

4.5.3. Les adjectifs.

“Il y a trop de choses, je vous dis, et le peuple fatigué s'épuise. Il y a trop de richesses, de clarté, de musique ; il y a trop de mots, d'adjectifs, d'adverbes et de participes. Il y a trop de mouvement.”

La guerre, page 225.

C'est un phénomène connu que la distribution des adjectifs suit de près celle des substantifs et que les deux catégories sont liées l'une à l'autre.

“D'une part, écrit Étienne Brunet³⁴², la frontière entre ces deux parties du discours est très largement ouverte au mouvement des frontaliers : beaucoup de ceux-ci, qui sont nés adjectifs, fonctionnent comme substantifs et le cas inverse se produit aussi. D'autre part l'un appelle l'autre par affinité syntaxique : l'adjectif a besoin d'un tuteur et le substantif aime la compagnie.”

Le substantif leclézien est-il plutôt un solitaire, ou cherche-t-il la compagnie ? Les courbes dans le diagramme ci-dessous témoignent, à quelques interruptions près, d'une amitié croissante entre les deux entités :

³⁴² É. Brunet (1981) : p. 307.

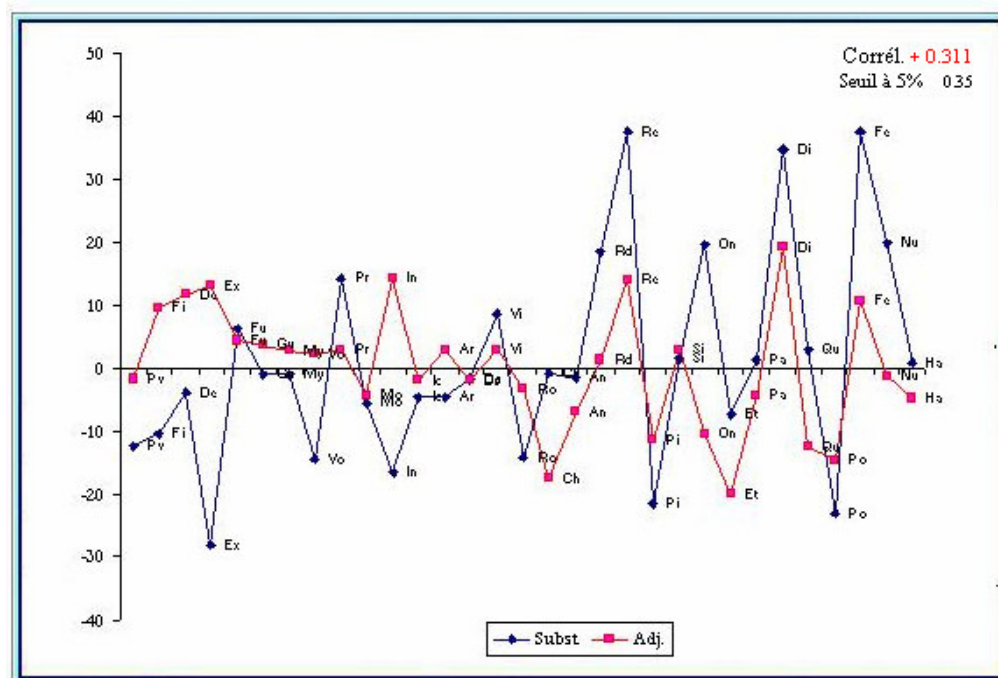


Figure n°100 : La distribution des adjectifs et des substantifs dans le corpus B (écarts réduits).

Les courbes sont en effet presque parallèles. Les deux catégories, divergentes au début de l'œuvre, se rencontrent - après une aversion déclarée au niveau de *L'extase matérielle* - dans *Le livre des fuites* où elles se rejoignent autour de l'axe des abscisses, pour ne plus se quitter. L'adjectif, après ses débuts dans l'excès, culminant dans *L'extase matérielle*, modère ses apparitions après sa rencontre décisive avec le substantif et, à partir de leur union, il n'y aura plus d'excès, plus de folie, il va se contenter du second rôle, suivant le mouvement du substantif, en fidèle compagnon.

Il convient encore, avant de préciser les résultats de cette étude, de rappeler la difficulté d'une catégorisation ferme étant donné que les catégories grammaticales n'ont pas de délimitations absolues, que les homographes sont nombreux et qu'il existe des ambiguïtés, ce qui rend souvent l'analyse syntaxique automatique

délicate. La catégorie de l'adjectif serait une des plus difficiles à traiter de façon formelle. Selon Étienne Brunet³⁴³ :

“... l'écrivain et son lecteur reconnaissent l'étymologie où les divergences puisent leur source commune – ce qui permet des ambiguïtés, des creux, des glissements sémantiques et des déguisements grammaticaux. Or l'adjectif est la catégorie qui se prête le plus volontiers à ces changements de costume.”

En effet, non seulement les nombreux adjectifs substantivés en français sont délicats à analyser, mais la confusion entre les adjectifs et les participes pose un réel problème qu'aucun chercheur n'a réellement pu résoudre. Dans notre étude, l'analyse syntaxique étant “automatique” grâce à l'ordinateur, nous nous bornons aux résultats de l'analyse proposée par l'étiqueteur, tout en reconnaissant l'existence de ces ambiguïtés.

Le logiciel Hyperbase assisté de Cordial signale que notre corpus contient 112.987 adjectifs, c'est-à-dire 4,9 % du corpus. Dans le théâtre classique du XVII^{ème} siècle on a relevé à peu près les mêmes taux. Chez Molière, les adjectifs occupaient 4,63 % du texte, chez Corneille 4,68 % et chez Racine 4,85 %. Chez Hugo, en revanche, le taux d'adjectifs est de 5,64 %.³⁴⁴

La répartition chronologique dans *Frantext*³⁴⁵ montre que la distribution des adjectifs n'est pas constante à travers les époques. On reconnaît un sommet à la fin du XIX^{ème} siècle et une dépression au XX^{ème} qui dure jusque dans les années 50 (la tranche extrême de l'étude) où l'on relève une poussée très forte d'adjectifs. L'histogramme de la distribution relative des adjectifs du corpus pourrait témoigner d'une continuation de cette tendance :

³⁴³ É. Brunet (1988) : p. 182-183.

³⁴⁴ Il est à noter que lemmatisation n'est pas faite de la même façon dans ces corpus, n'étant pas fournie par Cordial, ce qui pourrait introduire une certaine marge d'erreur dans cette comparaison.

³⁴⁵ É. Brunet (1981) : p. 338.

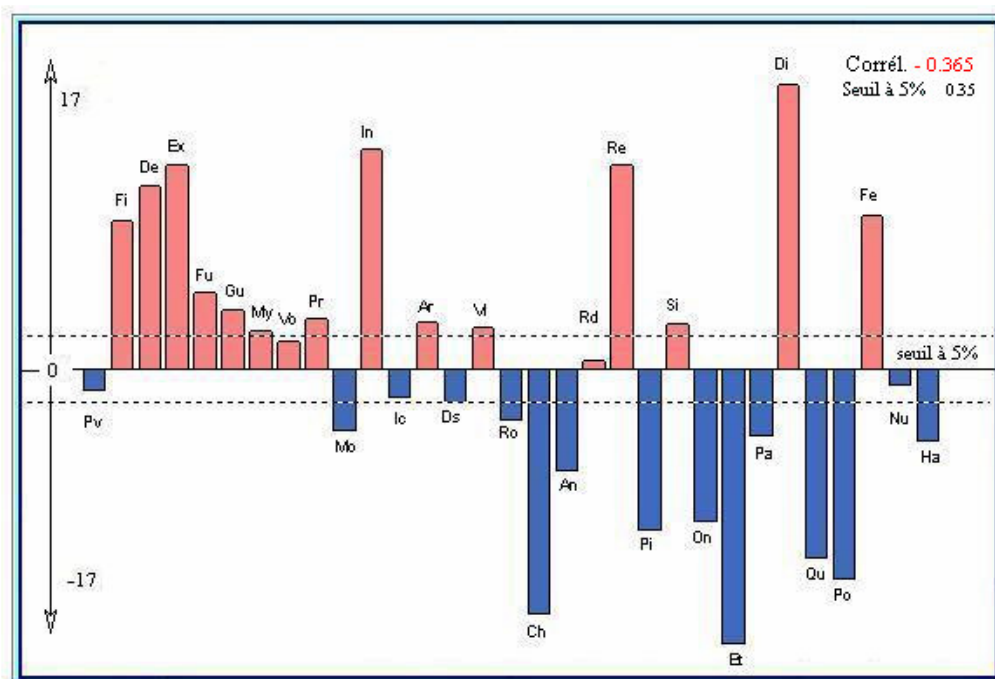


Figure n°101 : La distribution relative des adjectifs à travers le corpus B (écarts réduits).

Bien que la tendance générale soit la même que pour les substantifs lorsque l'on superpose les courbes de ces deux catégories, la distribution interne de l'adjectif seul a des traits spécifiques. Le graphique permet de distinguer deux traits principaux.

Premièrement, comme nous en avons maintenant l'habitude, nous observons l'opposition des genres littéraires. La biographie et les ouvrages d'ethnologie sont tous très riches en adjectifs tout comme ils l'étaient en substantifs. Ces genres d'écriture, avec leurs nombreuses descriptions, sont nettement favorables à l'adjectif. Les essais, souvent d'un caractère poétique, le sont aussi, mais il s'agit plus d'un effet de style comme dans l'exemple ci-dessous³⁴⁶ :

“Et la lumière, à nouveau : ineffaçable, ineffable, immatérielle. Tout apparaît. Elle vient du fond de l'espace, jusqu'à la terre. La lumière est douce et précise. Elle s'appuie sur les formes terrestres. Elle vibre et

³⁴⁶ *L'inconnu sur la terre*, p. 111.

danse sa danse continuelle, sans qu'on la voie bouger. Il n'y a pas de différence entre la lumière et la vie.

Lumière blanche du ciel brumeux.

Lumière bleue de la mer, lumière froide du nord.

Lumière jaune et chaude dans les chambres à trois heures de l'après-midi.

Lumière pâle et neuve de l'aurore, lumière glissante du soleil couchant sur les murs de pierre.”

Deuxièmement, le graphique fait apparaître une tendance chronologique. L'adjectif, après avoir été excédentaire au début de l'œuvre, régresse au fur et à mesure que l'œuvre progresse dans le temps, comme le suggère le coefficient de corrélation affecté à la courbe (-0,356).

Pour mieux cibler cette tendance dans l'œuvre romanesque, nous avons recours à notre corpus “roman” (Corpus E) dont l'histogramme ci-dessous permet de distinguer nettement deux périodes, et non pas trois, dans l'œuvre de Le Clézio :

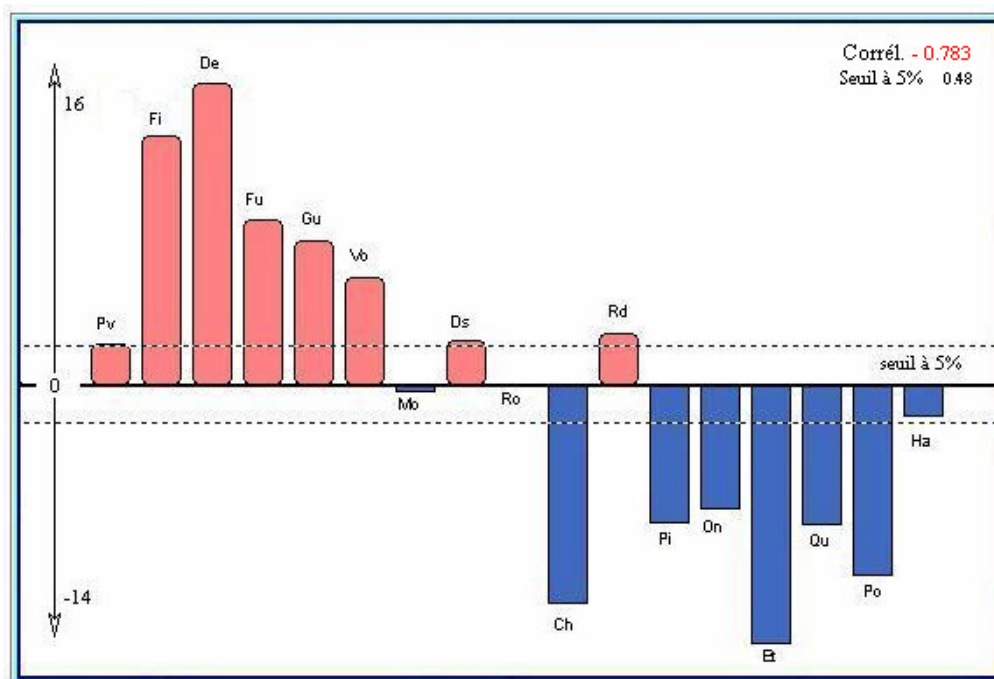


Figure n°102 : La distribution relative des adjectifs à travers le corpus E (écarts réduits).

En effet, les valeurs excédentaires sont très importantes au début de l'œuvre dans les livres qui suivent le premier roman *Le procès-verbal*. C'est dans les romans appartenant à l'école du "nouveau roman" *La fièvre*, *Le déluge* (écart réduit de +16), *Le livre des fuites*, *La guerre* et *Voyages de l'autre côté* que nous relevons de positifs d'adjectifs. L'exemple ci-dessous du roman *Le déluge* illustre jusqu'à quel point cette écriture excessive est favorable à l'adjectif³⁴⁷ :

"Mollesse et cruauté, paupières lourdes voilant les regards, bouches épaisses où vivent les dents pourries, cheveux gras collés par la sueur mauvaise, odeurs de suif, odeurs de pieds humides, crasse sous les ongles, faces encore, faces dégénérées, meurtrières et gonflées, peut-être sorties de l'enfer, là, pour se coller à votre lucarne."

Il s'agit en effet d'un trait particulier et caractéristique de ce style d'écriture singulière qu'est le "nouveau roman". Selon Maurice Nadeau³⁴⁸ :

"Puisqu'elle est astreinte au linéaire, la description est vouée au digressif. Le référent auquel elle se confronte ne saurait jamais se réduire à une ligne : il l'excède, ainsi, foncièrement."

C'est au niveau de *Mondo* que nous pouvons, encore une fois, constater la rupture dans l'œuvre avec un chavirement total dans l'usage des adjectifs chez Le Clézio. En effet, à partir du recueil de nouvelles *Mondo*, nous ne rencontrons pratiquement plus d'excédents d'adjectifs dans l'écriture romanesque de l'écrivain. Les romans et les nouvelles (à part *Désert*) sont déficitaires en adjectifs aussi bien qu'ils l'étaient en substantifs. Même *Le chercheur d'or* et *Hasard*, qui étaient excédentaires en substantifs, sont déficitaires d'adjectifs. Il est à remarquer que vers la fin de l'œuvre nous retrouvons le même mouvement légèrement progressif

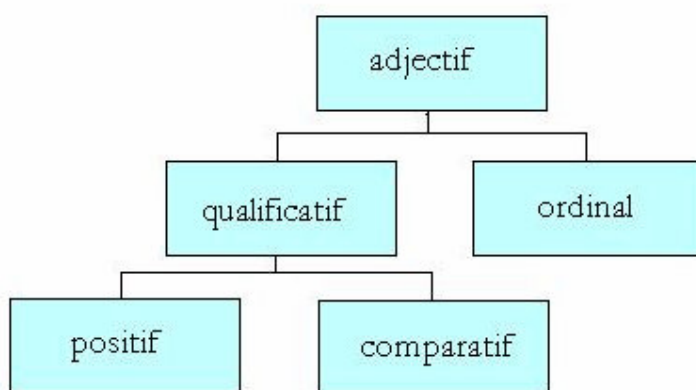
³⁴⁷ *Le déluge*, p. 210.

³⁴⁸ Cité par J. Ricardou (1990) : p. 135.

que nous avons déjà observé maintes fois ; néanmoins il est ici très atténué. Les valeurs excédentaires dans *Voyage à Rodrigues* sont à rapprocher de l'excédent important de substantifs de l'œuvre. Le style particulier de cet ouvrage est en fait favorable aux deux catégories.

On peut encore affiner l'analyse de cette catégorie adjectivale en prenant en compte ses diverses composantes.

Le logiciel Hyperbase dans sa version lemmatisée permet en effet de diviser la catégorie des adjectifs en plusieurs sous-catégories :



Parmi nos 112.987 adjectifs la plus grande partie, celle des adjectifs qualificatifs, occupe 98% avec ses 110.972 effectifs.

L'adjectif qualificatif se divise en deux catégories dans notre analyse : le positif qui représente la quasi-totalité du groupe avec 110802 effectifs (99,99%) et le comparatif presque exclusivement dans les quatre premiers livres du corpus qui ne compte que 170 effectifs (0,1%) dans le corpus entier.

L'adjectif ordinal occupe également très peu de place avec ses 1494 effectifs, mais la distribution n'est pas sans intérêt :

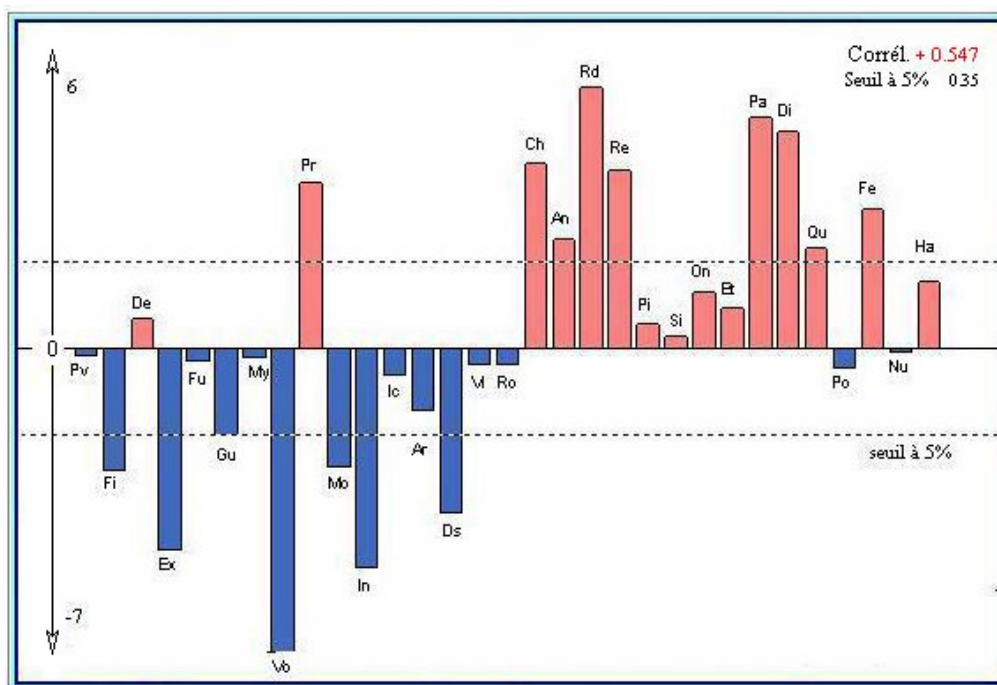


Figure n°103 : Distribution relative de l'adjectif ordinal dans le corpus B (écarts réduits).

L'histogramme divise encore une fois le corpus en deux parties distinctes, mais cette fois-ci dans le sens inverse : c'est le début qui est déficitaire et ce n'est qu'à partir de *Le chercheur d'or* que nous relevons des valeurs positives avec une forte concentration – qui n'étonnera aucun lecteur de Le Clézio – dans *Voyage à Rodrigues*. Une petite régression dans l'utilisation de l'adjectif numéral s'installe tout de même vers la fin de l'œuvre.

Nous avons vu que la distribution des adjectifs dans notre corpus suit de près celle des substantifs, avec une forte opposition entre les différents genres littéraires et des valeurs excédentaires dans la première moitié de l'œuvre qui deviennent déficitaires dans la deuxième moitié. La rupture était dans les deux études plus visible dans le corpus exclusivement voué à l'écriture romanesque.

Passons maintenant à l'étude du syntagme verbal et de ses différents composants, en commençant par les pronoms.

4.6. Les pronoms.

Les pronoms constituent, avec leurs 243.150 occurrences, presque 10% de notre corpus. Nous étudierons d'abord, dans ce chapitre, la distribution de la catégorie dans son ensemble. Nous nous intéresserons ensuite à la division en sous-catégories et nous étudierons individuellement la catégorie des personnels : enfin, nous appliquerons l'indice pronominal proposé par Charles Muller.

La distribution de la catégorie des pronoms varie beaucoup dans le corpus. L'histogramme ci-dessous rend compte de leur répartition à l'intérieur du corpus B :

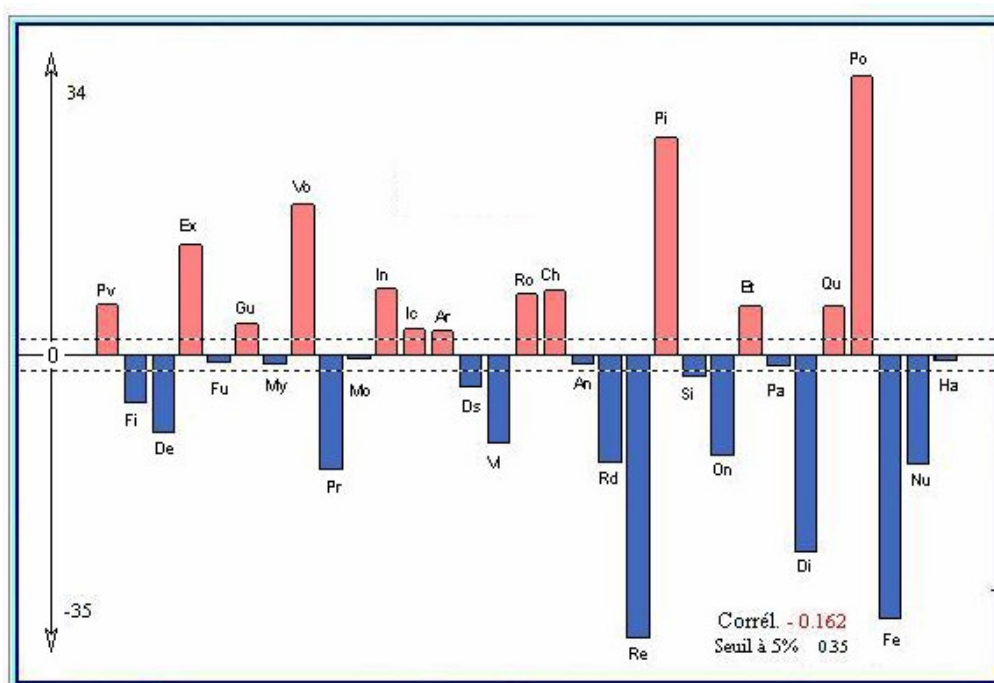


Figure n°104 : La distribution des pronoms dans le corpus B (écarts réduits).

Si l'on compare ce graphique avec celui des substantifs³⁴⁹, on constate le même type d'écarts importants et alternés, donnant une ligne générale assez chaotique,

³⁴⁹ Cf. figure n°85, p. 310.

mais les deux distributions sont inversées, notamment pour les écarts les plus grands (*L'extase matérielle*, *Voyages de l'autre côté*, *Prophéties de Chilam Balam*, *Voyage à Rodrigues*, *Le rêve mexicain*, *Printemps et autres saisons*, *Onitsba*, *Diego et Frida*, *Poisson d'or* et *La fête chantée*). Et les écarts autour de l'axe des abscisses, comme dans le graphique des substantifs, deviennent de plus en plus amples avec le temps, du côté négatif aussi bien que du côté positif. Cette relation entre les deux courbes est intéressante, car s'il est vrai que le pronom est "mis pour" un nom, on voit ici qu'en réalité la distribution de ces deux catégories se fait de manière complémentaire et qu'elles n'apparaissent donc pas dans les mêmes contextes.

Dans cet histogramme les valeurs déficitaires sont à chercher dans les ouvrages ethnologiques, la biographie et les essais – à l'exception de *L'extase matérielle* et de *L'inconnu sur la terre* qui sont excédentaires en pronoms. Les déficits les plus importants se trouvent dans *Le rêve mexicain* et dans *La fête chantée* avec des écarts réduits de -34,8 et de -32,3.

En revanche, la tendance générale des romans et des recueils de nouvelles est à la hausse avec pour point culminant *Poisson d'or* (écart réduit de +34,3). Toutefois, nous trouvons également des déficits importants – par exemple dans *Onitsba* – qui méritent d'être étudiés de près en prenant en compte uniquement l'œuvre romanesque de Le Clézio.

Lorsqu'on regarde le même histogramme en s'appuyant sur le corpus E, la dynamique du corpus romanesque est plus aisément observable :

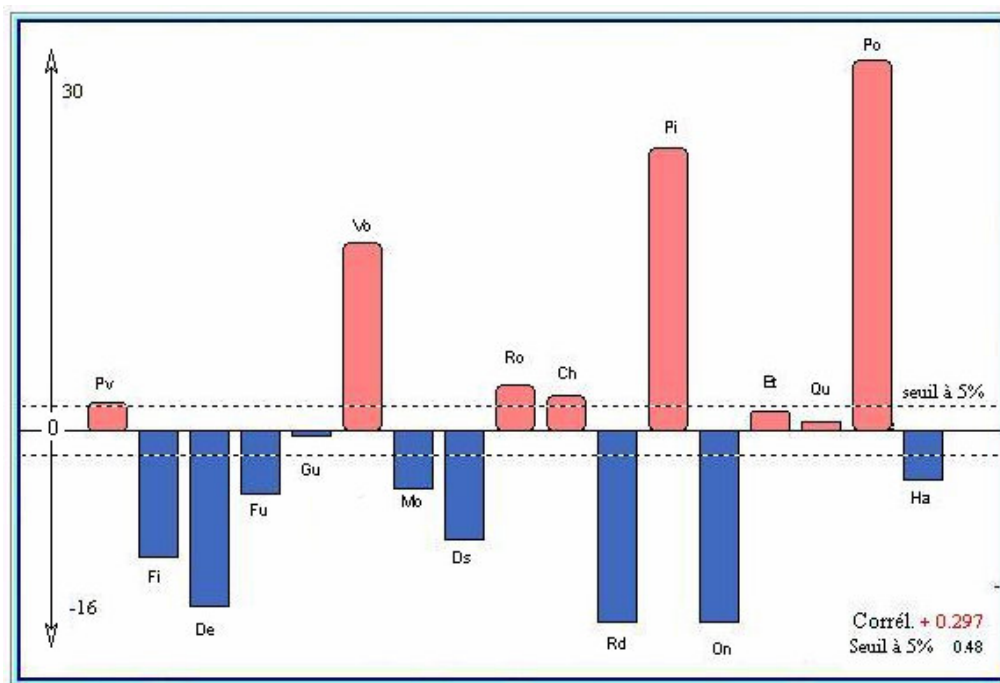


Figure n°105 : La distribution des pronoms dans le corpus E (écarts réduits).

En effet, la figure montre qu'il n'y a pas de tendance significative vers un usage plus prononcé du pronom chez Le Clézio. Bien que l'histogramme semble avoir une tendance à la hausse, les valeurs déficitaires des romans *Voyage à Rodrigues* et *Onitssha* sont trop importantes.

Afin de mieux comprendre ces fluctuations, il est intéressant d'étudier la distribution des sous-catégories pronominales.

4.6.1. Les sous-catégories pronominales.

Pour diviser la catégorie des pronoms en sous-catégories, nous avons encore une fois eu recours au logiciel Hyperbase qui grâce à l'analyseur syntaxique Cordial divise nos pronoms en sept sous-catégories : *pronoms personnels réfléchis*, *pronoms personnels non réfléchis*, *pronoms possessifs*, *pronoms démonstratifs*, *pronoms interrogatifs*, *pronoms indéfinis* et *pronoms relatifs*.

Le tableau ci-dessous indique la fréquence de ces sous-catégories pronominales dans l'œuvre de Le Clézio.

	Pers. réfléchi.	Pers. non réfléchi.	Possessifs	Démonstr.	Interrog.	Indéfin.	Relatifs
Pv	911	6525	7	1137	48	595	1199
Fi	1119	5917	2	1084	72	756	1224
De	1576	6785	4	1336	105	766	1502
Ex	734	5831	14	1552	81	763	2086
Fu	964	6210	14	1236	138	702	1709
Gu	1024	7074	7	1443	105	813	1695
My	62	570	0	146	13	66	171
Vo	1141	8683	2	1459	94	810	1995
Pz	34	225	0	86	10	24	134
Mo	1049	6348	2	1079	81	438	1420
In	962	7860	6	1825	151	793	2687
Ic	14	211	1	36	1	14	70
Ar	35	219	0	45	1	23	53
Ds	1470	8831	7	1660	166	780	2713
Vi	96	673	0	137	27	51	226
Ro	889	5719	5	1029	77	504	1383
Ch	777	10141	9	1282	165	512	2395
An	283	1550	4	220	19	119	350
Rd	218	1579	3	443	68	134	821
Re	440	2739	6	803	159	276	1528
Pi	470	6767	8	1004	50	379	1095
Si	13	109	0	29	3	8	63
On	696	4234	2	747	64	298	980
Et	1045	8285	17	1353	111	673	1847
Pa	89	649	3	107	15	52	180
Di	404	2943	8	551	62	241	1494
Qu	1247	12212	17	1381	148	747	2680
Po	493	8951	8	1131	62	429	1310
Fe	375	2316	11	549	85	276	1172
Nu	129	794	4	203	37	90	418
Ha	736	4354	5	718	43	359	1012
Tot.	19495	145304	176	25511	2261	12491	37612

Tableau n°34 : Les effectifs des différentes sous-catégories pronominales dans le corpus B.

Ce regroupement permet de faire la même observation que dans presque toutes les autres études de même caractère : les plus grands nombres d'occurrences se trouvent dans les deux premières classes, celles des pronoms personnels. Pour mieux illustrer l'irrégularité de la distribution des différentes catégories, nous

avons fusionné les deux premières colonnes en une seule catégorie : *pronoms personnels*.

Le diagramme ci-dessous rend compte de la distribution des sous-catégories en ordre hiérarchique, leur partie dans la catégorie exprimée en pourcentage :

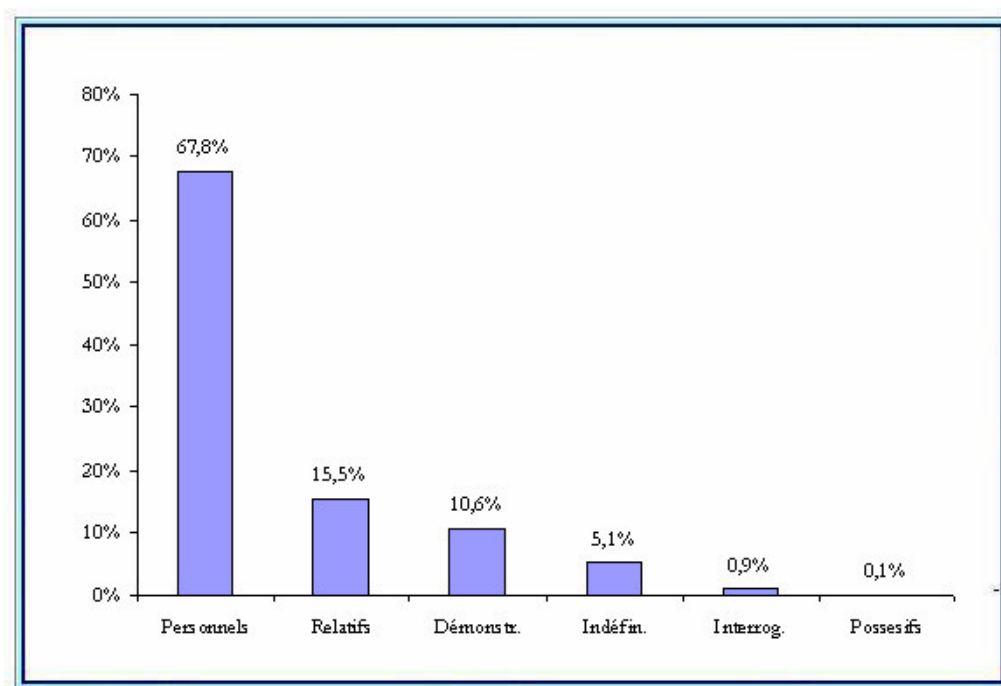


Figure n°106 : La proportion des différentes sous-catégories pronominales dans le corpus B (valeurs absolues).

Nous constatons que la grande majorité des occurrences est fournie par les pronoms personnels qui, à eux seuls, occupent 67,8% du diagramme. Les pronoms relatifs en occupent 15,5%, les démonstratifs 10,6%, et les 6% restants sont partagés entre les autres pronoms.

C'est pour cela que nous avons choisi la catégorie des pronoms personnels, pour une étude plus détaillée, afin de mieux comprendre l'usage qu'en fait Le Clézio et leur évolution dans l'œuvre.

4.6.2. Les pronoms personnels.

Les pronoms personnels sont sujets à de nombreuses ambiguïtés. Une difficulté dans l'analyse statistique est leur homographie avec les articles : *le*, *la*, *l'* et *les* sont facilement confondus avec les articles et les formes élidées sont souvent difficiles à désambigüiser.

Dominique Labbé rappelle à juste titre un autre problème : celui de leur signification³⁵⁰ ; la phrase de Flaubert : “Madame Bovary c’est moi”.

“Cependant, écrit-il³⁵¹, le fait même que ces procédés attirent l’attention prouve leur caractère inhabituel par rapport à une “norme” qui voudrait que le locuteur utilise la première personne pour se désigner, la seconde pour l’interlocuteur et la troisième pour parler des tiers ou des choses. Dès lors, ce sont des sortes de symboles conventionnels dont la substance dépend, dans une plage prédéfinie, de la situation dans laquelle ils sont prononcés et de la subjectivité du locuteur. Cette grille simple et très souple explique l’importance des pronoms dans la langue mais aussi les difficultés de l’analyse lexicale.”

Dans la littérature la situation est un peu différente étant donné qu’il s’agit souvent de fiction et de narration³⁵². Certains phénomènes stylistiques notamment, dans la littérature de l’école “nouveau roman”, mettent en jeu les conventions romanesques et l’auteur “joue” avec ses lecteurs. Maurice Nadeau³⁵³ écrit ceci à propos des innovations auxquelles se livre Le Clézio durant cette époque :

“... une narration indirecte qui tourne brusquement à la confidence, substitution du “je” au “il” et au “nous”, passages brusques chez le même individu de l’enfance à l’âge adulte et réciproquement, relations

³⁵⁰ D. Labbé (1990) : p. 32-33.

³⁵¹ *Ibid.*, p. 33.

³⁵² Cf. É. Benveniste (1966) : p. 251-257.

³⁵³ M. Nadeau (1992) : p. 184.

de situations invraisemblables (dans *Terra Amata*, Chancelade décrit sa propre mort et les sentiments qu'il éprouve en tant que cadavre)..."

Néanmoins, malgré ces difficultés, les données numériques apportent quelques informations intéressantes. Si nous regardons la distribution des pronoms personnels dans notre grand corpus, nous pouvons observer une structure très semblable à celle de la distribution de l'ensemble des pronoms :

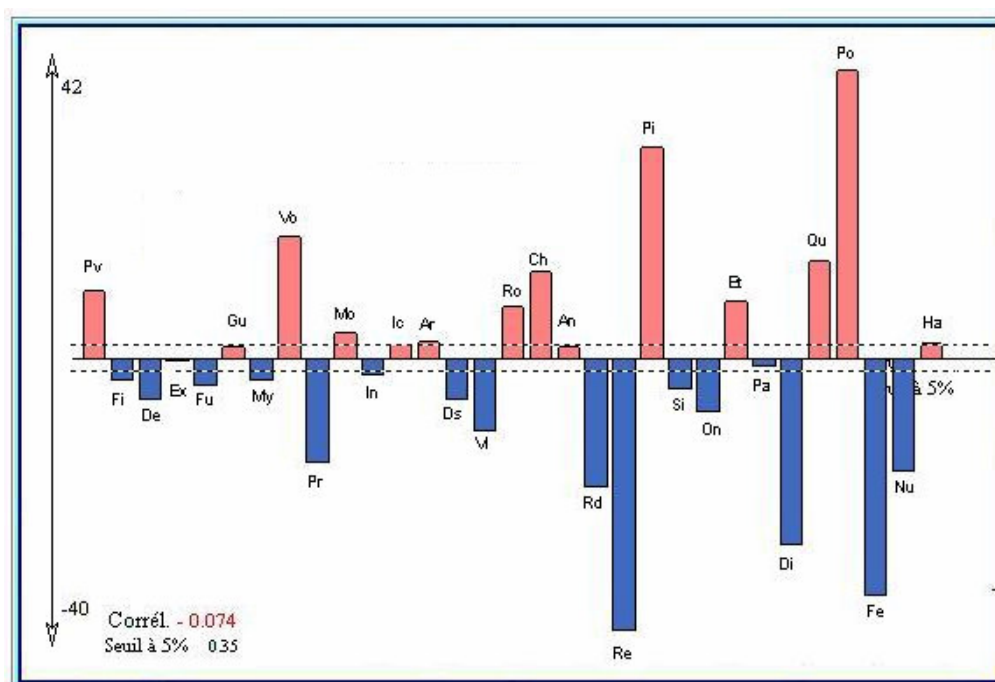
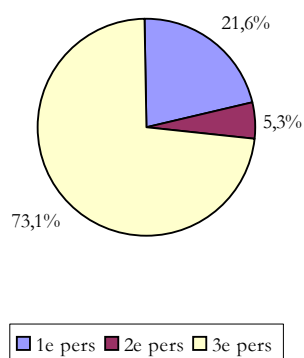


Figure n°107 : La distribution des pronoms personnels dans le corpus B (écarts réduits).

Cette ressemblance n'a rien d'étonnant sachant que les pronoms personnels occupent pratiquement 70% de la catégorie. Nous trouvons les mêmes déficits et les même excédents avec des amplitudes qui deviennent de plus en plus importantes au fur et à mesure que l'œuvre progresse. Il y toutefois quelques exceptions : les valeurs excédentaires de *L'extase matérielle* ont une autre source les pronoms personnels. *Mondo*, déficitaire dans le premier histogramme, est ici légèrement excédentaire (écart réduit de +3,7), comme l'est aussi *Angoli Mala*.

Avec l'aide du logiciel Hyperbase nous pouvons fractionner encore la catégorie pour nous intéresser aux différentes personnes. Des 164.799 pronoms personnels,

la première personne compte 35.573 occurrences, la deuxième 8.794 et la troisième 120.432. Le diagramme circulaire ci-dessous permet de mieux visualiser l'importance de la troisième personne dans l'œuvre de Le Clézio :



Ces résultats n'étonneront aucun lecteur de Le Clézio, les résultats de l'analyse statistique confirment ici ce qui est visible à l'œil nu : que l'écriture leclézienne est très favorable au récit à la troisième personne.

Les courbes ci-dessus comparent la distribution de la première, de la deuxième et de la troisième personne :

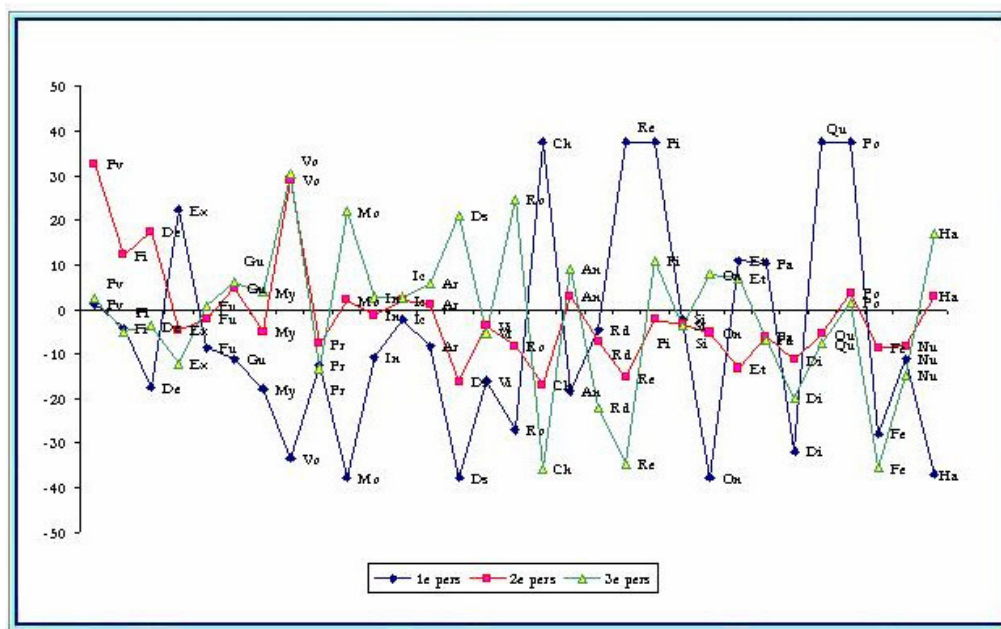


Figure n°108 : La distribution des pronoms à la première, à la deuxième et à la troisième personne dans le corpus B (écarts réduits).

tour des pronoms de la deuxième personne, et les pronoms de la troisième personne se trouvent à droite du tableau entourés des romans plus traditionnels de la deuxième période de Le Clézio. Le second axe sépare le singulier du pluriel, qui semble être favorisé par les ouvrages ethnologiques.

Il est intéressant d'analyser les différentes personnes individuellement afin de mieux connaître leur distribution propre.

Nous avons vu que la proportion des pronoms personnels à la première personne n'occupe que 21,6% des effectifs de la catégorie. L'histogramme ci-dessous montre que Le Clézio limite leur usage (relatif) à huit livres du corpus :

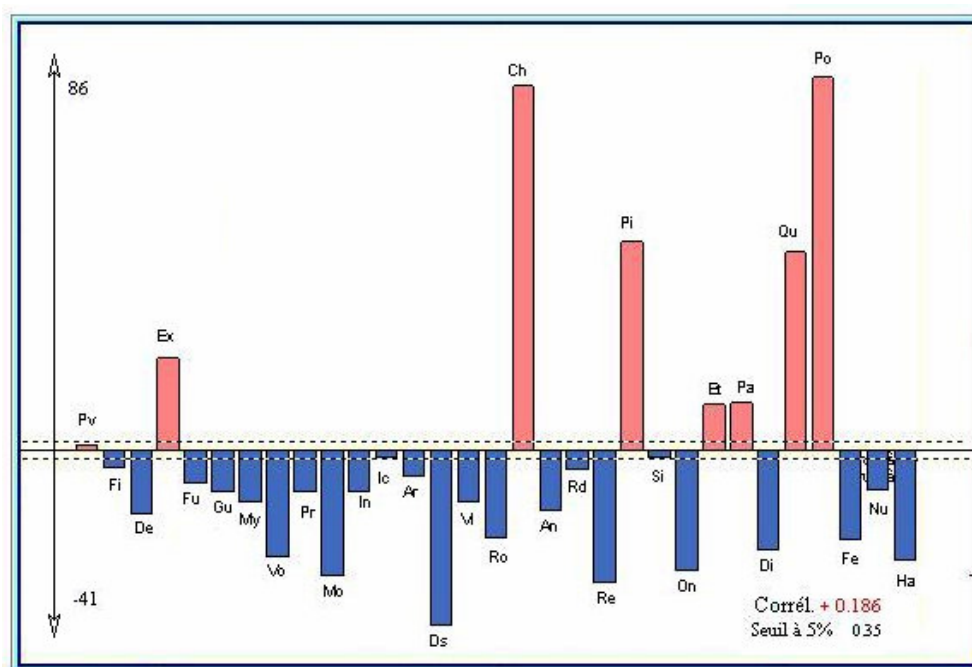


Figure n°110 : La distribution des pronoms personnels à la première personne dans le corpus B (écarts réduits).

En effet, l'usage de la première personne est excédentaire (sans pour autant atteindre un niveau significatif) dans le premier livre *Le procès-verbal*. Ensuite l'histogramme permet de constater que Le Clézio exploite de façon dominante le récit à la première personne dans *L'extase matérielle*, *Le chercheur d'or*, *Printemps et*

autres saisons, *Etoile errante*, *Pavana*, *La quarantaine* ainsi que dans *Poisson d'or* (écart réduit de +86,3).

Si nous regardons uniquement l'œuvre romanesque l'histogramme est encore plus parlant :

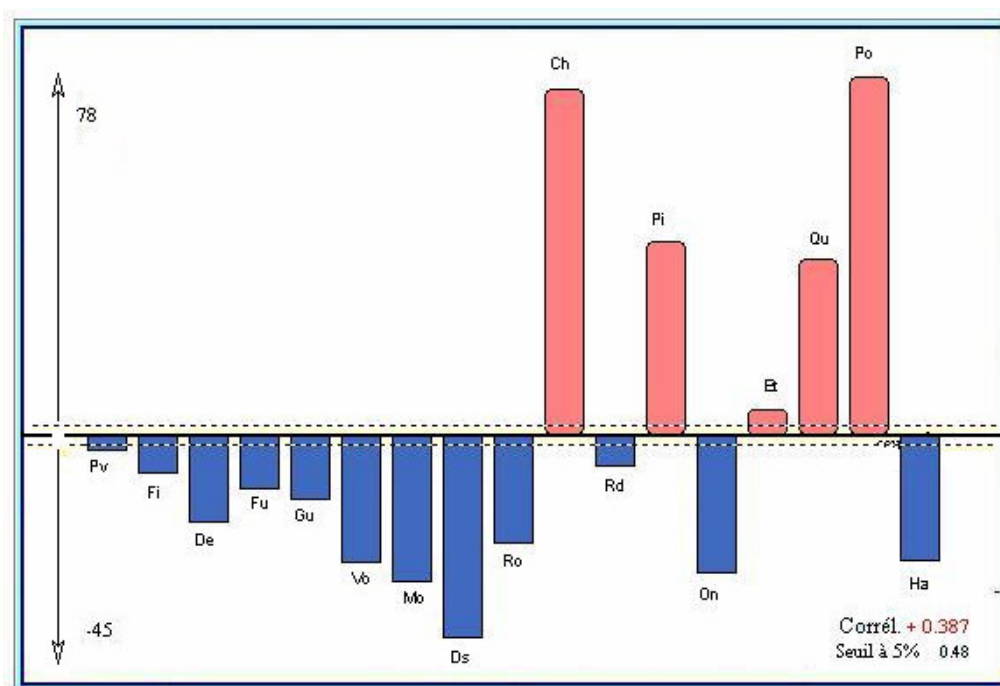


Figure n°111 : La distribution des pronoms personnels à la première personne dans le corpus E (écarts réduits).

L'histogramme souligne cette fois-ci l'importance de la première personne dans la seconde moitié de la production, surtout dans *Le chercheur d'or* et dans *Poisson d'or* mais aussi dans *Printemps* et dans *La quarantaine*, et à un moindre degré dans *Etoile errante*. Est-ce à mettre en rapport avec l'acceptation d'un certain classicisme, avec l'abandon des principes du "nouveau roman" ?

"Mon langage ne serait plus seulement celui des hommes et de leurs livres, écrit Le Clézio à cette époque³⁵⁴, mais s'adresserait à n'importe qui sur la terre, et n'importe qui pourrait le comprendre."

³⁵⁴ *L'inconnu sur la terre*, p. 314.

Toutefois, les taux déficitaires de la première période ne sont peut-être pas seulement l'effet d'un style recherché, mais témoignent aussi de ce sentiment profond chez notre écrivain d'être à l'écart des autres, de "non-activité", comme l'appelle Jean Onimus³⁵⁵ et de "non-participation" à la société avec les autres.

L'écriture leclézienne vise, au moins dans la première partie de la production, à s'inscrire dans un espace de silence qui est produit par l'effacement du "je" de l'énonciation. Il semble que Le Clézio cherche ainsi à trouver une neutralité du langage et de la langue, en accumulant les mots les uns après les autres, sans l'intervention d'aucun auteur.

"Cette idée, écrit Jean-Xavier Ridon³⁵⁶, que les mots sont toujours habités par d'autres dans le sens où ils se présentent comme le résultat d'un héritage culturel dont Le Clézio aimerait sortir. Effacer le 'je' de l'énonciation serait alors une façon de laisser parler les autres dans le langage."

Par ailleurs, le "je" de la première période n'est pas le même que celui de la fin. L'évolution du "je" du premier roman *Le procès-verbal* jusqu'à la fin du corpus est bien celle de toute une vie et de toute une œuvre. Entre le "je" du jeune homme Adam Pollo et celui du capitaine décadent Moguer dans *Hasard* le lecteur fait un long voyage à travers le monde et beaucoup d'expériences, aboutissant à ce "je" de la maturité, d'un homme d'un certain âge, parfois résigné.

Regardons maintenant la situation de la deuxième personne, totalement différente. La deuxième personne est la moins fréquente dans l'œuvre leclézienne et sa distribution diffère de celle de la première personne de façon significative. L'histogramme ci-dessous rend compte de la différence :

³⁵⁵ J. Onimus (1994) : p. 15-20.

³⁵⁶ J.-X. Ridon (1995) : p. 85.

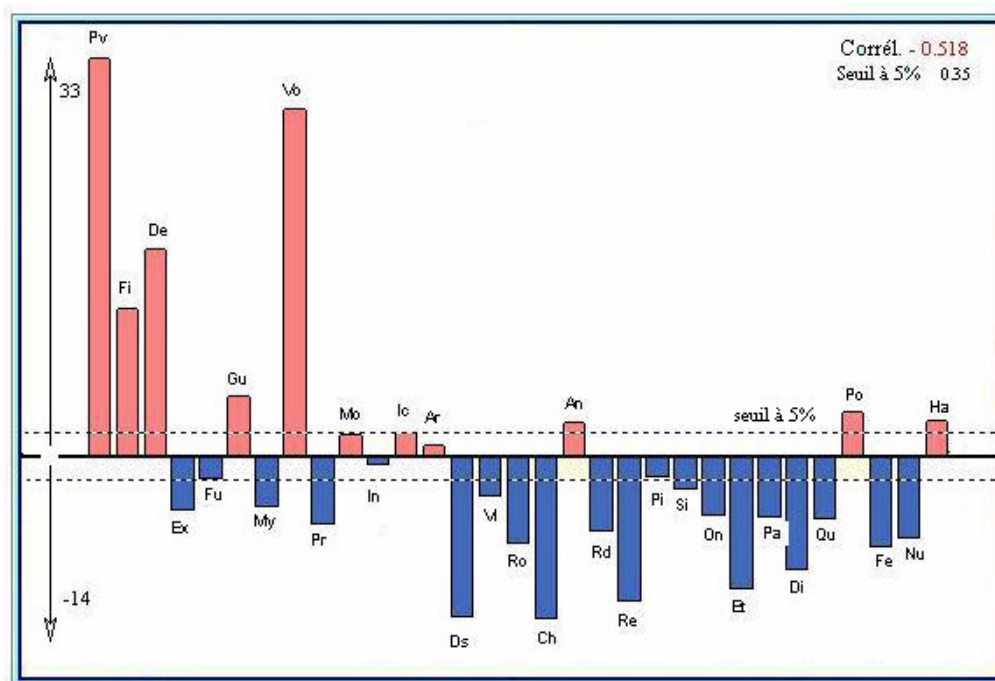


Figure n°112 : La distribution des pronoms personnels à la deuxième personne dans le corpus B (écarts réduits).

Il est aisé de constater que Le Clézio fait l'usage de la deuxième personne surtout au début de sa carrière littéraire, dans les premiers ouvrages ; *Le procès-verbal*, *La fièvre*, *Le déluge* ainsi que *Voyages de l'autre côté* en sont des témoins significatifs. Par la suite, nous pouvons observer un déclin de son usage à travers le corpus : le coefficient de corrélation chronologique atteint ici un taux significatif.

Les valeurs d'usage des pronoms de deuxième personne peuvent naturellement être assez variées : le pronom *vous*, ambigu, peut être une forme de politesse ou un vrai pluriel ; le pronom *tu* peut apparaître dans les dialogues entre personnages, comme dans *Poisson d'or*, ou lorsque l'auteur s'adresse fictivement au lecteur, comme dans *Voyages de l'autre côté* où l'auteur mène une sorte de dialogue avec le lecteur et l'"implique" ainsi dans le récit³⁵⁷ :

³⁵⁷ *Voyages de l'autre côté*, p. 99.

“Naja Naja voyage dans les arbres. Ça, ce n’est pas facile. Voici comment il faut faire : tu sors de la ville, et tu montes tous ces escaliers qui conduisent vers le haut des collines bleues. Tu montes les marches des escaliers sans te presser. C’est même préférable de monter lentement, pour garder ton souffle. Il y a de moins en moins de maisons, de plus en plus de jardins. De temps à autre, tu te retournes et tu vois la ville qui s’étend au pied des collines bleues.”

Néanmoins le regroupement des formes au singulier au début de l’œuvre manifeste aussi une part d’homogénéité dans leur emploi. Il convient donc de rappeler que cette catégorie du pronom personnel à la deuxième personne, tout en étant minoritaire dans l’ensemble, est quand même significative d’un style et une technique d’expression que Le Clézio semble abandonner au fur et à mesure que l’œuvre progresse.

C’est effectivement la troisième personne qui domine la catégorie des pronoms personnels avec ses 120.432 effectifs, englobant 73,1% des occurrences. Le lecteur de Le Clézio sait déjà qu’il favorise globalement le récit traditionnel à la troisième personne. Cependant, l’histogramme de la figure n°113 montre que la distribution interne n’est pas régulière à travers le corpus :

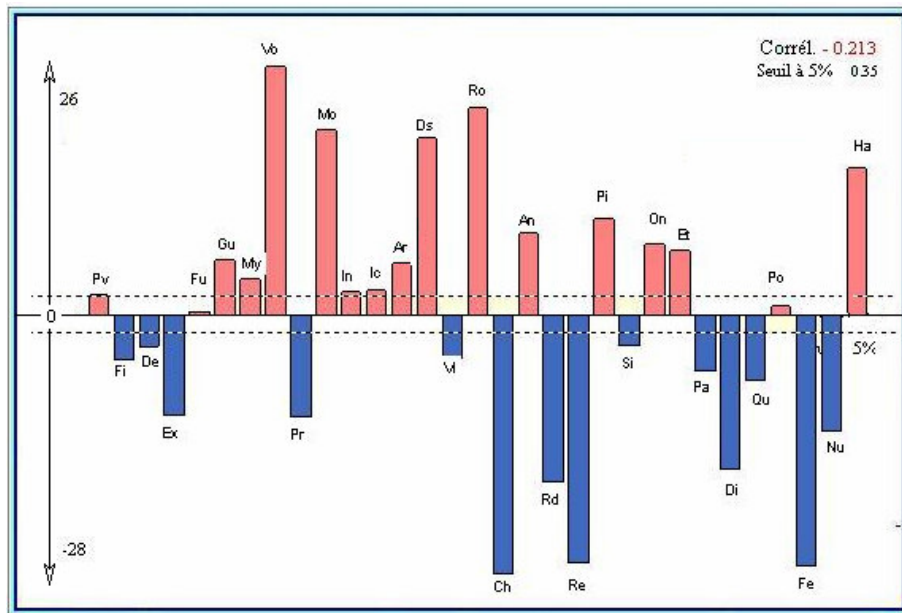


Figure n°113 : La distribution des pronoms personnels à la troisième personne dans le corpus B (écarts réduits).

Les pronoms personnels à la troisième personne sont, soit légèrement excédentaires, soit déficitaires dans les premiers ouvrages inspirés par l'école du "nouveau roman" qui privilégie des formes moins conventionnelles que le récit traditionnel à la troisième personne. L'histoire de Naja Naja dans *Voyages de l'autre côté* exploite en revanche de façon significative la troisième personne. *Mondo*, qui marque le tournant dans l'écriture romanesque de Le Clézio, est aussi largement excédentaire. Les romans qui suivent – à l'exception du roman *Le chercheur d'or*, de *Voyage à Rodrigues* et de *La quarantaine* - sont également en excédent. Nous avons vu que la réticence à l'égard de la première personne s'estompe à la fin de son œuvre ; corollairement, la troisième personne perd alors de son importance.

En revanche, Le Clézio exploite moins la troisième personne dans les genres non romanesques, comme l'on pourrait s'y l'attendre, où nous trouvons des valeurs déficitaires parfois importantes. L'histogramme ci-dessous, s'appuyant sur le corpus E, rend mieux compte de l'évolution de l'utilisation de la troisième personne dans l'écriture uniquement romanesque :

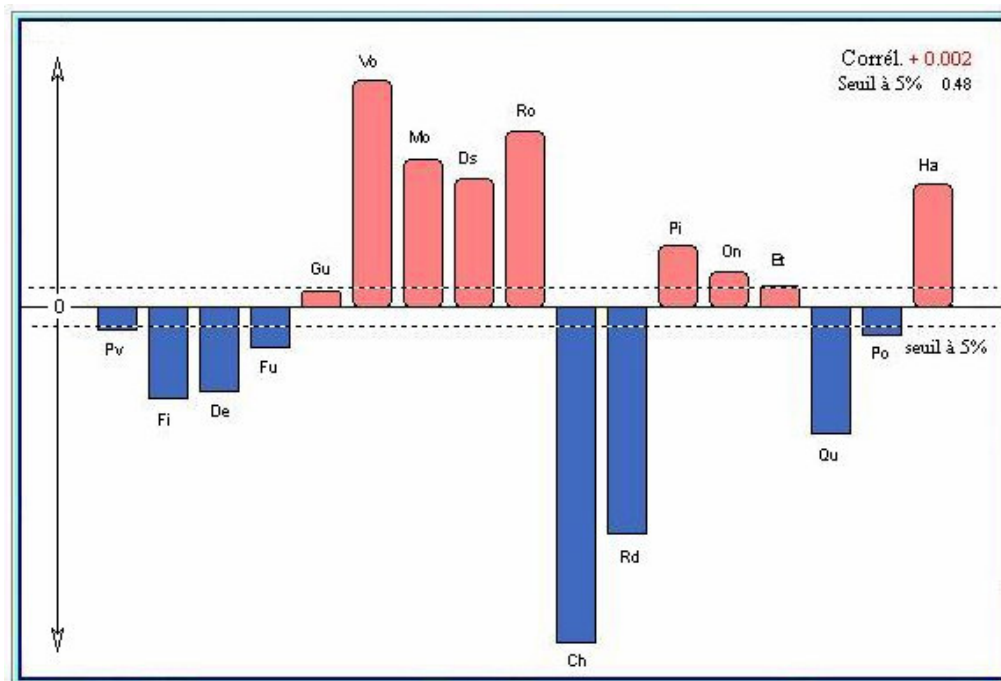


Figure n°114 : La distribution des pronoms personnels à la troisième personne dans le corpus E (écarts réduits).

La figure permet en effet de distinguer plusieurs périodes dans l'œuvre : une première déficitaire, un changement après *Le livre des fuites* vers un emploi significatif de la troisième personne jusqu'à *Le chercheur d'or*, où l'emploi de la première personne dominait nettement³⁵⁸. Par la suite, notons un nouvel emploi significatif, suivi des valeurs déficitaires dans *La quarantaine* et dans *Poisson d'or* mais qui se termine par une hausse dans le dernier livre du corpus.

Dans *Le livre des fuites* Le Clézio nous livre la première rupture de style. Après avoir déclaré au début du livre la phrase³⁵⁹ : “Malhonnêteté désespérante de celui qui n'ose pas dire ‘je’”, Le Clézio passe progressivement dans les livres suivants à un récit à la troisième personne qui se rapproche parfois du conte. Jean Onimus est persuadé que l'œuvre de J.M.G. Le Clézio est avant tout celle d'un conteur³⁶⁰ :

“Dire que Le Clézio est plutôt un conteur qu'un romancier peut paraître un paradoxe quand on pense au caractère typiquement “baroque” de ses livres, avec leur intense, leur exubérante vitalité, la variété des tons et de styles, la tragique obsession de la mort, le dédain pour les stylisations classiques. Pourtant n'est-il pas l'auteur de contes pour enfants ?”

Ce récit à la troisième personne semble correspondre aussi au personnage de Le Clézio qui, connu par sa modestie, a souvent évité le “je” au profit de l'écriture à la troisième personne. Jean-Xavier Ridon consacre une étude de cet effacement de soi dans son ouvrage³⁶¹ :

“Cet effacement de la voix propre à l'auteur se réalise constamment dans les textes de Le Clézio par le jeu des pronoms personnels mais aussi par l'utilisation des différents discours (direct, indirect, indirect

³⁵⁸ Cf. figure n°111, p. 350.

³⁵⁹ *Le livre des fuites*, p. 54.

³⁶⁰ J. Onimus (1994) : p. 183-184.

³⁶¹ J.-X. Ridon (1995) : p. 82.

libre) qui constituent la langue. On ne sait pas toujours qui parle ou qui s'écrit. Est-ce l'auteur ? est-ce les personnages ou bien nous-mêmes, lecteurs, auxquels on prêterait la parole ?”

Cependant, l'utilisation répétée du pronom personnel "on" (9158 occurrences) met en évidence une écriture qui s'élabore sans le pronom "je", qui, comme le souligne Jean-Xavier Ridon, "dépassé le cadre de cette instance énonciatrice tout en gardant une forme d'honnêteté envers le lecteur". L'histogramme ci-dessous montre la distribution relative de ce pronom à travers le corpus :

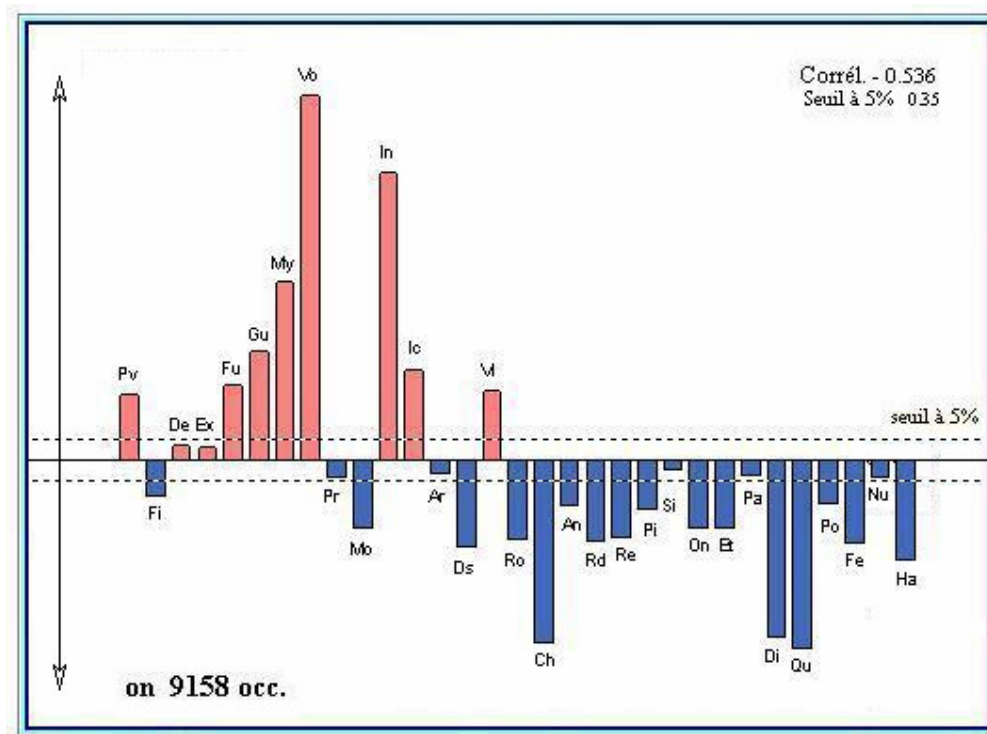


Figure n°115 : La distribution du pronom on dans le corpus B.

L'emploi du pronom *on*, très important dans la première partie de l'œuvre, est abandonné à la fin des années 70, au même moment que le récit à la troisième personne est introduit de façon significative dans l'œuvre. L'utilisation du pronom *on*, qui culmine dans *Voyages de l'autre côté*, semble, vraiment caractéristique de la première période de l'écriture leclézienne, influencée par l'école du "nouveau roman".

Le pronom “on” permet d’allier toutes les voix (celle de l’auteur, celle des personnages ou bien celle des lecteurs) car en même temps il peut représenter le narrateur, son personnage, l’écrivain et le lecteur. Il permet en effet - dans un champ identique - l’élaboration d’un jeu de voix différentes, technique caractéristique de l’école du “nouveau roman”. “Mais “on” est aussi, comme le souligne de manière juste Jean-Xavier Ridon³⁶², un principe de différence et de multiplicité puisque la voix qui parle par son intermédiaire est rendue générale ou parfois même multiple, ce qui donne au récit un caractère de doute”. L’extrait suivant de *La guerre*³⁶³ illustre bien ce phénomène :

“Peut-être qu’on ne verrait jamais plus le jour, peut-être que c’était ça qui était écrit sur le petit ticket de carton jaune percé d’un trou. Ou bien on s’abriterait là, au fond de la terre, pendant des mois, et on entendrait très loin les rumeurs et les grognements de la guerre qui faisait rage à la surface [...] Et on attendrait là, dans les souterrains, on se coucherait par terre sur le sol plein de parcelles de mica, et on espérerait pendant des mois, des années, des siècles.”

En effet, est-il ici possible de savoir qui est l’énonciateur ? Est-ce l’auteur en tant que narrateur du livre ? Ou bien est-ce le personnage qui pense ? La première hypothèse offre un excellent exemple de la technique de Le Clézio qui cherche à s’identifier, voire à se confondre, avec son personnage, ce qui est caractéristique de la première période de l’œuvre leclézienne. Le pronom “on” inclut aussi le lecteur dans un mouvement, ou une pensée. Le lecteur est ainsi incité à prendre part à cette opération de lecture-interprétation, fonctionnant selon le même principe de dérive que l’on voit chez les auteurs appartenant à l’école du “nouveau roman”.

Le recensement systématique d’occurrences pronominales permet ici de cerner ces techniques d’expression particulières qui souvent laissent chez le lecteur de Le

³⁶². J.-X. Ridon, *Henri Michaux: J.M.G. Le Clézio, L’exil des mots, op. cit.*, p. 83.

³⁶³. *La guerre*, p. 280.

Clézio une sensation de doute ou de mystère, sans pour autant pouvoir tenir compte de toute l'ambiguïté du pronom on, bien documentée par ailleurs.

L'étude des pronoms permet également d'étudier un autre facteur de style, celui du niveau de langue, mesurable grâce à l'indice pronominal.

4.6.3. L'indice pronominal.

En 1962 Charles Muller a mis au point un "indice du style familier"³⁶⁴. Il avait trouvé que le quotient entre les première et deuxième personnes des pronoms personnels et les mêmes personnes des possessifs dans les pièces de théâtre qu'il étudiait pouvait donner une mesure de la valeur stylistique du texte, indiquant son appartenance à un langage familier ou à une langue plus soutenue. "D'une façon générale, écrit-il³⁶⁵, l'indice pronominal est élevé quand le style est familier, bas quand le style tend à la noblesse ou au lyrisme." Bien que la pertinence de cet indice s'explique par des alternances typiques du théâtre classique ("je suis déshonorée" *vs.* "Mon honneur est perdu"), son emploi sur d'autres types de corpus s'est avéré également pertinent³⁶⁶.

Charles Muller avait observé des valeurs variant pour la plupart entre 1,6 et 3 dans le théâtre classique ; par exemple l'indice moyen pour *Le Cid* était de 1,64 et celui de *Nicomède* était encore en-dessous (1,07). Étienne Brunet a en revanche relevé des valeurs de 5,72 chez Zola, 4,08 chez Giraudoux, 3,56 chez Hugo et 2,63 chez Chateaubriand, dont l'indice témoigne d'un style très soutenu. Chez Proust il a observé un indice de 3,30 dans son premier texte qui passait à 5,03 dans son dernier. Dans *Frantext* l'indice pronominal est de 4,33 pour la prose littéraire et de 2,004 pour la poésie. Qu'en est-il du corpus leclézien ? Notre corpus B affiche un

³⁶⁴ Ch. Muller, "Sur quelques scènes de Molière, essai d'un indice du style familier", in *Langue française et linguistiques quantitatives*, Genève, Slatkine, 1979, p. 107-124.

³⁶⁵ Ch. Muller (1979) cité par E. Brunet (1988) p. 205.

³⁶⁶ Cf. les études d'É. Brunet. Cependant, la transposition de cet indice à d'autres genres littéraires nous paraît devoir être faite avec beaucoup de prudence et, pour notre part, nous contrôlerons les conclusions auxquelles conduira son application au corpus qui nous occupe.

indice moyen 5,81 qui indique un style relativement familier. Le tableau n°35 rend compte de l'indice pronominal dans les différents ouvrages du corpus :

	PERSONNELS			POSSESSIFS			Indice pronominal	
	1e pers.	2e pers.	Total	1e pers.	2e pers.	Total	Indice	
Pv	1469	950	2419	135	103	238	10,16	
Fi	1415	611	2026	162	52	214	9,47	
De	1293	833	2126	154	86	240	8,86	
Ex	2204	225	2429	610	54	664	3,66	
Fu	1351	339	1690	198	40	238	7,10	
Gu	1338	505	1843	122	83	205	8,99	
My	4	10	14	0	6	6	2,33	
Vo	758	1032	1790	222	117	339	5,28	
Pr	23	1	24	5	1	6	4,00	
Mo	412	399	811	59	59	118	6,87	
In	1527	431	1958	267	91	358	5,47	
Ic	29	16	45	5	0	5	9,00	
Ar	4	15	19	1	0	1	19,00	
Ds	438	256	694	208	47	256	2,71	
Vi	53	32	85	32	9	41	2,07	
Ro	571	181	752	72	10	82	9,17	
Ch	5867	175	6042	1047	23	1070	5,65	
An	109	116	225	15	11	26	8,65	
Rd	489	17	506	269	4	273	1,85	
Re	240	102	342	83	33	116	2,95	
Pi	2666	239	2905	437	30	467	6,22	
Si	27	1	28	9	0	9	3,11	
On	239	199	438	34	44	78	5,62	
Et	2301	216	2517	525	19	544	4,63	
Pa	310	7	317	31	1	32	9,91	
Di	369	112	481	113	24	137	3,51	
Qu	4827	483	5310	816	58	874	6,08	
Po	4439	393	4832	620	36	656	7,37	
Fe	418	140	558	106	48	154	3,62	
Nu	165	14	179	35	1	36	4,97	
Ha	245	304	549	37	41	78	7,04	
Tot.	35600	8354	43954	6429	1131	7561	5,81	

Tableau n°35 : L'indice pronominal.

Le tableau permet encore une fois d'observer l'importance du genre littéraire. Les essais, les ouvrages ethnologiques et la biographie ont un indice très inférieur aux œuvres romanesques. *Mydriase*, le récit poétique, a une valeur très basse, comme l'essai *Trois villes saintes* où l'auteur insère, dans le récit très poétique, des extraits de

textes sacrés maya ainsi que les mots des conquérants et des prêtres, comme ici Juan de la Cruz³⁶⁷ :

“Ainsi donc, mes chers Chrétiens, fils des villages, gardez dans votre cœur mes commandements, parce que moi-même, mes enfants, je ne puis rester immobile, sans cesse je suis en route, ma gorge et mon ventre sont desséchés par une soif inextinguible, car sans cesse je suis en marche à travers le Yucatan pour vous défendre ...”

Les autres ouvrages qui traitent du monde amérindien, par exemple *Le rêve mexicain* ou *La fête chantée*, sont aussi souvent riches en extraits de ce type, qui contribuent à donner une valeur très poétique, presque lyrique à ces textes bien qu'ils soient classés comme étant des ouvrages ethnologiques.

Notons que le seul livre pour enfants du corpus, *Voyage au pays des arbres*, a un indice pronominal de 19,00 qui reflète bien son style particulier. Les variations de l'indice pronominal à l'intérieur de l'œuvre romanesque sont d'une très grande ampleur. La courbe ci-dessous en témoigne :

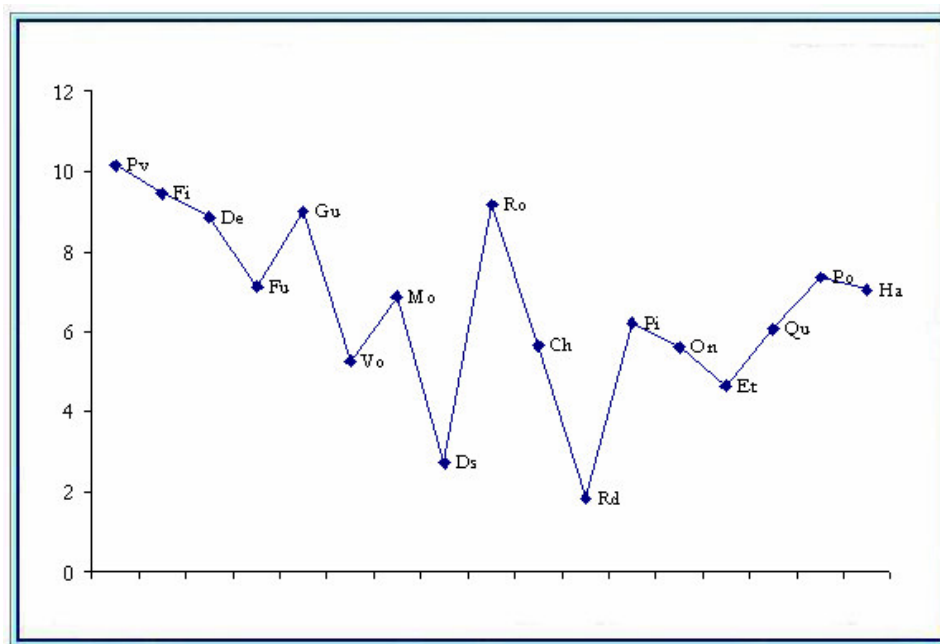


Figure n°116 : Courbe de l'indice pronominal du corpus romanesque (le corpus E).

³⁶⁷ *Trois villes saintes*, p. 55.

Malgré quelques saccades de la courbe, nous pouvons observer un mouvement chronologique qui après des valeurs initiales élevées tend vers une baisse au milieu de l'œuvre, c'est-à-dire vers un style plus soutenu. Le Clézio semble laisser de côté le style très familier de ses débuts, puis – après cette période centrale au ton soutenu – revenir vers la fin de son œuvre à un style moins soutenu, sans pour autant atteindre les niveaux initiaux.

La courbe permet en effet d'observer un indice très élevé au début de l'œuvre dans les ouvrages "nouveau roman". C'est le premier livre du corpus *Le procès-verbal*, mais également *La guerre*, *La ronde* et *Poisson d'or*, qui ont les indices les plus élevés du corpus. C'était dans ces livres – à l'exception de *Poisson d'or* – que l'usage de la deuxième personne était excédentaire. Ils ont en commun des personnages adolescents qui s'expriment dans une langue de jeunes teintée de langue orale. L'extrait ci-dessous témoigne de ce style³⁶⁸ :

“Puis d'un seul coup cela s'en va, et c'est elle maintenant qui parle, la voix un peu rauque, sans savoir très bien ce qu'elle dit. “Bon. Alors, on y va ? On y va maintenant ?”

Le garçon descend de son vélomoteur. Il embrasse Titi sur la bouche, puis il s'approche de Martine qui le repousse avec violence.

“Allez, laisse-la.”

Les valeurs les plus basses de l'indice, témoignant d'un style soutenu ou lyrique (selon la qualification proposée par Ch. Muller), se trouvent dans *Désert* et dans *Voyages à Rodrigues* mais pour des raisons diverses. L'indice de 2,71 est remarquable, bien qu'il n'étonnerait guère le lecteur leclézien connaissant la veine poétique du roman *Désert* pour lequel Le Clézio a reçu de nombreuses récompenses. Quant à *Voyages à Rodrigues*, les raisons de l'indice très faible sont non seulement à chercher dans le style très particulier, déjà décrit auparavant, de

³⁶⁸ *La ronde et autres faits divers*, p. 14.

ce livre, mais également dans le fait qu'il s'agit du récit à la première personne de l'aventure du grand-père de Le Clézio qu'il nomme "mon grand-père" tout au long du livre³⁶⁹ :

"La présence de mon grand-père dans ce lieu solitaire, c'est cela qui me trouble, me retient. C'est mon unique lien avec lui, car je ne sais rien de lui, hormis ces papiers et quelques photos jaunies."

L'indice pronominal semble avoir un rapport avec la proportion de substantifs et de verbes dans notre corpus : lorsqu'un texte est riche en substantifs il a un indice pronominal bas et vice versa. Il s'agit, pour les valeurs basses, aussi souvent d'un effet du genre littéraire que d'un signe de noblesse ou de lyrisme. La fréquence totale des pronoms d'un texte, qui va de pair avec celles de verbes, se reflète également sur les résultats de l'indice. Les textes riches en verbes, pleins d'action, sont aussi souvent riches en dialogues et en oralité produisant un indice pronominal élevé. En revanche, les possessifs (du moins les adjectifs qui sont très largement majoritaires dans cette catégorie) accompagnent nécessairement un substantif. L'exemple ci-dessous tiré de *L'inconnu sur la terre* illustre bien cette conjonction, cette affinité entre les diverses catégories morpho-syntaxiques³⁷⁰ :

"Et la parole venue de l'espace, qui va vers l'espace, me traverse. Elle me donne mon vrai nom, ma vraie pensée, mon unique langage. [...] et mon corps et mon esprit sont une mécanique dont tous les rouages s'entraînent, et j'avance, je respire, je vis."

Finalement, il s'est avéré que l'indice pronominal auquel nous croyions très peu après avoir pris connaissance des critiques émises sur son efficacité – sur d'autres genres que le théâtre classique – fonctionne parfaitement sur le corpus Le Clézio et fournit un complément intéressant à l'analyse stylistique.

³⁶⁹ *Voyage à Rodrigues*, p. 38.

³⁷⁰ *L'inconnu sur la terre*, p. 138.

Dans ce chapitre nous avons pu constater que la fréquence des pronoms, qui occupent environ 10% de notre corpus, est très variable. A l'intérieur de la catégorie des pronoms personnels nous avons pu constater la dominance de la 3^e personne qui occupe presque trois quarts des personnels. Le Clézio privilégie cette forme d'écriture, et l'usage de la 1^e et de la 2^e personnes est beaucoup plus restreint et bien limité à certains livres du corpus. C'est dans ces livres que nous avons relevé un indice pronominal élevé, témoignant d'un style encore plus familier que celui du corpus dans son ensemble.

Les pronoms étaient, nous l'avons vu dans les analyses factorielles de la distribution des parties du discours³⁷¹, étroitement liés à une autre catégorie grammaticale, celle des verbes que nous allons étudier dans le chapitre suivant.

³⁷¹ Cf. figure n°79, p. 299.

4.7. La catégorie verbale.

“Il n’y a pas plusieurs formes de conscience, mais une seule. Jamais le langage ne pourra restituer cette évidence, puisque le langage est une interprétation, une aliénation, une action. Ce qu’il nous faut deviner est caché derrière les mots et les actes, caché derrière les œuvres. Il nous faut revenir au point de départ pour connaître le lieu de l’arrivée. Entre la réalité brute, et la reconquête de la réalité pure, il y a tout le voyage du verbe.”

L’extase matérielle, page 173.

“Mon intention est de plaider ici, écrit Charles Muller³⁷², pour une indexation grammaticale, et plus précisément pour l’identification des formes verbales dans les catégories traditionnelles de la personne, du temps et du mode. C’est que je pose en principe que l’analyse stylistique d’un texte gagnerait à ne pas borner ses données quantitatives au seul lexique, et que le maniement par l’écrivain des catégories verbales mérite une attention d’autant plus grande qu’il constitue certainement un élément de grande mobilité, et qui fournit des indications précises sur la composition de l’œuvre, sur sa structure stylistique et thématique. [...] On souhaiterait donc des index où la catégorisation des occurrences se double, au moins pour les verbes, d’une sous-catégorisation grammaticale, qui permettrait à l’usager de connaître non seulement la fréquence de chaque verbe, mais celle des ensembles grammaticaux constitués par la réunion de toutes formes d’un même temps, d’un même mode, d’une même personne.”

La difficulté de l’analyse verbale est bien documentée ; les ambiguïtés lexicales aussi bien que celles qui relèvent de la grammaire posent de nombreux problèmes.

³⁷² Ch. Muller, “Lexicologie, statistique lexicale et critique littéraire”, in *Langue française, débats et bilan, recueil d’articles*, Honoré Champion, Paris, 1993, p. 62.

En effet, la catégorie verbale comporte de nombreuses homographies. Le bilan des homographies internes du verbe français qu'a effectué Charles Muller comptabilise 51 formes simples (participes compris) générées par la "conjugaison" d'un verbe français. Ces 51 formes simples comprennent : 7 "paradigmes" de 6 formes ou "personnes" chacun : indicatif présent, imparfait, futur, passé simple, conditionnel présent ; subjonctif présent ; un paradigme de 3 formes, l'impératif ; un ensemble de 4 formes, le participe passé ; deux formes uniques, l'infinitif et le participe présent. De ces formes, il n'y en a que 18 qui ne sont jamais ambiguës dans un nombre important de verbes. Si l'on tient compte de la fréquence des formes graphiques ambiguës, on peut donc les estimer à 20% des formes verbales, soit 5% environ des mots du texte³⁷³.

La forme graphique *courait* par exemple, hors contexte, nous suffit pour reconnaître une des occurrences du verbe *courir*, et un élément de l'ensemble de l'indicatif imparfait et de la 3^e personne du singulier. La forme *court* est ambiguë d'un point de vue lexical (*courir* verbe, ou bien *court* adj.), mais rattachée au verbe, elle n'offre pas d'ambiguïté grammaticale ; elle ne peut être qu'un indicatif présent, une 3^e personne du singulier. En revanche, la forme *courus* est claire lexicalement, mais peut être soit un participe, soit un passé simple, et dans ce dernier cas, une 1^{ère} ou une 2^e personne³⁷⁴.

Il y a donc deux préparations importantes à effectuer lors du traitement des formes verbales pour que la recherche statistique soit fiable et valable : d'abord la lemmatisation, qui a pour tâche de lever les ambiguïtés lexicales, ensuite l'étiquetage morpho-syntaxique pour lever les ambiguïtés grammaticales.

³⁷³ Ch. Muller fait aussi une remarque sur ces fréquences : "la proportion en formes ambiguës est la plus faible dans les verbes les plus fréquents (les auxiliaires, les semi-auxiliaires, les verbes irréguliers du 3^e groupe) ; elle est la plus forte dans les conjugaisons régulières (1^{er} et 2^e groupe)." Cf. Ch. Muller, "Lexicologie, statistique lexicale et critique littéraire", in *Langue française, débats et bilan, recueil d'articles*, Honoré Champion, Paris, 1993, p. 63.

³⁷⁴ Cf. chapitre 1.5. "La lemmatisation".

“Souhaitons, écrit Charles Muller³⁷⁵, qu’avec l’aide des machines on nous donne des index aussi riches d’information lexicale et grammaticale que possible, qui ouvriraient des perspectives prometteuses à l’analyse quantitative des œuvres littéraires.”

Son souhait est aujourd’hui en train de se réaliser grâce au progrès de l’informatique. Le lemmatiseur et l’analyseur syntaxique intégrés dans certains logiciels permettent désormais de trier l’ensemble des formes verbales de façon “automatique”.

Avec l’aide du logiciel Hyperbase nous avons pu effectuer une analyse de l’ensemble des verbes dans le corpus Le Clézio. Etant donné qu’une étude exhaustive de l’ensemble des temps, des modes et des aspects serait extrêmement longue et fastidieuse, nous nous intéresserons, après avoir situé l’ensemble de la catégorie dans le corpus, à certaines des sous-catégories qui nous ont semblé pertinentes et révélatrices de l’écriture leclézienne.

4.7.1. La distribution des verbes.

L’analyse factorielle des parties du discours nous avait déjà indiqué que la distribution des 320.747 occurrences verbales n’était pas stable à travers le corpus. L’histogramme ci-dessous permet de mieux évaluer ses variations et son évolution :

³⁷⁵ Ch. Muller, “Lexicologie, statistique lexicale et critique littéraire”, in *Langue française, débats et bilan, recueil d’articles*, Honoré Champion, Paris, 1993, p. 65.

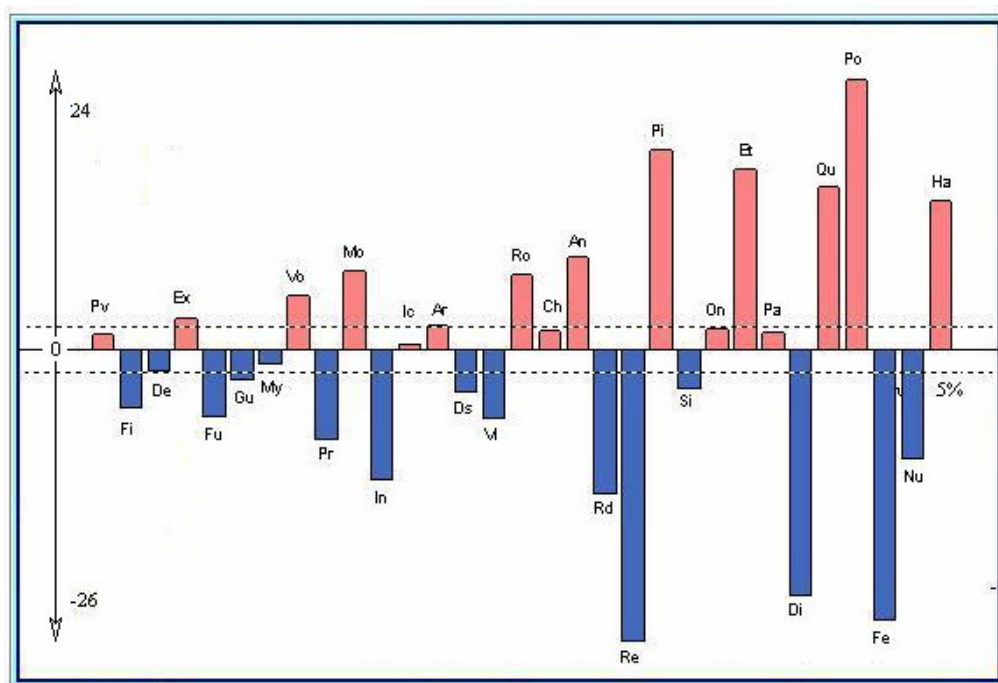


Figure n°117 : La distribution de la catégorie de verbes dans le corpus B (écarts réduits).

Nous observons premièrement le mouvement chronologique des verbes. Les écarts autour de l'axe des abscisses s'amplifient dans les deux sens, celui des excédents et celui des déficits, au fur et à mesure que l'œuvre progresse. Cette amplification se modère toutefois vers la fin du corpus, et les écarts diminuent dans les deux derniers livres du corpus.

Dans le chapitre précédent nous avons observé que le substantif aimait la compagnie de l'adjectif. En revanche, il n'est pas du tout attiré par le verbe qu'il a plutôt tendance à repousser. Les histogrammes des deux catégories sont pratiquement inversés³⁷⁶. Il y a toutefois quelques exceptions de cette tendance générale dans les premiers romans de la première période "nouveau roman" : *La fièvre* et *Le déluge*, qui sont déficitaires dans les deux catégories. Dans ces deux livres où l'auteur expérimente des procédés pour mettre à jour les sensations, Le Clézio se moque des conventions littéraires et entend le manifester.

³⁷⁶ Cf. figure 85, p. 310.

“La littérature, écrit-il à cette époque³⁷⁷, n’est-elle pas en ce cas (et particulièrement la fiction) l’effort désespéré et permanentement mis en échec pour produire une expression unique ?”

Étienne Brunet avait déjà repéré³⁷⁸, dans *Frantext*, le progrès des verbes au détriment des substantifs sur la période 1789-1960³⁷⁹. Dans notre corpus les verbes progressent dans les œuvres de fiction - romans et recueils de nouvelles, mais en revanche ils reculent dans les essais et les ouvrages d’ethnologie. *Le rêve mexicain*, *La fête chantée* et *Diego et Frida*, qui comptaient des valeurs excédentaires importantes en substantifs, sont fortement déficitaires en verbes.

Les textes qui favorisent les substantifs sont souvent caractérisés comme “descriptifs”, “pittoresques” et “substantiels” tandis que ceux qui favorisent les verbes sont qualifiés de “pleins d’actions”. Dans notre corpus “l’action” ne semble pas être située dans les ouvrages que nous venons de citer, mais plutôt dans les romans.

Écartons donc pour un instant les ouvrages d’ethnologie et les essais, pour regarder de plus près l’œuvre romanesque. L’histogramme de la distribution de verbes dans le corpus Roman (E) permet d’affiner l’analyse des variations constatées auparavant :

³⁷⁷ *Le livre des fuites*, p. 191.

³⁷⁸ É. Brunet (1981) : p. 203.

³⁷⁹ Cela n’est pas vrai pour une période plus longue, car si on remonte jusqu’au XVI^{ème} siècle le substantif augmente.

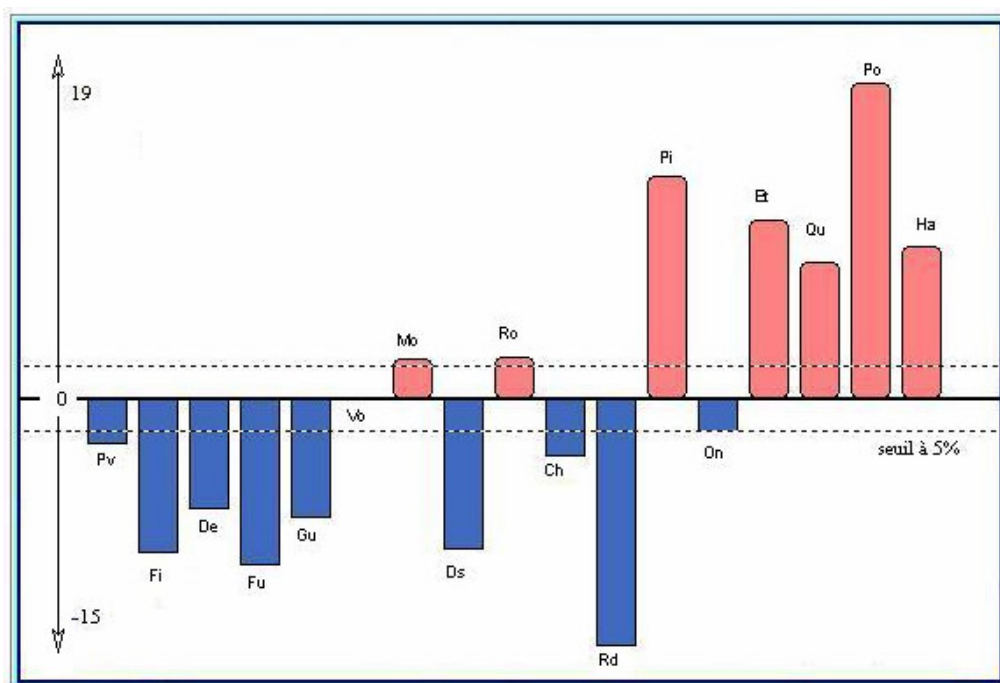


Figure n°118 : La distribution relative de verbes dans le corpus romanesque (le corpus E).

L'histogramme divise cette fois-ci plus clairement l'œuvre leclézienne en plusieurs périodes. Les verbes sont déficitaires dans toute la première période, ce n'est qu'à partir de *Mondo* que nous pouvons observer des valeurs excédentaires. Cette variation correspond bien au changement du style dans l'écriture leclézienne et à la rupture dans l'œuvre dont les différentes études statistiques témoignent. Par la suite, durant la deuxième période, *Désert*, *Le chercheur d'or*, *Voyage à Rodrigues* et *Onitsha* – les romans riches en substantifs – sont déficitaires en verbes et vice versa. Dans le dernier tiers du corpus nous constatons l'augmentation progressive des verbes, correspondant au retour vers le roman traditionnel, et une écriture plus conventionnelle. Ces livres semblent privilégier l'action au détriment du substantif, plus "contemplatif". L'opposition la plus marquée est à observer dans *Poisson d'or*, le livre le plus riche de verbes (écart réduit de +19), dans lequel la jeune Laïla affronte la vie, un roman qui abonde véritablement en action³⁸⁰.

"Quand je sentais qu'il était bien fatigué, je m'en allais. J'embrassais sa main, et lui appuyait un instant sa paume sur mon front, descendait sur

³⁸⁰ *Poisson d'or*, p. 166.

mes yeux, sur mon nez, mes joues, mes lèvres. Il disait : “Au revoir, ma fille”, comme si j’étais vraiment Marima.

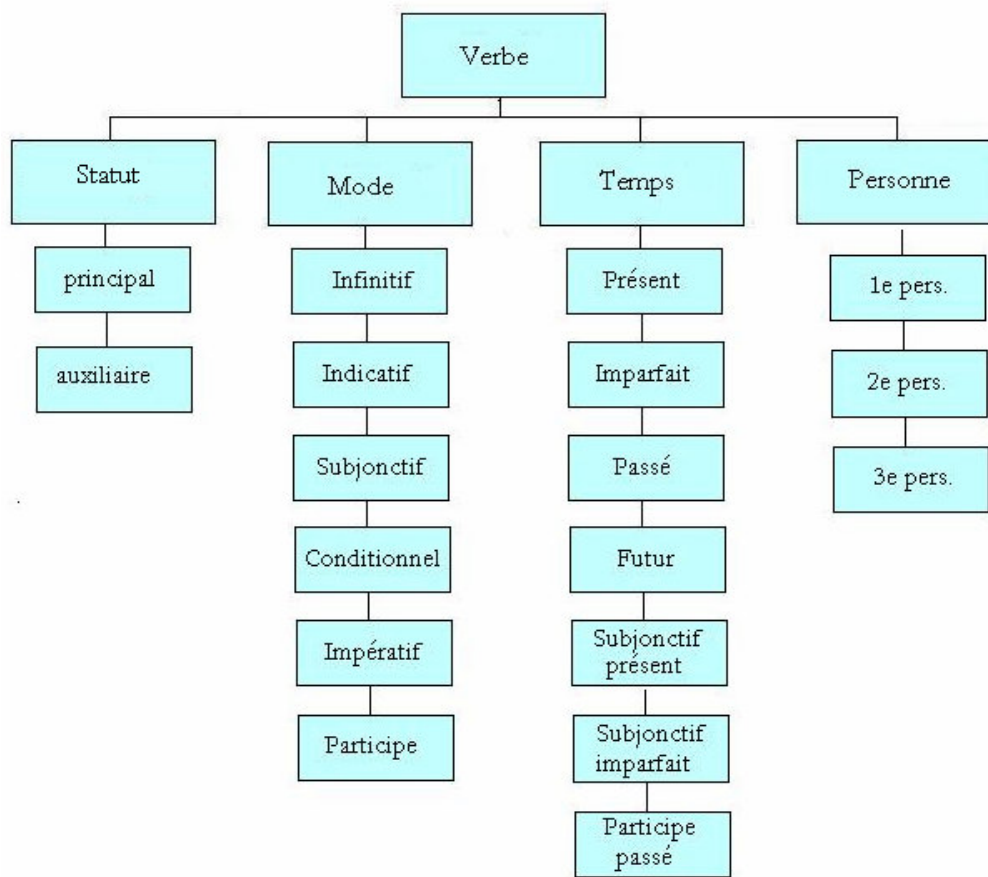
Peut-être qu’il croyait vraiment que j’étais elle. Peut-être qu’il avait oublié. Peut-être que c’était moi qui étais devenue semblable à elle, à force de venir auprès de son grand-père, à force de l’écouter raconter ce qu’il avait vécu là-bas, au bord du fleuve.”

Toutefois, lorsque l’on s’intéresse à la fréquence et à la distribution de verbes dans un corpus, il convient de prendre en considération la distinction des différentes catégories de verbes.

4.7.2. Les sous-catégories verbales.

Le logiciel Hyperbase permet désormais de distinguer et de regrouper les sous-catégories de verbes de façon automatique³⁸¹. L’analyse regroupe les verbes selon leur statut de principal ou d’auxiliaire, selon le mode, selon le temps exprimé ou bien selon la personne, de la façon suivante :

³⁸¹ Cf. chapitre 1.5.6. “Cordial”.



Il convient encore de souligner la difficulté des ambiguïtés et la relative opacité d'une analyse automatique qui ne permet pas à l'utilisateur de faire son choix et de trancher selon son propre jugement dans les cas ambigus³⁸². La distinction ne va pas de soi, notamment dans la première sous-catégorie. En effet, pour distinguer entre les emplois auxiliaires et les emplois pleins parmi toutes les formes du verbe *être*, peut-on se fier au tri automatique des temps composés ?

Lorsque l'on s'intéresse aux modes verbaux, peut-on se fier entièrement à la distinction du logiciel entre le participe et l'adjectif ? Ne risque-t-on pas d'inclure des adjectifs dans les participes passés et vice versa ? Les questions sont nombreuses et l'automatisme du traitement ne permet pas toujours d'avoir des réponses ni le contrôle nécessaire.

³⁸² Les règles sont opaques mais non le détail des analyses que l'on peut contrôler (sans les corriger).

Par conséquent, nous avons choisi d'exclure les études détaillées dans les domaines qui touchent de près les aspects évoqués ci-dessus, comme l'étude des auxiliaires et des verbes principaux.

Nous nous intéresserons en premier lieu à un caractère fondamental de la forme verbale : le mode.