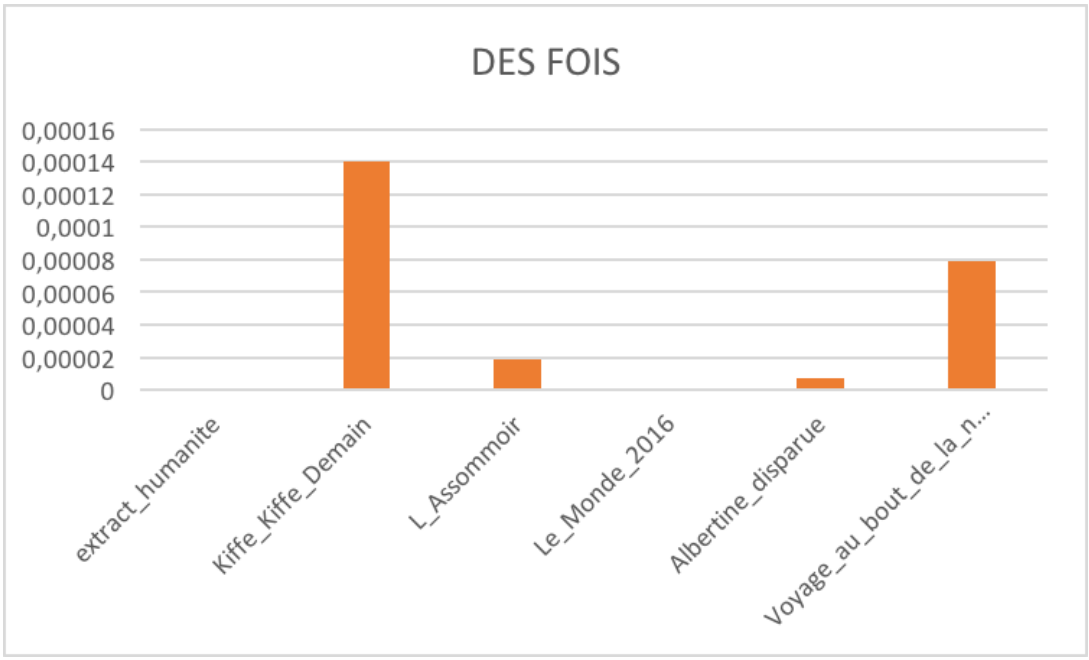


Figure 34 graphe des fréquences relatives du terme dans les différents textes/registres

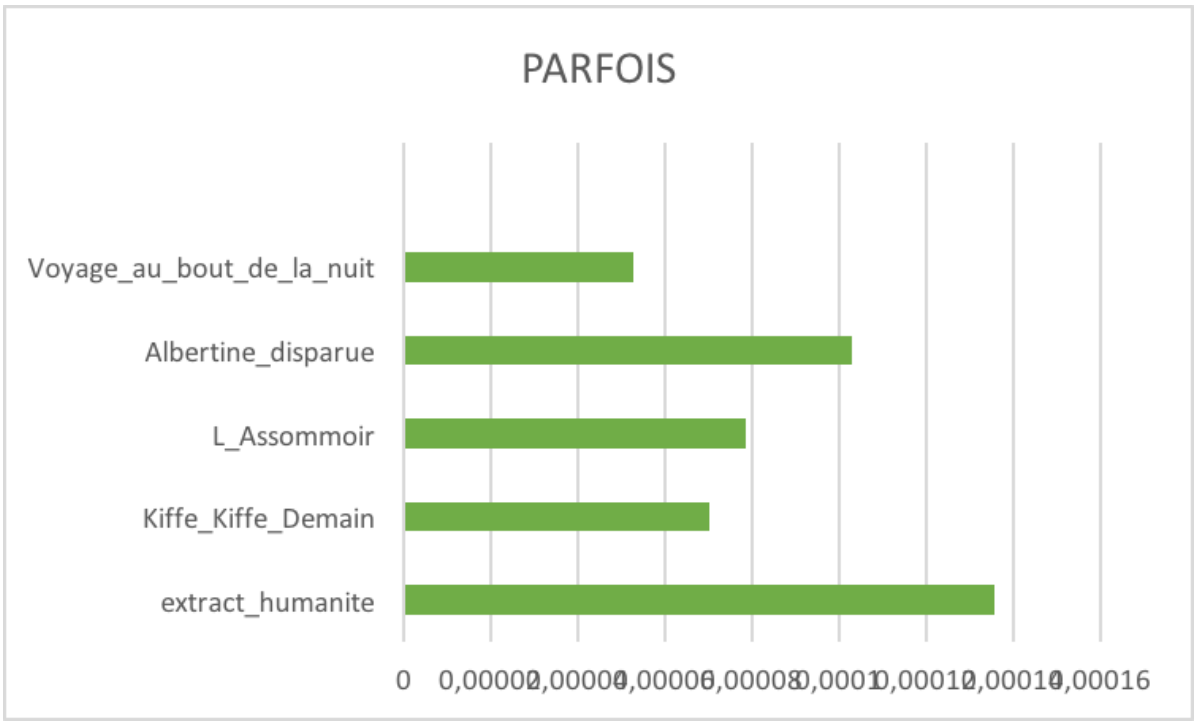


Figure 35 graphe des fréquences relatives du terme dans les différents textes/registres sans Le Monde

→ dans la figure 35, j'ai sciemment enlevé les résultats du monde car la fréquence relative était tellement supérieure aux autres que le graphe ne donnait rien :

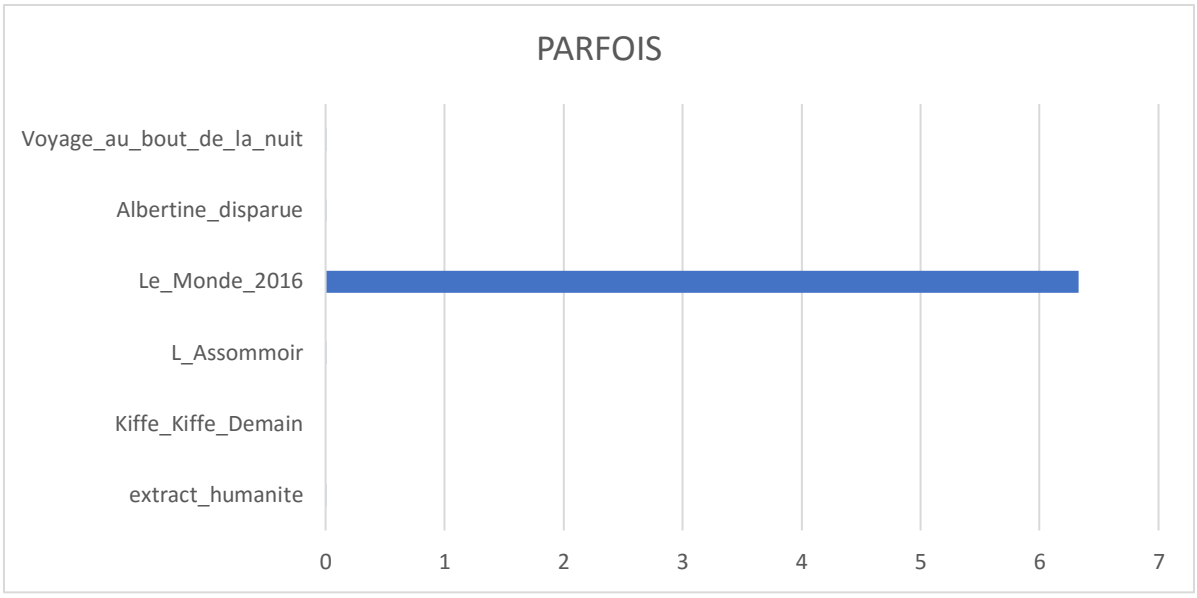


Figure 36 Figure 35 graphe des fréquences relatives du terme dans les différents textes/registres avec Le Monde

→ Le descripteur est confirmé car la forme non standard se retrouve en plus grand nombre dans les œuvres associées au familier (*Kiffe Kiffe Demain* < *Voyage au bout de la nuit* < *L'Assommoir*) tandis que les formes standards dans les œuvres du registre courant (*Le Monde* < *L'Humanité* < *Albertine disparue*).

2.25 – (sa | son) x à (elle | lui)

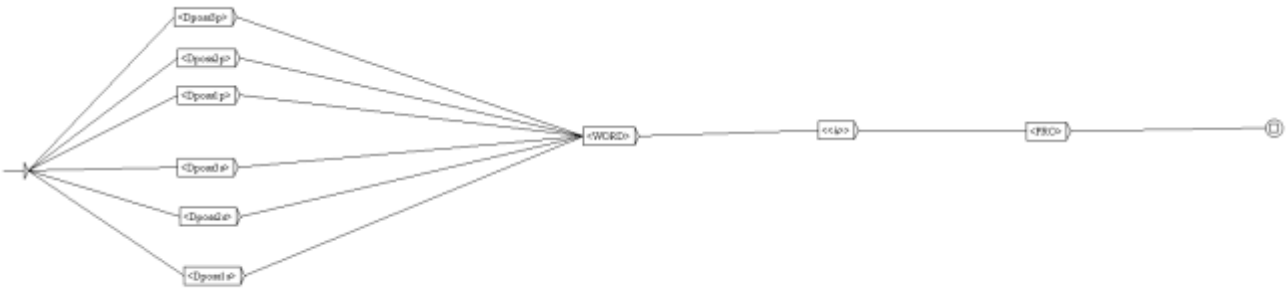


Figure 37 automate pour identifier le motif

L'Assommoir

{S}Le soir, quand Gervaise se retrouva chez elle, elle resta abêtie sur une chaise.{S} Il lui semblait que les pièces étaient désertes et immenses.{S} Vrai, ça faisait un fameux débarras.{S} Mais elle n'avait bien sûr pas laissé que maman Coupeau au fond du trou, dans le petit jardin de la rue Marcadet.{S} Il lui manquait trop de choses, ça devait être un morceau de sa vie à elle, et sa boutique, et son orgueil de patronne, et d'autres sentiments encore, qu'elle avait enterrés ce jour-là.{S} Oui, les murs étaient nus, son coeur aussi, c'était un déménagement complet, une dégringolade dans le fossé.{S} Et elle se sentait trop lasse, elle se ramasserait plus tard, si elle pouvait.

- Présence de discours indirect libre identifiable avec des éléments oraux tels que "vrai, ça faisait"

- Présence des "ça" notable car en sur présents dans l'extrait
- Descripteur lié au DIL

{S}Dès que le zingueur eut mangé sa soupe, il ronfla.{S} Le lendemain, il s'éveilla très bon enfant.{S} Il avait un reste de la veille, tout juste de quoi être aimable.{S} Il assista à la toilette de la petite, attendri par la robe blanche, trouvant qu'un rien du tout donnait à cette vermine un air de vraie demoiselle.{S} Enfin, comme il le disait, un père, en un pareil jour, était naturellement fier de sa fille. {S} Et il fallait voir le chic de Nana, qui avait des sourires embarrassés de mariée, dans sa robe trop courte.{S} Quand on descendit et qu'elle aperçut sur le seuil de la loge Pauline, également habillée, elle s'arrêta, l'enveloppa d'un regard clair, puis se montra très bonne, en la trouvant moins bien mise qu'elle, arrangée comme un paquet.{S} Les deux familles partirent ensemble pour l'église.{S} Nana et Pauline marchaient les premières, le paroissien à la main, retenant leurs voiles que le vent gonflait ;{S} et elles ne causaient pas, crevant de plaisir à voir les gens sortir des boutiques, faisant une moue dévote pour entendre dire sur leur passage qu'elles étaient bien gentilles.{S} Madame Boche et madame Lorilleux s'attardaient, parce qu'elles se communiquaient leurs réflexions sur la Banban, une mange-tout, dont la fille n'aurait jamais communié si les parents ne lui avaient tout donné, oui, tout, jusqu'à une chemise neuve, par respect pour la sainte table.{S} Madame Lorilleux s'occupait surtout de la robe, son cadeau à elle, foudroyant Nana et l'appelant « grande sale », chaque fois que l'enfant ramassait la poussière avec sa jupe, en s'approchant trop des magasins.

Kiffe Kiffe Demain

{S}Le problème, c'est que notre scénariste à nous, il a aucun talent.
{S}Il sait pas raconter de belles histoires.
{S}Ma mère, elle s'imaginait que la France, c''était comme dans les films en noir et blanc des années soixante.
{S}Ceux avec l''acteur beau gosse qui raconte toujours un tas de trucs mythos à sa meuf, une cigarette au coin du bec.
{S}Avec sa cousine Bouchra, elles avaient réussi à capter les chaînes françaises grâce à une antenne expérimentale fabriquée avec une couscoussière en Inox.
{S}Alors quand elle est arrivée avec mon père à LivryGargan en février 1984, elle a cru qu'ils avaient pris le mauvais bateau et qu'ils s'`étaient trompés de pays.
elle m'a dit que la première chose qu''elle avait faite en arrivant dans ce minuscule F2,

{S}Il lui manque plus que son abonnement à elle et c'est une femme accomplie.
{S}Qu'est-ce que je peux demander de plus ?{S} Vous pensiez que j'al ais dire « rien » ?{S} Eh ben, si, il me manque encore des tas de trucs.
{S}Ici, y a plein de trucs à changer… Tiens, ça me donne une idée, ça.
{S}Pourquoi je ferais pas de la politique ? « Du CAP coiffure à l'élection présidentielle, il n'y a qu'un pas… » c'est le genre de phrase qui reste.
{S}Faut que je pense à en faire plus des comme ça, comme les citations qu'`on peut lire dans les livres d'histoire de quatrième, style ce bouffon de Napoléon qui a dit : « À tout peuple conquis, il faut une révolte.

- Dans les deux cas le motif se trouve dans des relatives
- Forme contractée : « c'est »
- Interrogative[33] qui commence par « Qu'est-ce que » dans le contexte proche du motif.

Voyage au bout de la nuit

{S}Voyager, c'est bien utile, ça fait travailler l'imagination.{S} Tout le reste n'est que déceptions et fatigues.{S} Notre voyage à nous est entièrement imaginaire.{S} Voilà sa force.
{S}Il va de la vie à la mort.{S} Hommes, bêtes, villes et choses, tout est imaginé.{S} C'est un roman, rien qu'une histoire fictive.{S} Littré le dit, qui ne se trompe jamais.
{S}Et puis d'abord tout le monde peut en faire autant.{S} Il suffit de fermer les yeux.
{S}C'est de l'autre côté de la vie.

[33] Utilisation d'"est-ce que" pour les interrogatives - Maarit Ilmola 2012

- Et plus que ça encore, Lola, si peur, voyez-vous, que si je meurs de ma mort à moi, plus tard, je ne veux surtout pas qu'on me brûle !{S} Je voudrais qu'on me laisse en terre, pourrir au cimetière, tranquillement, là, prêt à revivre peut-être… Sait-on jamais !{S} Tandis que si on me brûlait en cendres, Lola, comprenez-vous, ça serait fini, bien fini… Un squelette, malgré tout, ça ressemble encore un peu à un homme… C'est toujours plus prêt à revivre que des cendres… Des cendres c'est fini !… Qu'en dites-vous ?… Alors, n'est-ce pas, la guerre…

- Présence de la forme contractée : « on »
- Relative populaire « je ne veux pas qu'on me brûle », « Je voudrais qu'on me laisse en terre »

« Ma mort, à moi ! qu'elle hurlait à présent la mère Henrouille, je veux la voir ma mort à moi !{S} Tu m'entends !{S} J'ai des yeux pour la voir, moi !{S} Tu m'entends !{S} J'ai des yeux encore moi !{S} Je veux la regarder bien ! »

- Redondance du pronom et du nom commun
- Répétition du syntagme qui débute le discours direct
- Contexte de prise de parole directe

{S}J'ai fini par m'endormir sur la question, dans ma nuit à moi, ce cercueil, tellement j'étais fatigué de marcher et de ne trouver rien.
{S}Autant pas se faire d'illusions, les gens n'ont rien à se dire, ils ne se parlent que de leurs peines à eux chacun, c'est entendu.{S} Chacun pour soi, la terre pour tous.{S} Ils essayent de s'en débarrasser de leur peine, sur l'autre, au moment de l'amour, mais alors ça ne marche pas et ils ont beau faire, ils la gardent tout entière leur peine, et ils recommencent, ils essayent encore une fois de la placer. « Vous êtes jolie, Mademoiselle », qu'ils disent.{S} Et la vie les reprend, jusqu'à la prochaine où on essayera encore le même petit truc. « Vous êtes bien jolie, Mademoiselle !… »

- Descripteur redoublé par un second descripteur en fin de phrase
  - "et de ne trouver rien" → inversion "et de ne rien trouver"

Albertine disparue

{S}Peut-être, si elle l'avait su, eût-elle été touchée de voir que son ami ne l'oubliait pas, maintenant que sa vie à elle était finie, et elle eût été sensible à des choses qui auparavant l'eussent laissée indifférente. $
{S}Mais comme on voudrait s'abstenir d'infidélités, si secrètes fussent-elles, tant on craint que celle qu'on aime ne s'en abstienne pas, j'étais effrayé de penser que, si les morts vivent quelque part, ma grand'mère connaissait aussi bien mon oubli qu'Albertine mon souvenir. $
{S}Et tout compte fait, même pour une même morte, est-on sûr que la joie qu'on aurait d'apprendre qu'elle sait certaines choses balancerait l'effroi de penser qu'elle les sait _toutes_? $

- Le seul motif identité se trouve dans une phrase avec une chaîne de co-référents qui reprennent tous la 3 personne du singulier.

L'Humanité

{S}Toutes les contributions sont souhaitées.{S} Modestes ou importantes, elles nous permettront d'offrir avant Noël à Yuxiao le cadeau que Chunlan voulait lui offrir, notre manière à nous de réparer l'irréparable.
{S}Notre manière à nous de dire à toutes et à tous ceux, souvent sans visage et sans identité, que nous sommes à leurs côtés pour
leur apporter du réconfort et une aide concrète, qu'ils s'appellent Chunlan, Yasmina, Souleymane, Boris...
{S}Notre manière à nous de dire à toutes et à tous ceux, souvent sans visage et sans identité, que nous sommes à leurs côtés pour
leur apporter du réconfort et une aide concrète, qu'ils s'appellent Chunlan, Yasmina, Souleymane, Boris...

- Intéressant de noter que nous avons ici un descripteur présent deux fois à une phrase d'intervalle : effet de style et création d'un rythme audible.

```
→ Il apparaît que dans les œuvres dîtes littéraires mais représentatives du registre familier, cela est soit associé au discours indirect libre (pour faire entendre l'oralité dans l'écrit dans un passage narratif sans descripteur explicite de prise de parole).
→ un rythme est souvent audible lorsque nous avons le descripteur dans les textes : cela peut être donc fortement associé au style de l'auteur ou bien à l'exercice littéraire (ce qui semble être confirmé par la faible présence du descripteur dans le registre standard).
→ les descripteurs sont souvent fortement renforcés par un contexte avec soit des co-référents soit d'autres descripteurs pour le même registre.
```

→ Ce descripteur semble donc être utilisé dans un souci d'insistance. Cela nous renvoie au besoin d'expressivité évoqué par Henri Frei. Cette notion d'expressivité est ici problématique puisque nous avons affaire à des œuvres littéraires… Ainsi, les descripteurs relevés sont-ils symptomatiques d'une variation linguistique et donc discriminant pour un registre de langue ou bien sont-ils simplement des outils sciemment utilisés pour construire un registre ?

III – problèmes rencontrés et perspectives de travaux ultérieurs

A – Difficultés soulevés par les travaux

En effet, non seulement le corpus est écrit mais surtout chaque texte est représentatif à la fois d'un registre et d'un genre particulier.

- *Albertine Disparue* de Proust
  - Registre soutenu
  - Genre littéraire
- L'*humanité*
  - Registre courant / standard
  - Genre journalistique
- *Kiffe Kiffe Demain* de Guene
  - Registre familier
  - Genre littéraire
- Histoires à parler
  - Registre enfantin
  - Genre littéraire (jeunesse)
- *Voyage au bout de la nuit* - Céline
  - Registre familier
  - Genre littéraire
- *L'Assommoir* - Zola
  - Registre familier
  - Genre littéraire

La difficulté découle donc de la motivation du style. Effectivement, comme nous venons de le voir la variation est toujours causée par un élément extérieur qui influence alors la manière de dire d'un locuteur. Or, dans notre corpus, le style est construit car écrit. Autrement dit, chaque registre représenté est reconstitué. Dès lors, la variation n'est plus motivée par un « besoin » ou bien par une « régularisation » de la langue mais par un souhait de mimétisme avec un registre particulier. Cette

altération de la variation se répercute sur le registre lui-même : si la variation est motivée par le registre alors la manière d'écrire devient un effet de style.

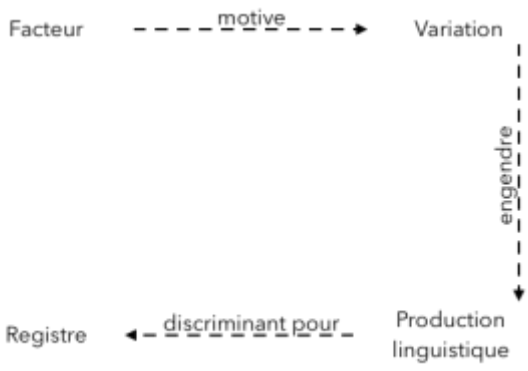


Figure 38 schéma de la motivation de la variation

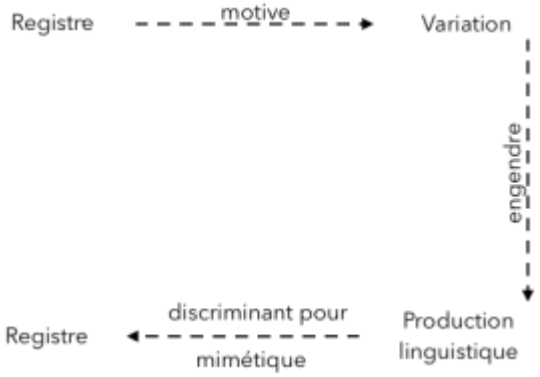


Figure 39 schéma de la variation motivée par le registre

Ainsi, non seulement le support écrit fausse la variation mais le genre également puisqu'il permet de motiver la variation par le genre lui-même.

B - Travaux sur les discours rapportés

Toutefois, si la difficulté de travailler sur un medium écrit lorsque l'oral est implicitement lié au familier est toujours là, j'ai essayé de solutionner cette impasse en m'intéressant à la représentation de cette oralité. Autrement dit, si les éléments oraux eux-mêmes ne peuvent être considérés comme symptomatique d'une variation car motivés par la représentation d'un registre, en revanche la façon d'amener ces éléments et leur place dans le récit peuvent être des éléments résultant d'une motivation échappant à l'auteur et son désir de mimétisme. Ainsi, nous allons dans un dernier temps nous intéresser à la manière dont les discours rapportés sont représentés.

Tout d'abord voici les définitions des différents discours rapportés :

- Discours rapporté : « En grammaire, le discours rapporté est un type de discours permettant à un énonciateur de citer les paroles de quelqu'un d'autre. » (source)

    - Discours direct – DD : le discours direct permet un accès direct aux paroles rapportées, cela donne une impression « d'objectivité ». Ses marques sont un verbe de parole (moyen qui permet d'influencer le discours), des marques typographiques (de ponctuations notamment) : « — », « : » par exemple.

    - Discours indirect – DI : contrairement au discours direct il n'a pas d'indépendance syntaxique. Autrement dit, il prend la forme d'un subordonnée complément d'une proposition principale dont la tête est un verbe de parole (au sens large cela peut être « penser » par exemple). Sa marque principale est donc un pronom relatif ce qui implique un lien intrinsèque entre une subordonnée relative et la présence du discours

indirect. Il engendre également une modification du temps (du discours rapporté) qui s'accorde le plus souvent au passé.

- Discours indirect libre – DIL : c'est un style indirect implicite. En d'autres termes, nous ne pouvons pas identifier de marqueurs formels indicateurs de la présence du DIL. Il n'utilise donc pas de verbe introducteur comme les discours précédents. Toutefois la présence d'une prise de parole dans le contexte du DIL est un indice.

- Discours narrativisé – DN : c'est un discours qui rapporte des évènements. Le lecteur n'a pas d'accès au discours rapporté puisque ce dernier est narrativisé. Il n'y a pas de monstration de la parole de l'autre.

Cette dernière étape qui n'est qu'une ébauche de réflexion par rapport à l'identification automatique des différents types de discours rapporté me permit d'appréhender les enjeux linguistiques informatiques d'une telle entreprise.

Tout d'abord, j'ai recensé grâce aux définitions ci-dessus des marqueurs explicites de la présence d'un discours rapporté : des listes de verbe de parole par exemple (an sachant que celle-ci ne pourrait pas être totalement exhaustive). Dans un premier temps j'ai donc « réussi » à dégager des proportions de discours direct et de discours indirect (les deux scripts sont dans les annexes).

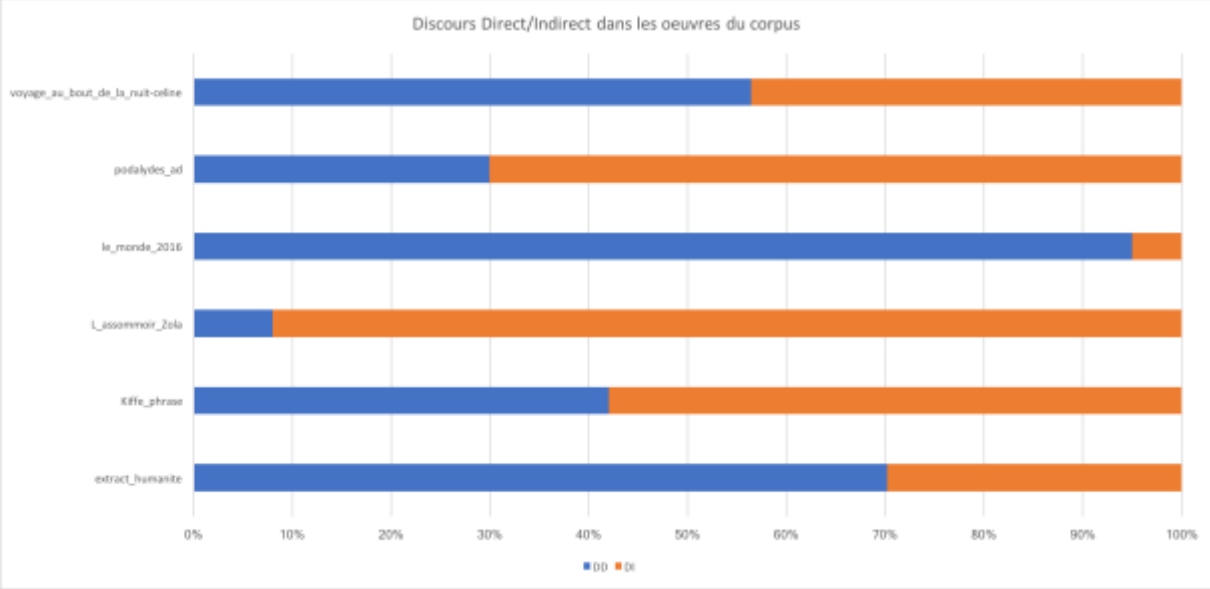


Figure 40 graphe des proportions de DD / DI dans chaque texte/registre

La forte présence de discours direct dans les deux textes de type journalistique sont expliqués par les règles très strictes de ces genres par rapport à la monstration de la parole de l'autre : cela rend les marqueurs forts et donc facile à identifier en traitement automatique des langues. Les deux œuvres dont la proportion de discours direct est la plus forte sont deux œuvres du registre familier : *Kiffe kiffe demain* et *Voyage au bout de la nuit*.

Les choses furent plus difficiles pour le discours narrativisé : l'approche était lexicale puisque ce type de discours rapporté est court et porté par le verbe de parole. Quant au discours indirect libre cela fut impossible pour moi d'arriver à un résultat concluant. Le problème ne se situait pas dans l'identification automatique : je ne suis pas arrivée à ce stage. Effectivement, la première étape de définition de marqueurs formels et systématiques ne put pas être dépassée.

J'ai donc concentré mon travail sur le discours direct mais j'ai décalé mon angle de recherche. En effet, au lieu de chercher à discriminer les registres par rapport à la proportion de discours direct dans chaque œuvre j'ai essayé d'identifier les verbes de parole qui introduisent les discours rapportés. Peut-être était-ce à cet endroit que se portait la dimension discriminante.

Tout d'abord, pour un premier « balayage » simple j'ai voulu vérifier la « pauvreté » ou la « richesse » lexicale des verbes de parole. Comme nous l'avons déjà vu, il est impossible de trouver de manière exhaustive tous les verbes de parole potentiellement présents dans un texte : je ne pouvais donc pas chercher une mesure de « richesse » lexicale. Toutefois, nous pouvions penser le problème en sens inverse : le texte avec la fréquence relative du verbe « dire » la plus haute est le texte où nous avons une mesure lexicale « pauvre » avec peu de diversité quant aux verbes de parole. J'ai alors décidé de rechercher toutes les occurrences de toutes les formes du terme « dire » : « j'ai dit », « tu dis », il lui dit », « nous lui avons dit » ... Encore une fois, trouver une multitude de termes dont j'ignorais la forme me posait problème : j'ai donc décidé d'utiliser les fichiers étiquetés retournés par Tree Tagger et de chercher le lemme du terme, c'est-à-dire, « dire ».

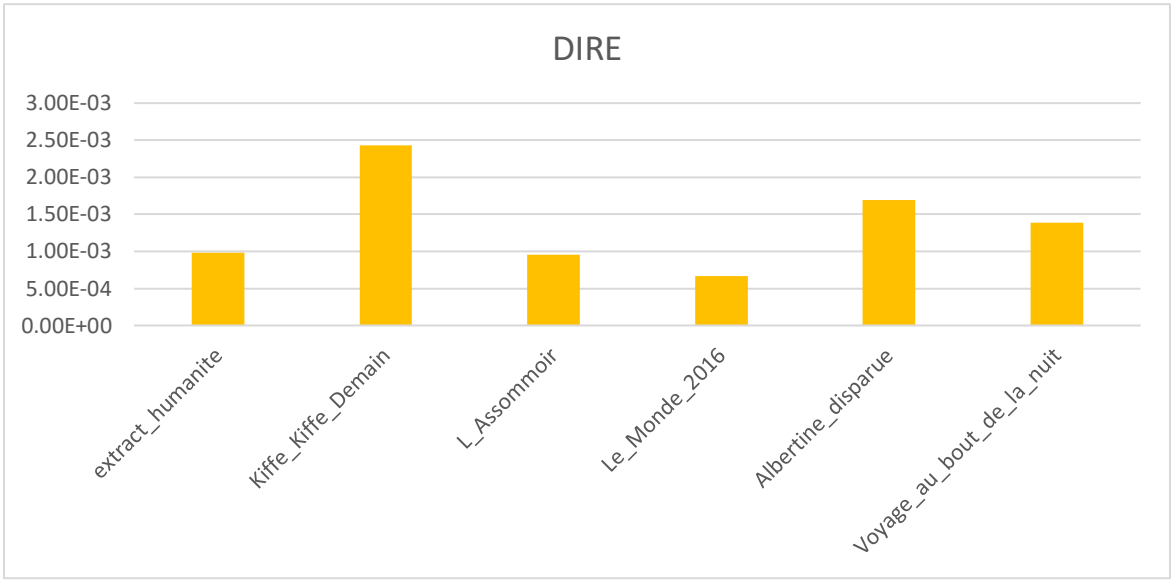


Figure 41 graphe des fréquences relatives du verbe "dire" dans toutes les œuvres

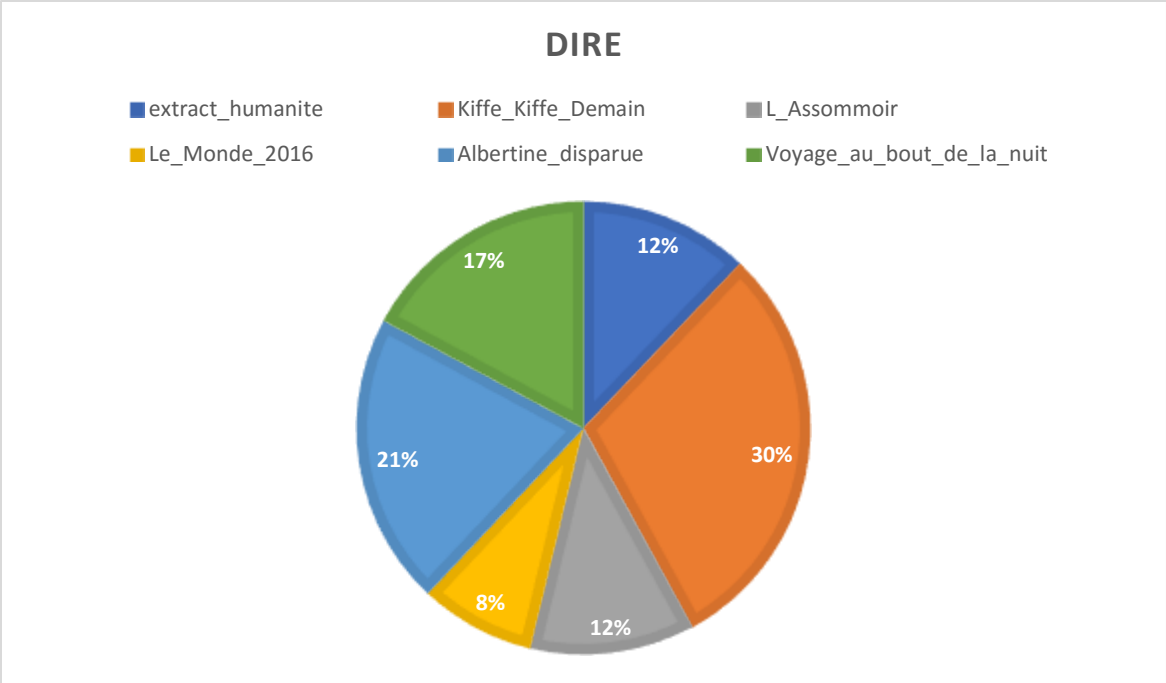


Figure 42 comparaison des fréquences relatives du lemme "dire" dans chaque registre

Il apparaît nettement une fréquence plus haute pour l'œuvre *Kiffe Kiffe Demain* associé au registre familier. Ce premier résultat est contrebalancé par la position en seconde place d'*Albertine Disparue* représentatif du registre soutenu. Toutefois, l'écart entre les deux fréquences permettrait de confirmer ce premier test : si la fréquence du verbe dire est haute cela sous-entend une certaine pauvreté lexicale quant aux verbes de parole et serait alors symptomatique du registre familier.

Pour compléter ce début de piste j'ai utilisé les marqueurs très forts du discours direct pour repérer grâce à unitex les verbes de parole :

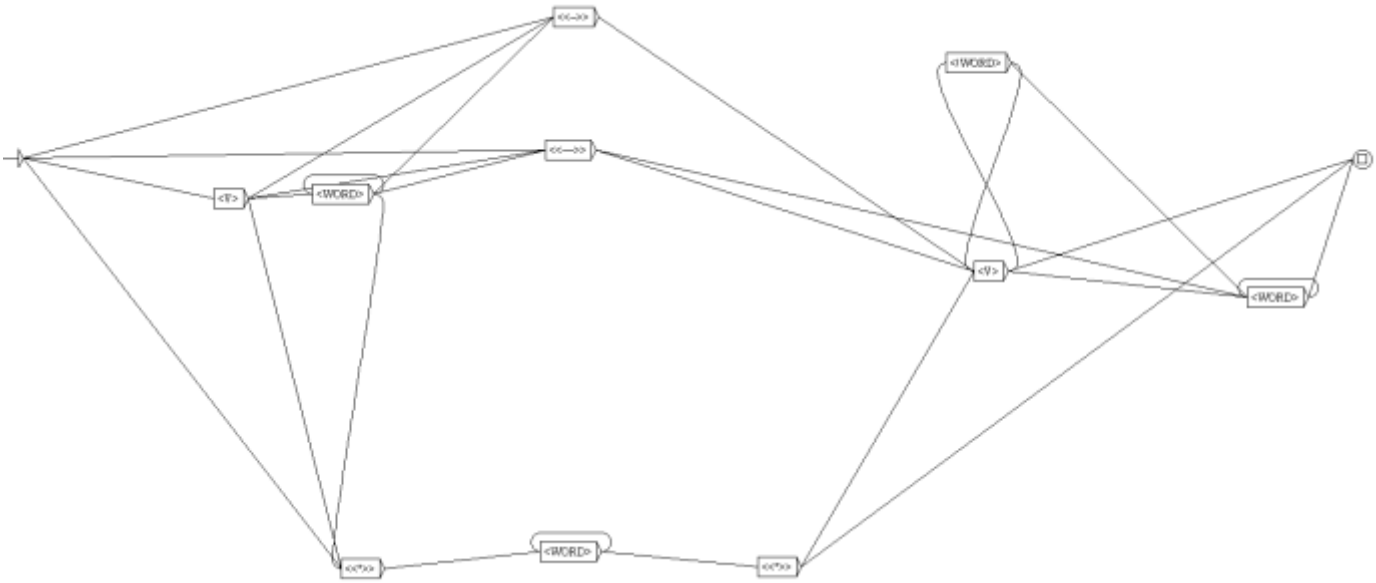


Figure 43 automate pour identifier un verbe de parole qui introduit du discours direct

Kiffe Kiffe demain

{S}Avant de partir, elle m'a regardée en fouillant dans son sac " Vieuthon " et Elle m'a fait
: - J'`ai quelque chose pour toi… elle a dit ça de sa voix aiguë, en séparant chaque syllabe
de la phrase, ça faisait débile mentale.

- Descripteur phonétique : jeu de mot « vieuthon » → « Vuitton »
- Aucun verbe de parole qui introduit le discours direct repéré par unitex
    - Verbe de parole pour le discours direct suivant : « Elle m'a fait » → verbe de parole non standard.

elle m'avait engueulée parce qu'`à mon tour de lecture, au lieu de prononcer Job, J'`ai dit "
Djob ".
{S}Jel'aiprononcé à l'anglaise.

- Monstration d'un emprunt[34] → dimension métalinguistique
- Insistance sur la prononciation explicite dans la phrase suivante
- Verbe de parole « dire » qui introduit le terme introduit entre guillemets
    - En contexte : autre verbe de parole qui exprime un discours narrativisé : « elle m'avait enguelée »

{S}Alors on dit " perf " pour signaler à quelqu'un qu'il est un peu con quand même… Voilà, Mme
Burlaud et moi, on était pas tout à fait sur la même longueur d'onde.
{S}Cela dit, je sais que c'est grâce à ça que J'`ai réussi à aller mieux.
{S}Je nie pas qu''elle m'a aidée énormément.
{S}Tiens, je lui ai même dit merci à Mme Burlaud.
{S}Un vrai merci.

- Monstration d'un terme lexical discriminant : explication du terme
    - Conscience d'un sociolecte et donc d'une variation non compréhensible par tout le monde → dimension métalinguistique
- Présence d'un lexique familier : « con », « on »
- Mal formation de la négation
- Verbe de parole dont le sujet est « on » : « dit »

{S}La " hchouma ".
{S}En plus, c'est même pas sûr qu'il nous invite.

- Emprunt encore mis en lumière avec des guillemets → dimension métalinguistique
- Aucun verbe de parole qui introduit le terme rapporté

{S}Aziz avait engagé deux " négafas ", ce sont des marieuses chargées de toute l'organisation
de la fête : décors, vêtements, maquil age, bijoux de la mariée, nourriture, tous ces
trucs-là.
c''était un super mariage, il a vraiment mis la patate Aziz.
{S}Enfin à ce qu'il paraît, parce que, effectivement, on n'a pas été invitées.
{S}Mme Dutruc, l'assistante sociale de la mairie, on ne la voit plus parce qu''elle est partie
en congé de maternité.
elle a dit qu''elle reviendrait après la naissance de son bébé

- Emprunt mis en exergue avec les guillemets → dimension métalinguistique
    - Explication du terme dans la phrase
- Aucun verbe de parole qui introduit le terme rapporté

[34]Françoise GADET 2003

{S}Ça a déjà dû lui coûter de me dire : " Ça… a ? " Même pas un vrai " ça va ".
{S}Alors " au revoir ", c''était trop demander.
{S}J'avoue que ça m'a foutu quand même un peu la haine.
{S}Mais le pire m'attendait.
j'`arrive au lycée Louis-Blanc nom propre du dico, et là, je me retrouve au milieu d''une trentaine de poufiasses décolorées, permanentées, et liberté, égalité, fraternité.
{S}Ça ressemblait pas à une rentrée des classes.
j'`avais l'impression d'attendre pour un casting.
{S}Ils étaient tous archibranchés, " fashion ", comme ils disent à la télé.
{S}Alors moi, avec mon khôl autour des yeux et mon jean contrefaçon, je me sentais pas tout à fait dedans.
{S}Ensuite, on nous a appelés pour monter dans les classes par groupes.
{S}Notre professeur principal, c'est une femme.

- Au sens strict du terme nous n'avons pas de verbe de parole qui introduit le discours rapporté. Toutefois les « : » et les guillemets permettent de lier le terme « dire » aux propos rapportés.
- Dans le contexte nous trouvons du lexique familier voire vulgaire : « ça m'a foutu (…) la haine ».
- Formes contractées présentes : « ça », « c'était »

L'Assommoir

{S}Elle soiffait à tirelarigot, et appelait Mes-Bottes " mon fiston ".
{S}Derrière elle, la machine à soûler fonctionnait toujours, avec son murmure de ruisseau

- Reprise d'une formulation associée à le discours d'un autre locuteur : monstration de la parole et mise à distance dans une même action.
- Verbe de parole : « appelait »

{S}Ils s'appelaient " sacré animal ", d'un ton où perçait une pointe de tendresse.
{S}Comme on les regardait, ils finirent par se promener doucement côte à côte, le long des maisons, tournant sur eux-mêmes tous les dix pas.
{S}Une conversation très-vive s'était engagée.
{S}Brusquement, Coupeau parut se fâcher de nouveau, tandis que l'autre refusait, se faisait prier.

- Mise en exergue de l'oralité avec une précision sur la manière de dire : « d'un ton où perçait une pointe de tendresse »
- Verbe de parole : « appelaient »

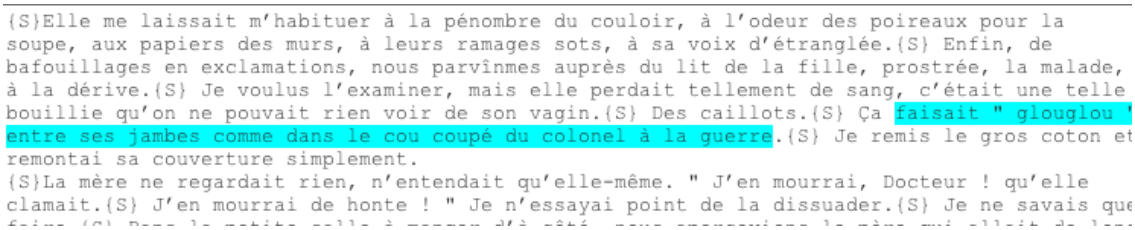
{S}Et, comme ce mot de " supplément " était accueilli par des exclamations furibondes, il donna le détail : vingt-cinq litres, au lieu de vingt, nombre convenu à l'avance ;{S} les œufs à la neige, qu'il avait ajoutés, en voyant le dessert un peu maigre ;{S} enfin un carafon de rhum, servi avec le café, dans le cas où des personnes aimeraient le rhum.
{S}Alors, une querelle formidable s'engagea.

- Aucun verbe de parole
- Monstration du terme avec « ce mot » → dimension métalinguistique

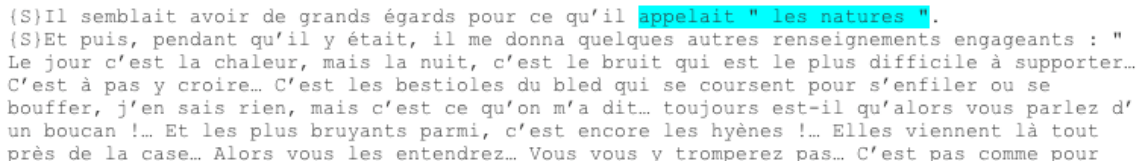
{S}Les dragées surtout, qu'il appelait des " amandes sucrées ", lui mettaient une petite mousse aux lèvres, tant elles lui chatouillaient la gargamelle.
{S}Depuis un an, il ne vivait plus que de bonbons.
{S}Il ouvrait les tiroirs, se fichait des culottes tout seul, quand Virginie le priait de garder la boutique.
{S}Souvent, en causant, devant des cinq ou six personnes, il ôtait le couvercle d'un bocal d'

- Verbe de parole « appelait »

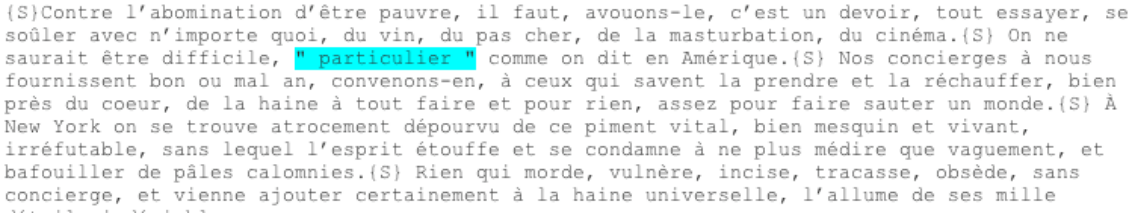
Voyage au bout de la nuit


{S}Elle me laissait m'habituer à la pénombre du couloir, à l'odeur des poireaux pour la soupe, aux papiers des murs, à leurs ramages sots, à sa voix d'étranglée.{S} Enfin, de bafouillages en exclamations, nous parvînmes auprès du lit de la fille, prostrée, la malade, à la dérive.{S} Je voulus l'examiner, mais elle perdait tellement de sang, c'était une telle bouillie qu'on ne pouvait rien voir de son vagin.{S} Des caillots.{S} Ça faisait " glouglou " entre ses jambes comme dans le cou coupé du colonel à la guerre.{S} Je remis le gros coton et remontai sa couverture simplement.
{S}La mère ne regardait rien, n'entendait qu'elle-même. " J'en mourrai, Docteur ! qu'elle clamait.{S} J'en mourrai de honte ! " Je n'essayai point de la dissuader.{S} Je ne savais que

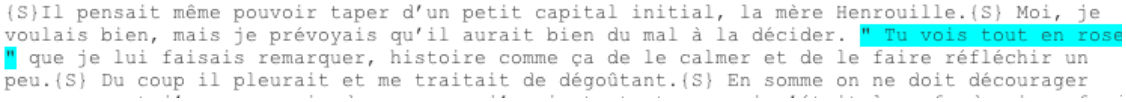
- Le terme à redoublement est mis en avant pour sa sonorité
    - Descripteur doublé directement par « ça » + VB
- Verbe de parole « faire » : « ça faisait »


{S}Il semblait avoir de grands égards pour ce qu'il appelait " les natures ".
{S}Et puis, pendant qu'il y était, il me donna quelques autres renseignements engageants : " Le jour c'est la chaleur, mais la nuit, c'est le bruit qui est le plus difficile à supporter… C'est à pas y croire… C'est les bestioles du bled qui se coursent pour s'enfiler ou se bouffer, j'en sais rien, mais c'est ce qu'on m'a dit… toujours est-il qu'alors vous parlez d'un boucan !… Et les plus bruyants parmi, c'est encore les hyènes !… Elles viennent là tout près de la case… Alors vous les entendrez… Vous vous y tromperez pas… C'est pas comme pour

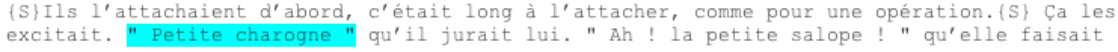
- Terme utilisé pour renvoyer à une réalité autre que celle du locuteur
    - Toujours ce double mouvement de monstration/mise à distance
    - Ici nous n'avons pas de définition de la réalité (étrangère au locuteur) évoquée par l'introduction du terme « les natures »
- Verbe de parole : « appelait »


{S}Contre l'abomination d'être pauvre, il faut, avouons-le, c'est un devoir, tout essayer, se soûler avec n'importe quoi, du vin, du pas cher, de la masturbation, du cinéma.{S} On ne saurait être difficile, " particulier " comme on dit en Amérique.{S} Nos concierges à nous fournissent bon ou mal an, convenons-en, à ceux qui savent la prendre et la réchauffer, bien près du coeur, de la haine à tout faire et pour rien, assez pour faire sauter un monde.{S} À New York on se trouve atrocement dépourvu de ce piment vital, bien mesquin et vivant, irréfutable, sans lequel l'esprit étouffe et se condamne à ne plus médire que vaguement, et bafouiller de pâles calomnies.{S} Rien qui morde, vulnère, incise, tracasse, obsède, sans concierge, et vienne ajouter certainement à la haine universelle, l'allume de ses mille

- Ici les marqueurs permettent de signaler une voix collective
    - Soutenue par les « on » qui généralisent le propos
    - Dimension de « dicton » construite par la lexie « comme on dit en X »
- Verbe de parole : « dit »

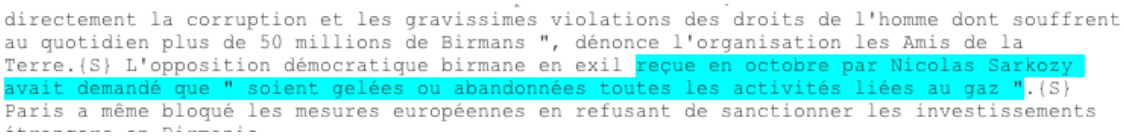

{S}Il pensait même pouvoir taper d'un petit capital initial, la mère Henrouille.{S} Moi, je voulais bien, mais je prévoyais qu'il aurait bien du mal à la décider. " Tu vois tout en rose " que je lui faisais remarquer, histoire comme ça de le calmer et de le faire réfléchir un peu.{S} Du coup il pleurait et me traitait de dégoûtant.{S} En somme on ne doit décourager

- Verbe de parole « faire » : « que je lui faisais »

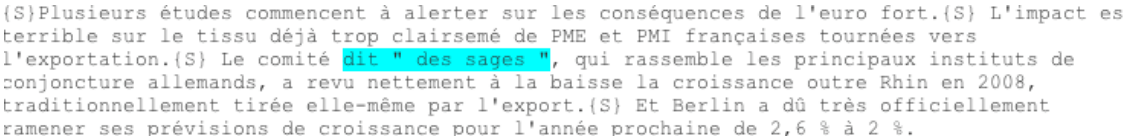

{S}Ils l'attachaient d'abord, c'était long à l'attacher, comme pour une opération.{S} Ça les excitait. " Petite charogne " qu'il jurait lui. " Ah ! la petite salope ! " qu'elle faisait

- Verbe de parole « jurer » : « qu'il jurait lui ».
    - La sémantique du terme est discriminante puisqu'elle introduit des propos familiers voire vulgaires
    - La forme lexicale du terme n'est pas discriminante pour le registre familier : elle l'associe au registre standard

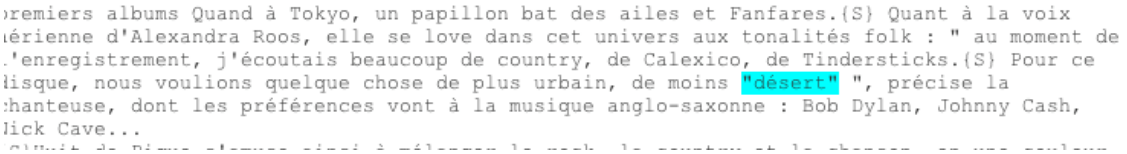
L'humanité


directement la corruption et les gravissimes violations des droits de l'homme dont souffrent au quotidien plus de 50 millions de Birmans ", dénonce l'organisation les Amis de la Terre.{S} L'opposition démocratique birmane en exil reçue en octobre par Nicolas Sarkozy avait demandé que " soient gelées ou abandonnées toutes les activités liées au gaz ".{S} Paris a même bloqué les mesures européennes en refusant de sanctionner les investissements

- Verbe de parole : « demander »
    - Bonne formation du discours rapporté avec la relative « avait demandé que x »


{S}Plusieurs études commencent à alerter sur les conséquences de l'euro fort.{S} L'impact es terrible sur le tissu déjà trop clairsemé de PME et PMI françaises tournées vers l'exportation.{S} Le comité dit " des sages ", qui rassemble les principaux instituts de conjoncture allemands, a revu nettement à la baisse la croissance outre Rhin en 2008, traditionnellement tirée elle-même par l'export.{S} Et Berlin a dû très officiellement ramener ses prévisions de croissance pour l'année prochaine de 2,6 % à 2 %.

- Verbe de parole « dit »
    - Pourrait être lié aux formes d'emprunts ou au cas de monstration d'un terme inconnu du lecteur des autres œuvres :
        - Dans les autres œuvres soit nous n'avions pas de verbe de parole et nous trouvions le terme alors introduit directement avec les guillemets ' La « Hchouma »' dans *Kiffe Kiffe Demain*.
        - Soit, nous avions le verne « appeler » : 'qu'il appelait « les natures » ' dans *Voyage au bout de la Nuit*, 'qu'il appelait « des amandes sucrées »' dans *L'Assommoir*.
    - Nous pouvons donc associer les formes précédentes à un registre plutôt familier tandis que la formulation 'dit « x »' serait plutôt représentative d'un registre standard.


premiers albums Quand à Tokyo, un papillon bat des ailes et Fanfares.{S} Quant à la voix térienne d'Alexandra Roos, elle se love dans cet univers aux tonalités folk : " au moment de 'enregistrement, j'écoutais beaucoup de country, de Calexico, de Tindersticks.{S} Pour ce lisque, nous voulions quelque chose de plus urbain, de moins "désert" ", précise la hanteuse, dont les préférences vont à la musique anglo-saxonne : Bob Dylan, Johnny Cash, lick Cave...

- Verbe de parole « précise »

→ De cette rapide comparaison en contexte nous pouvons dégager un descripteur pour le registre familier : « faire » utilisé comme verbe de parole. En outre, il semble que pour les termes utilisés de manière métalinguistique nous n'avons pas de verbe de parole pour le registre familier (contrairement au registre standard).

Conclusion

Ainsi, dans un premier temps l'état de l'art me permit de dissocier la variation linguistique du registre, puis de distinguer leurs dynamiques : la variation linguistique est un concept en mouvement qui produit un énoncé qui relève alors de tel ou tel registre. En outre, il se dégage de ces travaux préliminaires un paradigme où est lié implicitement l'oral avec le « non-standard ». Suite à ces premières constatations théoriques j'entrepris de trouver et de classer différents descripteurs pour