3.5. La segmentation interne de la phrase.

> "Ces petits signes tarabiscotés qui avancent tout seuls, presque tout seuls, qui couvrent le papier blanc, qui gravent sur les surfaces planes, qui dessinent l'avancée de la pensée. Ils rognent. Ils ajustent. Ils caricaturent. Je les aime bien, ces armées de boucles et de pointillés. Quelque chose de moi vit en eux. Même s'ils n'ont pas de perfection, même s'ils ne communiquent pas vraiment, je les sens qui tirent vers moi la force de la réalité.
> Avec eux, tout se transforme en histoires, tout avance vers sa fin. Je ne sais pas quand ils arrêteront. Leurs contes sont vrais, ou faux. Ça m'est égal. Ce n'est pas pour ça que je les écoute. Ils me plaisent, et c'est avec plaisir que je me laisse tromper par le rythme de leur marche, que j'abandonne tout espoir de les comprendre un jour."
>
> *L'extase matérielle*, page 105.

Pour avoir une idée plus précise du rythme et des pauses tout au long d'un texte, il convient aussi de prendre en considération la ponctuation faible. L'étude de la distribution des signes de ponctuation faible – de la virgule, des deux points, des paires de parenthèses et du point-virgule[243] – apporte des connaissances sur la segmentation interne de la phrase. Il convient, toutefois, avant d'effectuer une étude de la segmentation du discours permettant de caractériser le style particulier de notre auteur, de rappeler qu'elle ne dépend pas uniquement de la syntaxe mais aussi des contraintes visuelles ou respiratoires. Les signes de ponctuation faible sont également très sensibles aux traditions typographiques et leur présence est parfois assurée par les soins de l'éditeur plutôt que par la plume de l'écrivain. Les critiques ont souvent commenté le rythme particulier de l'écriture leclézienne. Maurice Nadeau écrit ceci à ce propos[244] :

> "L'auteur du *Procès-verbal* n'est pas venu au roman à la suite de longues réflexions théoriques ou par une sorte de pari intellectuel aujourd'hui à la portée de tout adolescent qui a fait des études et appris à manier le langage. A lire ses phrases pressées, au rythme parfois haletant,

[243] Nous avons exclu de cette étude l'analyse le tiret à cause de son ambiguïté difficile à résoudre de façon satisfaisante dans un corpus numérisé de taille importante.
[244] Maurice Nadeau (1992) : p. 181-182.

télescopique, ses descriptions qui, ou tombent dans la pure énumération des objets, des faits, des situations (comme pour prendre pied plus vite dans le concret) ou s'élargissent en visions cosmiques, on sent qu'un besoin essentiel le pousse : la nécessité de dire et de dire vite et fort (parce que les hommes sont sourds, parce que le temps est court), et le désir quasi désespéré d'alerter ses contemporains sur ce qui, pour un individu délivré de ses entraves, de son rôle fonctionnel, de la vue peu exaltante qu'il prend de lui-même en tant que lieu de rencontre des forces biologiques, historiques et sociales, constitue sa réalité primordiale : d'être un vivant parmi les vivants et dans un ensemble cosmique qu'une seule grande loi résume : le passage de la naissance à la mort, sur des rythmes qui vont de la pulsation éphémère à la permanence apparente."

Pour analyser la phrase leclézienne, avec le recours impartial des chiffres, nous avons recensé les différents signes de ponctuation au moyen du logiciel Hyperbase. La distribution des fréquences absolues des signes de ponctuation faible dans notre corpus est la suivante :

Texte	,	;	:	(et)	Texte	,	;	:	(et)
Pv	7560	711	340	53	Ch	10984	17	374	38
Fi	9725	554	218	17	An	1845	0	53	6
De	10455	670	352	31	Rd	3260	10	146	162
Ex	8107	302	176	51	Re	6719	163	412	749
Fu	8866	135	275	53	Pi	5999	2	172	24
Gu	8616	161	297	87	Si	228	3	10	11
My	892	9	24	0	On	6274	7	173	4
Vo	9038	99	146	46	Et	10231	7	312	33
Pr	770	22	25	36	Pa	922	0	6	2
Mo	6864	73	170	11	Di	5443	44	272	229
In	12333	212	183	18	Qu	12551	4	401	129
Ic	289	6	6	12	Po	6711	14	284	21
Ar	189	4	6	4	Fe	5588	59	279	549
Ds	13085	175	170	11	Nu	1799	14	50	61
Vi	1485	17	10	0	Ha	5217	3	155	20
Ro	7384	35	137	11					

Total	,	;	:	(et)
201074	189429	3532	5634	2479

Tableau n°28 : Les signes de segmentation interne chez Le Clézio, effectifs absolus.

Il est aisé de constater le score écrasant de la virgule qui à elle seule compte 94,2% des effectifs. Les points-virgules occupent 1,8%, les deux-points 2,8% et les signes de parenthèse 1,2 % du nombre total des signes de ponctuation faible.

Nous nous proposons, avant d'en tirer des conclusions, de regarder de près chacun des quatre signes qui constituent les marqueurs de la segmentation interne de la phrase.

3.5.1. L'emploi de la virgule.

Avec 189.429 effectifs, la virgule est le signe de ponctuation le plus important (en quantité) du corpus. Son effectif est presque deux fois plus important que celui du point final, avec ses 97682 effectifs. L'histogramme ci-dessous rend compte de la distribution du signe à travers le corpus :

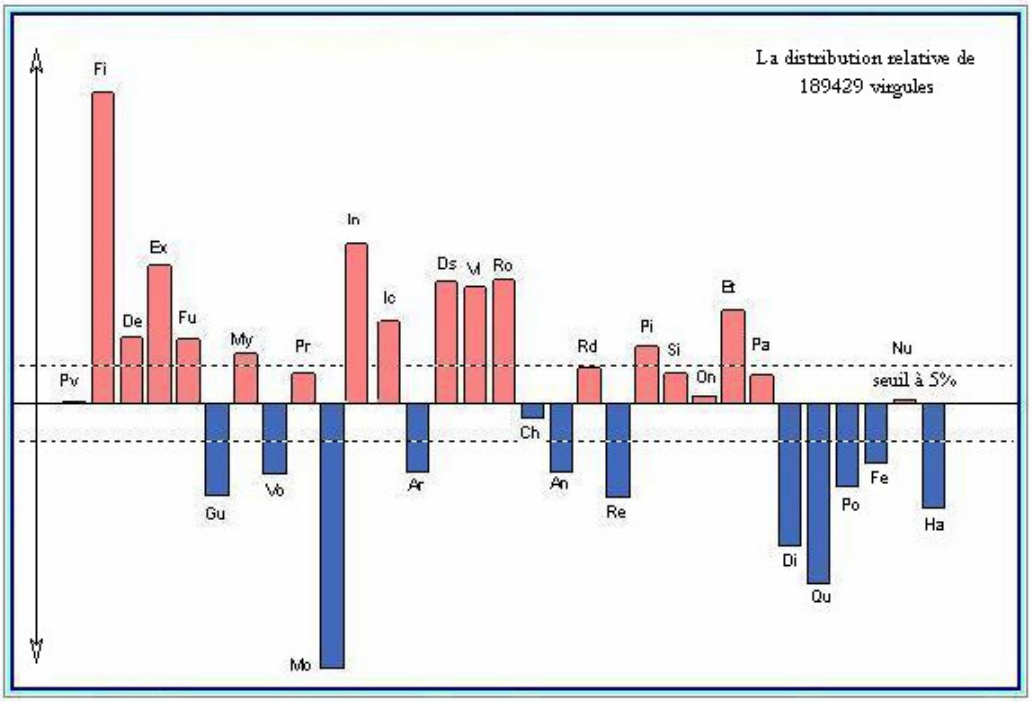


Figure n°58 : La distribution relative des virgules dans le corpus A.

La distribution de la virgule n'est pas constante, nous pouvons, en effet, observer des variations importantes. La virgule est excédentaire dans les premiers ouvrages avec l'emploi le plus fréquent du signe dans le deuxième livre du corpus : *La fièvre*. Dans ce livre contenant "neuf histoires de petite folie"[245] la virgule témoigne des courtes séquences qui contribuent à donner un rythme un peu vague au récit où les personnages glissent entre le conscient et l'inconscient. Le Clézio s'exprime ainsi à propos de cette écriture expérimentale[246] :

> "Je suis comme plongé dans une sphère étanche, je nage parmi les éléments de la pensée et de l'imagination. Ils reviennent toujours… ils sont cercle sans commencement, sans aboutissements… ils sont ivresse de roue, indéchiffrable mouvement de la vie sans fin qui me fait connaître l'éternité."

En effet, dans ce livre, les choses, les émotions, les visages, les dates, les titres de livres, les noms se combinent, s'enchaînent et se métamorphosent selon des rythmes imprévus au sein de longues phrases contenant de nombreuses virgules.

Dans *L'Inconnu sur la terre*, nous trouvons également une importante fréquence de la virgule (écart réduit…). Ce texte, difficilement classé, témoigne de la même volonté que dans *La fièvre* de créer un rythme particulier au récit. Il s'agit de créer l'ambiance de la rêverie et de la méditation autour du personnage principal ; l'enfant mythique[247] :

> "Je voudrais vous parler loin, longtemps, avec des mots qui ne seraient pas seulement des mots, mais qui conduiraient jusqu'au ciel, jusqu'à l'espace, jusqu'à la mer.
> J'entends ce langage, cette musique, ils ne sont pas étrangers, ils vibrent

[245] J.M.G. Le Clézio, dans sa présentation de l'ouvrage.
[246] *La fièvre*, p. 201.
[247] *L'inconnu sur la terre*, p. 7.

autour, ils brillent autour, sur les rochers blancs et sur la mer, ils brillent au centre des villes, même dans les yeux des passants.
Comment parler ? Les mots de cette musique viennent d'un pays où le langage n'existe pas, où le langage est scellé, enfermé en lui-même, est devenu comme la lumière, visible seulement de l'extérieur."

La virgule semble ici accompagner cette quête de l'expression juste, des mots qui se trouvent au-delà du langage : elle permet en effet la juxtaposition, l'accumulation des termes qui se succèdent et se remplacent parce qu'ils ne sont jamais totalement adéquats ; elle manifeste les retouches successives en quête d'une expression parfaite.

Dans le recueil de nouvelles *Mondo et autres histoires*, en revanche, la virgule se fait rare. C'est dans ce livre que nous avons constaté un changement important dans la structure lexicale[248] du corpus, correspondant à la rupture la plus forte dans l'œuvre. L'usage réduit de la virgule témoigne aussi de ce changement où Le Clézio abandonne la phrase compliquée comprenant plusieurs parties ou thèmes, dans sa recherche d'un langage simple et pur.

La fréquence faible de virgules dans *Mondo* témoigne également d'un autre phénomène bien connu dans la recherche lexicométrique : celui de la relation entre la virgule et le point. Les textes sont souvent pauvres en virgules lorsqu'ils sont riches en points et vice versa. En effet, la virgule a souvent une fonction de compensation et de complément à l'égard du point. Étienne Brunet a observé, dans son étude sur le vocabulaire de Hugo[249], l'image inversée de la virgule, en miroir, par rapport à celle du point.

Dans notre corpus ce n'est pas la règle générale bien que nous pouvons constater une opposition dans 18 des 31 ouvrages. La valeur du coefficient de corrélation

[248] Cf. les chapitres 2.4. "Le richesse du vocabulaire." et 2.5. "l'accroissement lexical.".
[249] É. Brunet (1988) : p. 81.

quasi nulle (0,030) indique l'indépendance mutuelle du point et de la virgule, comme l'illustre l'histogramme de la figure n°59 :

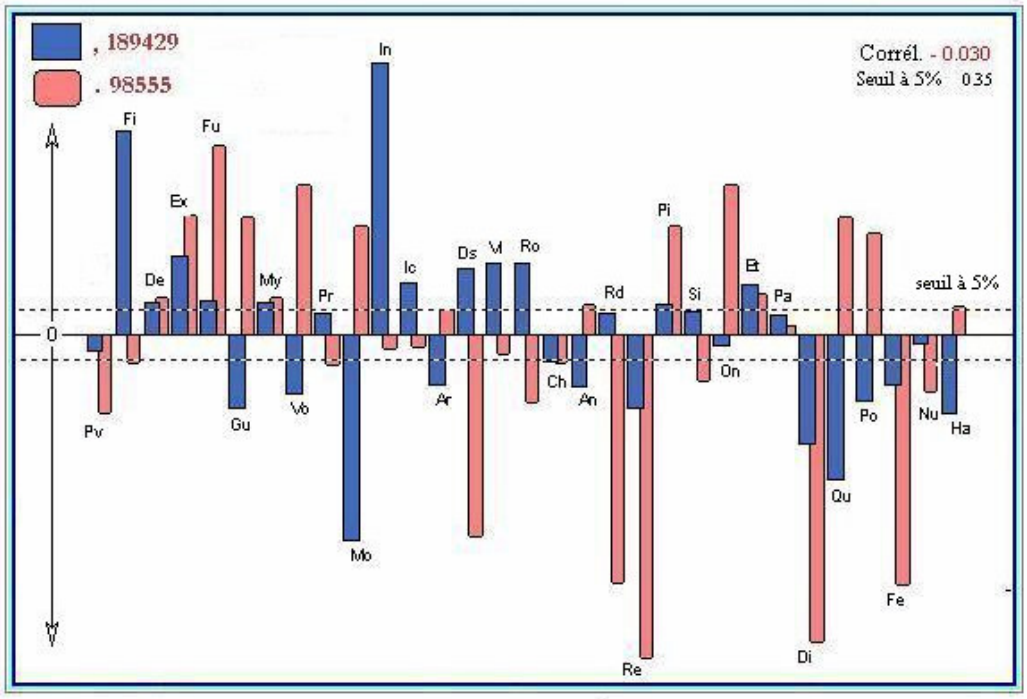


Figure n°59 : La corrélation entre points et virgules dans le corpus A (écarts réduits).

Dans les romans, nous pouvons observer que, lorsque les textes sont riches en virgules, ils sont pauvres en points, sauf dans *Le déluge*, *Le livre des fuites, Le printemps*, *Onitsha*, *Etoile errante* et *Pawana,* des livres à phrases relativement courtes, qui sont excédentaires des deux signes.

En revanche, dans les autres genres - les essais, les ouvrages ethnologiques et la biographie - où nous avons observé les phrases les plus longues du corpus, les deux signes sont déficitaires et il n'y a aucun signe d'effet de compensation par la virgule de l'espacement important entre deux points finaux.

Un autre signe qui marque la segmentation interne de la phrase est le point-virgule.

3.5.2. L'emploi du point-virgule.

Notre corpus compte 3532 points-virgules dont la distribution à travers le corpus est la suivante :

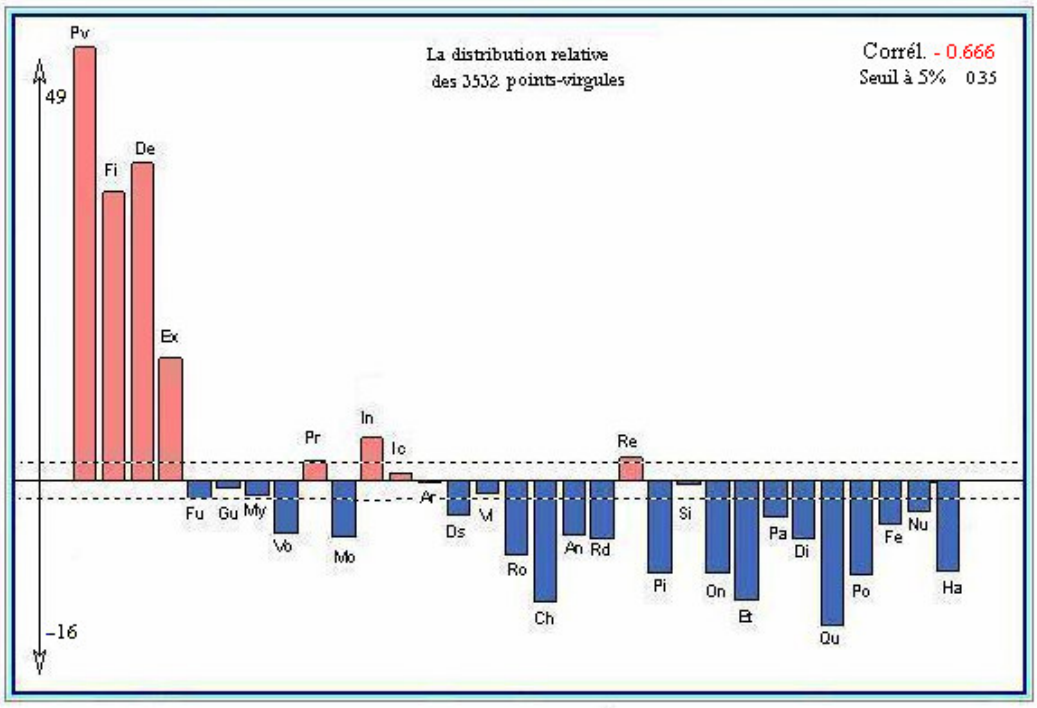


Figure n°60 : La distribution des points-virgules dans le corpus A.

L'histogramme montre que le point-virgule est surtout employé dans les premières œuvres du corpus et qu'il va pratiquement disparaître par la suite. C'est uniquement dans les essais *L'inconnu sur la terre* et dans *Le rêve mexicain* que nous pouvons noter un modeste emploi du point-virgule après l'abandon constaté à partir de *Le livres des fuites*.

Le Clézio en fait le plus grand usage dans son premier livre *Le procès-verbal*, où le signe sert d'outil dans le procédé de création du style très particulier du livre,

reflétant la personnalité du jeune homme tourmenté[250]. Déjà dans l'incipit nous pouvons noter l'emploi considérable du signe[251] :

> "Il y avait une petite fois, pendant la canicule, un type qui était assis devant une fenêtre ouverte ; c'était un garçon démesuré, un peu voûté, et il s'appelait Adam ; Adam Pollo."

Contrairement aux observations qu'a faites Étienne Brunet[252], il n'y pas d'effet net de parallélisme entre le point-virgule et la virgule. Il est vrai que les fréquences les plus importantes sont à observer, dans les deux cas, dans les premiers livres du corpus, mais les variations relativement importantes que nous avons relevées dans l'étude de la virgule sont absentes dans l'analyse du point-virgule.

Il est intéressant de comparer l'usage des point-virgules avec celui des deux-points, le troisième marqueur de segmentation interne de notre étude.

3.5.3. L'étude des deux-points.

En effet, la situation des deux-points est tout à fait différente de celle que nous venons de voir concernant les points-virgules. Non seulement l'effectif est considérablement plus important (5634 occurrences), mais la distribution est très différente :

---

[250] Rappelons que cet ouvrage a également un excédent remarquable de points d'exclamation, un nombre élevé de points d'interrogation et qu'il a les phrases les plus courtes du corpus. On voit à quel point la simple étude quantitative de données très formelles peut être révélatrice d'un phénomène stylistique original et significatif.
[251] *Le procès-verbal*, p. 3.
[252] É. Brunet (1988) : p. 83.

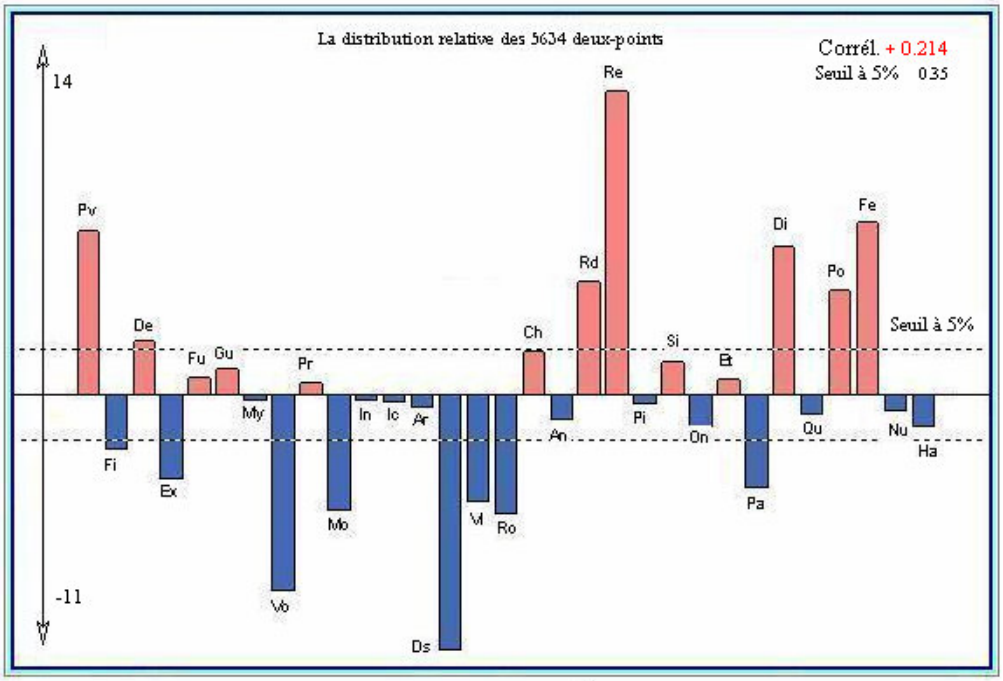


Figure n°61 : La distribution des deux-points dans le corpus A.

Les deux-points ont un rôle très différent de celui des points-virgules : ils servent non seulement à rapporter des propos, dans les dialogues par exemple, mais aussi à introduire une énumération ou une explication.

Ce sont probablement ces champs d'application du signe qui expliquent la différence dans l'usage que fait notre écrivain. En effet, Le Clézio les utilise surtout dans les ouvrages ethnologiques, la biographie et les essais, des ouvrages qui sont, par la nature de leur genre, riches en explications et en énumérations dont voici un exemple[253].

> "Quand Cortés, par exemple, orne son blason de l'orgueilleuse devise adressée au roi : "Moi, pour vous servir, sans égal" Bernal Díaz rectifie : lui-même, et ses compagnons ont aidé leur capitaine à "gagner cette gloire, cet honneur et cet état"."

[253] *Le rêve mexicain*, p. 13.

Dans les œuvres de fiction, nous pouvons observer, comme dans les études précédentes, la fréquence importante de deux-points dans *Le procès-verbal* et dans *Le déluge* de la première période. L'importance du signe dans *Voyage à Rodrigues* s'explique probablement par la présence de nombreuses explications liées au thème du roman (voyage initiatique, recherche d'un trésor, écriture d'une quête, récit d'une découverte) tandis que la fréquence considérable dans *Poisson d'or* est le reflet de ses nombreux dialogues. Les derniers signes de ponctuation faible que nous traitons dans cette étude sont les signes de parenthèse.

3.5.4. L'emploi des parenthèses.

L'usage de la parenthèse introduit, dans le corps d'une phrase, un élément qui, à la différence de l'incise, interrompt la construction syntaxique ; il s'agit souvent d'un complément ou d'un ajustement de la pensée. Dans notre corpus, nous recensons un total de 2476 occurrences tenant compte des deux signes qui limitent la parenthèse, c'est-à-dire que le couple parenthèse ouvrante et parenthèse fermante compte pour deux occurrences. L'histogramme suivant illustre leur distribution dans le corpus :

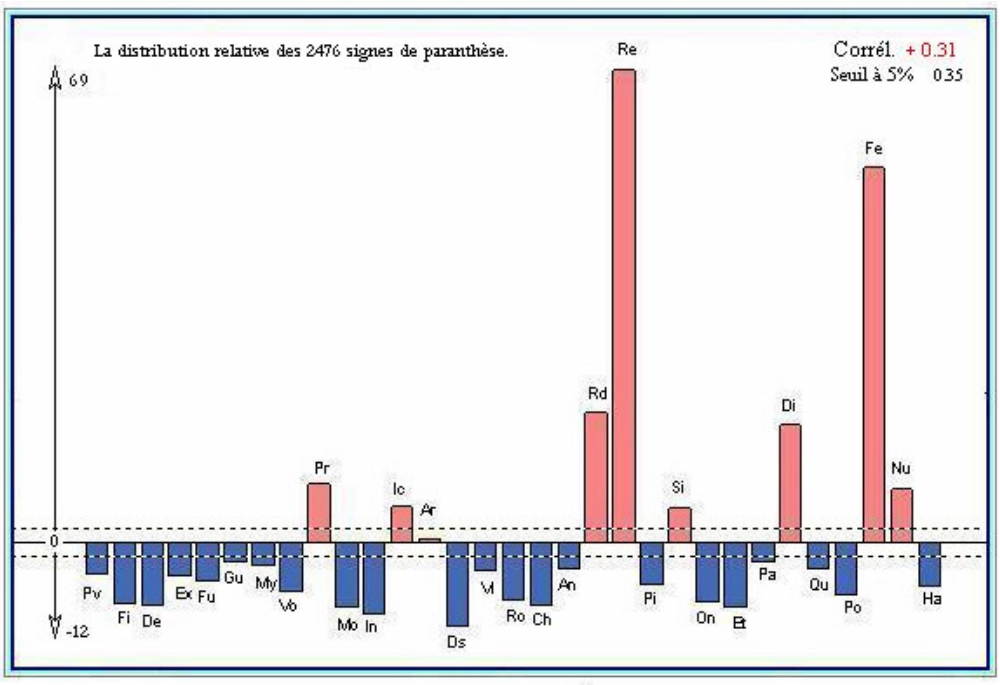


Figure n°62 : La distribution relative des signes de parenthèse dans le corpus A (écarts réduits).

L'histogramme permet d'observer que Le Clézio en fait un usage limité à quelques œuvres du corpus. Nous en trouvons dans les ouvrages ethnologiques, la biographie ainsi que dans le roman *Voyage à Rodrigues*. Dans ce livre, les nombreuses annotations et indications de plan qu'à laissées le grand-père sont souvent entourés de signes de parenthèse. Pour les autres ouvrages que nous venons de citer, l'explication de la relative richesse de parenthèses est bien plus triviale ; il s'agit presque toujours des renvois que fait l'auteur, dans le texte même, à d'autres passages du texte. Là encore, un procédé formel et quantifiable est révélateur d'un phénomène générique essentiel, en l'occurrence d'une différence dans le statut énonciatif du locuteur (ou plutôt du scripteur) : les parenthèses trahissent en effet une activité méta-énonciative de structuration du texte, une présence active du narrateur qui sont en principe interdites par les règles de la fiction classique.

Le recensement de ces différents signes de ponctuation faible permet de calculer et de connaître la longueur moyenne du segment interne de la phrase de la même façon que nous avons procédé pour dans le calcul de la longueur de la phrase, son entité supérieure, en tenant compte, cette fois, de tous les signes, ceux de la ponctuation faible aussi bien que de ceux de la ponctuation forte.

3.6. La segmentation interne dans l'œuvre, ponctuations fortes et ponctuations faibles.

La longueur moyenne des séquences intérieures de la phrase leclézienne est de 7.40 et la distribution dans les différents ouvrages est la suivante :

Texte	occurrences	ponct. faibles	ponct. fortes	tot. ponct.	longueur
Pv	92955	8664	4247	12911	7,20
Fi	101666	10514	4940	15454	6,58
De	122269	11508	6320	17828	6,86
Ex	90616	8636	5001	13637	6,64
Fu	104451	9329	6608	15937	6,55
Gu	110602	9161	5857	15018	7,36
My	9771	925	532	1457	6,71
Vo	113790	9329	6009	15338	7,42
Pr	8690	853	346	1199	7,25
Mo	99114	7118	5325	12443	7,97
In	124433	12746	5886	18632	6,68
Ic	2637	313	129	442	5,97
Ar	2910	203	161	364	7,99
Ds	150843	13441	5881	19322	7,81
Vi	15266	1512	664	2176	7,02
Ro	83380	7567	3584	11151	7,48
Ch	134435	11413	6251	17664	7,61
An	24313	1904	1209	3113	7,81
Rd	38238	3578	1110	4688	8,16
Re	87414	8043	2527	10570	8,27
Pi	70230	6197	3750	9947	7,06
Si	2463	252	92	344	7,16
On	76046	6458	4169	10627	7,16
Et	118756	10583	5650	16233	7,32
Pa	10593	930	488	1418	7,47
Di	73266	5988	2092	8080	9,07
Qu	165612	13085	8537	21622	7,66
Po	86051	7030	4480	11510	7,48
Fe	71025	6475	2215	8690	8,17
Nu	21563	1924	843	2767	7,79
Ha	68261	5395	3249	8644	7,90

Tableau n°29 : La longueur de la séquence intérieure de la phrase leclézienne (ponctuations faibles).

Ce tableau montre bien que la longueur moyenne du segment intérieur est assez homogène dans notre corpus. Les séquences les plus courtes se trouvent dans

Vers les icebergs avec une moyenne de 5,97 mots par segment et les plus longues sont à trouver dans *Diego et Frida* avec 9,07 mots par segment.

Il est intéressant de comparer ces résultats avec ceux obtenus dans l'étude de la longueur de la phrase afin de connaître le rapport entre les deux signes :

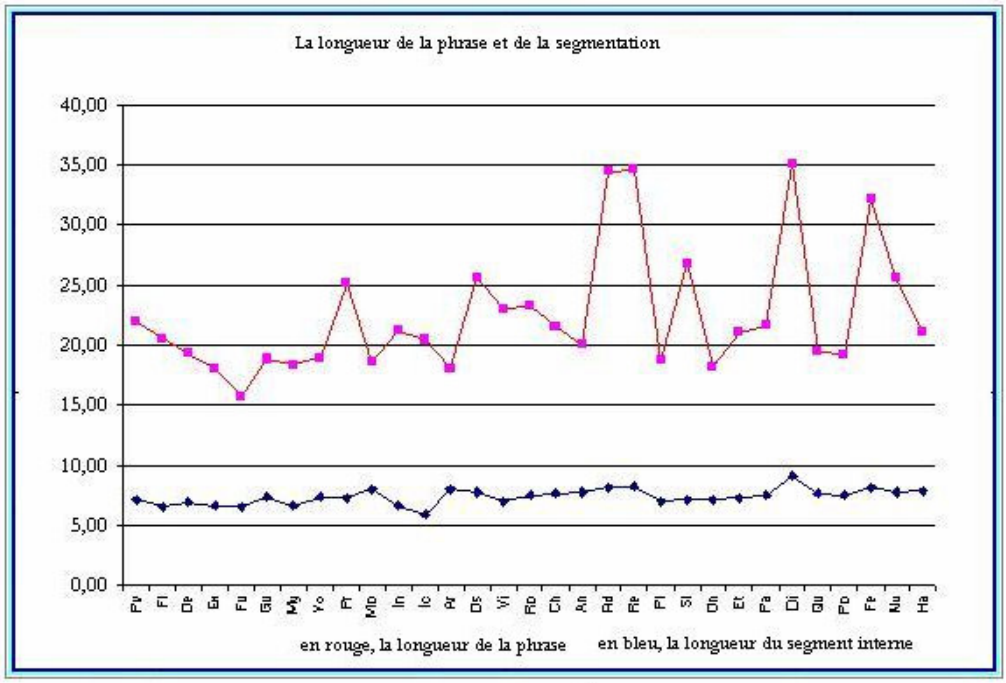


Figure n°63 : La segmentation du discours.

L'histogramme permet de constater qu'il n'y a aucun parallélisme entre les deux entités. Les variations importantes que nous avons pu observer dans la longueur de la phrase sont neutralisées lorsque l'on prend en considération également la ponctuation faible de la phrase. La relative symétrie des deux courbes que l'on peut observer chez d'autres auteurs fait défaut chez J.M.G. Le Clézio.

En effet, il semble que la segmentation, tenant compte des signes de ponctuation faible, est insensible au genre littéraire chez notre écrivain. Les séquences sont sensiblement de la même longueur indépendamment de la ponctuation forte.

L'analyse factorielle ci-dessous réunit les huit signes de ponctuation et restitue les oppositions et les alliances des différents signes et ouvrages du corpus :

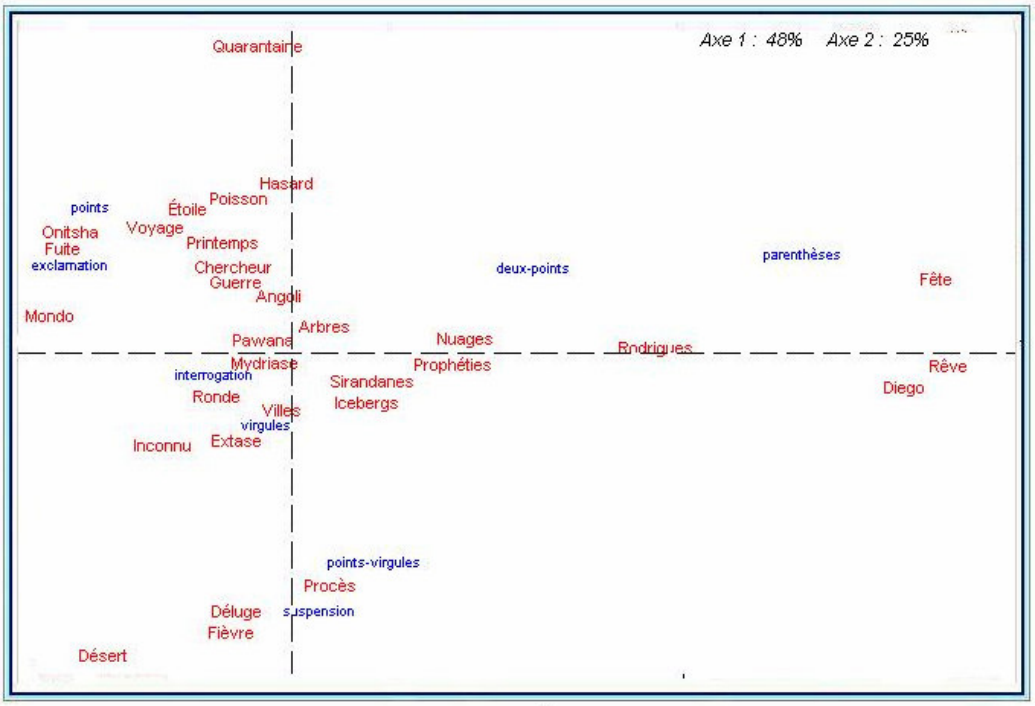


Figure n°64 : Analyse factorielle de la ponctuation chez Le Clézio.

Le facteur prédominant de l'analyse est, comme toujours, celui du genre, les ouvrages ethnologiques et la biographie se trouvent à droite du graphique près des parenthèses. Remarquons que *Voyage à Rodrigues* se trouve placé à proximité de ce groupe, pour les raisons évoquées ci-dessus. Les premiers romans de l'école du "nouveau roman" se regroupent en bas de la figure avec les points de suspension et les points-virgules. Les autres œuvres romanesques accaparent les autres signes de ponctuation au milieu du graphique, *Le livre des fuites* et *Mondo et autres histoires* très proches du point d'exclamation. L'analyse factorielle dégage également un second facteur qui parcourt la chronologie de l'œuvre du bas vers le haut.

Le recensement de tous les signes de ponctuation permet aussi d'étudier les proportions des différents signes de notre corpus et de les comparer avec l'usage qu'en font d'autres écrivains.

La figure ci-dessous illustre les proportions des principaux signes de ponctuation chez J.M.G. Le Clézio :

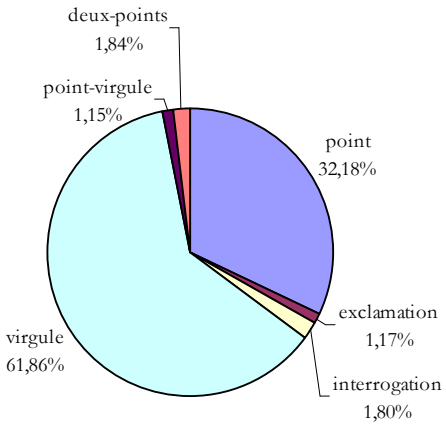


Figure n°65 : La proportion des signes de ponctuations dans le corpus (valeurs absolues).

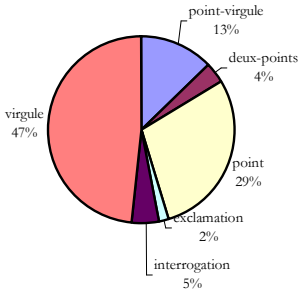
Nous constatons que la grande majorité du graphique est occupée par le point et la virgule qui à eux seuls occupent 94,04% du diagramme circulaire. Les 6% restants sont partagés en parts à peu près égales entre les quatre autres signes minoritaires[254].

Étienne Brunet a comparé, dans ses études[255], la proportion des signes de ponctuation chez plusieurs écrivains et il a obtenu les résultats suivants :

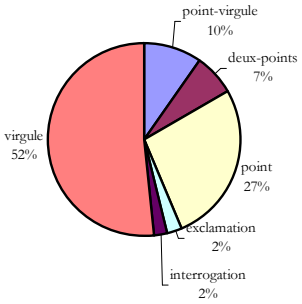
[254] Afin d'effectuer une comparaison adéquate avec les résultats des études d'Étienne Brunet, nous avons exclu de cette étude les points de suspension et les signes de parenthèse.
[255] É. Brunet (1988) : p. 87-90 et "La phrase de Zola", *Actes du Colloque de Victoria*, in *La critique littéraire et l'ordinateur*, B. Derval, M. Lenoble (éds.), Montréal, 1987, p. 117-120.

Figure n°66 : La proportion des signes de ponctuations des différents auteurs.

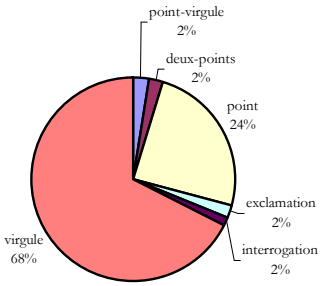
Rousseau



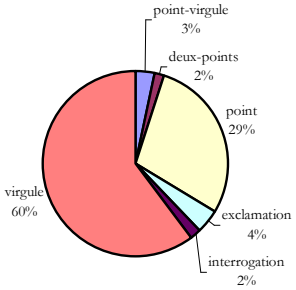
Chateaubriand



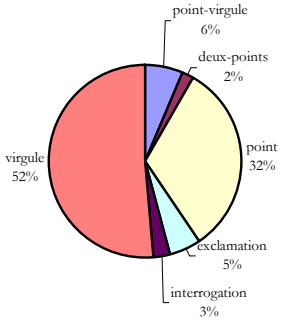
Proust



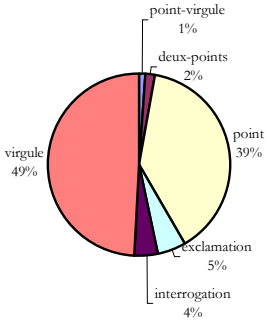
Zola



Hugo



Giraudoux



Il a pu constater que la part du point était à peu près d'un tiers de l'ensemble et celle de la virgule de la moitié chez les six écrivains faisant partie de son étude. Le Clézio respecte *grosso modo* ces proportions : 32.18% et 61.86% respectivement. Nous constatons toutefois la part importante de virgules, plus considérable que celle de Zola et qui s'approche même de la part exceptionnelle de virgules chez Proust, sans pour autant diminuer la part de points finaux. Quant aux autres signes, Le Clézio les exploite moins que les autres écrivains ; s'agit-il d'une question de tempérament de notre auteur qui refuserait l'exubérance des émotions ? Il est difficile d'en juger, mais leur nombre réduit correspond bien à la sobriété et à la modestie du style de Le Clézio.

Nous avons vu à travers ces différentes analyses que la phrase leclézienne est relativement longue et qu'elle varie selon les genres littéraires dans lesquels s'exprime l'auteur. En revanche, la segmentation du discours n'atteint jamais des valeurs extrêmes et sa stabilité est notable. La proportion importante de la virgule chez Le Clézio semble être le reflet d'une phrase longue et peut-être complexe.

Toutefois, un autre facteur important quant à la complexité de la phrase est celui de la subordination et de la coordination, qui méritent d'être étudiées avant de tirer des conclusions très précises sur la phrase leclézienne.

3.7. La complexité de la phrase.

> "Pour moi, Descartes ou Malebranche n'avaient jamais cessé de représenter des mondes étrangers, ceux du raisonnement et de la dialectique. Je les lisais, je les comprenais, mais s'il fallait mettre quelque chose en doute, c'étaient eux, c'était leur étonnante prétention à tout régulariser, à bâtir un monde sur les fondements du langage humain, ce langage si pauvre, si malhabile. Vous savez, la pensée divisée en deux parties, parce que la phrase se divise en deux parties, que la cause appelle la conséquence, le thème le prédicat, la principale la subordonnée."
>
> *La fièvre,* page 146.

Nous avons vu, dans le chapitre précédent, que la longueur de la phrase dépend étroitement de sa complexité. De façon générale, plus les expansions et les ramifications faites à partir du noyau de la phrase seront riches, plus la phrase sera longue. Les expansions peuvent être rattachées au noyau par subordination ou par coordination. Comme l'exprime Dominique Labbé[256] : "on imbrique ou on empile." Suivant la technique qui prédominera, les structures de phrases et les styles seront assez différents. Dans son étude sur la phrase, Conrad Bureau propose de considérer la phrase comme une construction linguistique et pas seulement comme une construction typographique avec les limites graphiques que sont les signes de ponctuation. Il écrit ceci concernant le relatif artifice que constituent les signes de ponctuation[257] :

> "C'est pourquoi tout segment, monème ou syntagme, qui est dans un lien d'implication réciproque ou d'implication à sens unique avec un autre segment et qui est séparé de ce dernier par un artifice de ponctuation ou de typographie n'est pas une phrase, mais un élément de

[256] D. Labbé (1991) : p. 179.
[257] C. Bureau (1976) : p. 72.

> phrase ; la relation d'implication du premier type ou du second, en effet, prive ce segment du caractère d'autonomie syntaxique totale, d'autosuffisance qui est nécessaire pour qu'il y ait phrase ; dans un tel cas, la ponctuation relève du style. La phrase est donc syntaxiquement une structure close, indépendante d'une autre structure qui l'engloberait."

Il propose une définition de la longueur de la phrase qui se fonde non pas sur des critères arbitraires non linguistiques comme les blancs d'imprimerie et la ponctuation, mais sur un critère uniquement linguistique, celui de la fonction. Sa technique consiste à mesurer la longueur d'une phrase en fonction de la somme des syntaxèmes[258] qui la composent.

Cette idée, que nous exploiterons en premier, est intéressante en raison du caractère particulier de l'écriture leclézienne, surtout des premiers livres appartenant à l'école du "nouveau roman". Le lecteur de ces œuvres a, en effet, souvent l'impression que la ponctuation est utilisée pour l'effet de style plutôt que pour délimiter la phrase.

Il est évident que la technique que propose Conrad Bureau, qui demande une analyse manuelle de chaque phrase, ne peut pas s'appliquer à un corpus de grande taille[259]. En effet, en l'état actuel des instruments informatiques, ces procédures ne sont pas automatisables, ce qui rend cet examen impossible sur l'ensemble du corpus.

Toutefois, l'outil informatique nous permet d'effectuer une étude de la phrase complexe (sans pour autant prétendre à l'exhaustivité d'une telle analyse) et de détecter la proportion des expansions développées à partir d'un noyau de phrase, dans un corpus lemmatisé, non seulement par le recensement des signes de

---

[258] Pour plus de détails sur ses définitions cf. C. Bureau (1976) : p. 92.

[259] Le corpus de Conrad Bureau est très limité, il comprend 3000 phrases, constitué de 2000 phrases de *Combray* et de *Du côté de chez Swann* ainsi que de 500 phrases tirées de *La guerre* de Le Clézio et de 500 de *La porte étroite* de Gide.

ponctuation faible comme nous l'avons fait dans le chapitre précédent, mais aussi par l'inventaire des marqueurs de coordination et de subordination. Cet inventaire nous permettra d'apporter enfin quelques précisions supplémentaires concernant la phrase leclézienne.

3.7.1. La phrase incomplète.

Dans une comparaison de la construction de la phrase de Georges Perec avec celle de J.M.G. Le Clézio[260], Gunnel Engwall étudie la relation qui existe entre la phrase graphique et le macrosyntagme[261]. Elle distingue, dans son article, quatre types de phrases graphiques par rapport à ce macrosyntagme :

1. La phrase graphique coïncide avec le macrosyntagme.
2. La phrase graphique se compose de plusieurs macrosyntagmes.
3. La phrase graphique ne contient qu'une partie d'un macrosyntagme.
4. La phrase graphique comporte une partie d'un macrosyntagme et un autre macrosyntagme complet.

L'étude s'appuie sur des échantillons de deux textes[262] comportant 20000 occurrences chacun. Elle constate que la majeure partie des phrases des deux écrivains sont attestées dans le premier groupe (64% des phrases de Le Clézio). La proportion des phrases graphiques qui contiennent plusieurs macrosyntagmes est de 26%. Ce sont les résultats de l'analyse des exemples du troisième groupe qui sont les plus remarquables de son étude et qui différencient nettement notre auteur. En effet, l'effectif de cette classe compte 10% du corpus tiré de *Le déluge* de Le Clézio tandis qu'il ne représente que 2% du corpus de *Les choses* de Perec. Ces phrases syntaxiquement incomplètes et qui ne sont pas construites selon les règles

---

[260] G. Engwall, "Valeurs quantitatives et impressions du lecteur", in *Papers from the Fifth Scandinavian Conference of Linguitics*, Frostvallen, 27-29 avril 27-29, 1979, T. Pettersson (éd.), Stockholm, Acta Universitatis Lundensis Sectio i Theologica Juricica Humaniora, tome I, p. 93-103.
[261] Pour plus de détails sur la définition du macrosyntagme cf. B. Loman et N. Jörgenssen (1971) et U. Teleman (1974).
[262] *Le déluge* de J.M.G. Le Clézio et *Les choses* de G. Perec.

de ponctuation habituelles sont intéressantes par leur nombre aussi bien que par leur construction.

Conrad Bureau s'intéresse également dans son analyse à ces types de phrases. Dans le fragment suivant, par exemple, il considère qu'il n'y a qu'une seule phrase, malgré le point et, parce qu'il y a implication réciproque entre le prédicat à base nominale *les lèvres* et l'actualisateur spécialisé *il y avait*, relation identique à celle qui existe entre *il y avait* d'une part, et *ce creux*, *cette rigole* d'autre part, au début de la phrase[263] :

> "Au-dessous du nez, il y avait ce creux, cette rigole, par où coulait la morve.
> Puis les lèvres, les deux bourrelets violacés striés de petites rides, de petites coupures."

Il considère également le fragment suivant comme une seule phrase : "Ils se retiennent de respirer. Pendant des heures."[264] Il considère que puisque le segment *pendant des heures* est dans une relation d'implication à sens unique par rapport à ce qui précède, qu'il lui est donc subordonné par l'intermédiaire du monème fonctionnel *pendant*, il n'y a qu'une seule phrase.

D'autres procédés souvent exploités par l'auteur comme *et* ou *or* à l'initiale de phrase marquent un certain rapport entre deux phrases, impliquant des liens forts, cette fois-ci sémantiques, et non syntaxiques[265] :

> "C'est ça. Je peux vous inventer des fables. Sur-le-champ. Tenez. Je vais vous donner des titres. Écoutez. La légende du palmier nain qui voulait voyager en Europe Orientale. Ou. L'ibis qu'un représentant de commerce métamorphosa en fille. Ou. Asdrubal l'homme-aux-deux-

[263] *La guerre*, p. 16-17 .
[264] *Le procès-verbal*, p. 249.
[265] *Ibid.*, p. 251-252.

> bouches. Et. L'Histoire d'amour d'un roi de carnaval et d'une mouche. Ou. Comment Zoé reine du Peloponnèse trouva le trésor du syrinx en dentelles. Ou. Le courage des typanosomes. Ou. Comment tuer les serpents à sonnette. C'est simple. Il faut savoir trois choses. Les serpents à sonnettes. Sont très orgueilleux. N'aiment pas le jazz. Et dès qu'ils voient un edelweiss ils tombent en catalepsie."

En effet, Le Clézio se sert souvent de cette technique pour mettre des mots en relief comme l'illustre également l'exemple suivant[266] :

> "Ni harmonieux. Ni efficaces. Ni beaux ni laids. Ni vastes ni resserrés. Ni terribles. Ni humains. Ni délicats. Ni joyeux."

ou bien pour donner au récit un caractère oral[267] :

> "Et lui, il. Et lui, il s'en foutait."

C'est évidemment une tâche presque impossible de relever toutes ces phrases dans un corpus de l'ampleur du nôtre pour en constituer un inventaire. Néanmoins, le traitement automatique effectué par le logiciel Hyperbase permet de repérer certaines particularités concernant la ponctuation dans notre corpus, notamment celle de commencer la phrase par une coordination de subordination ou de coordination.

3.7.2. La conjonction à l'initiale de la phrase.

Nous avons établi l'inventaire des phrases commençant par une conjonction de subordination ou de coordination. Toutefois, nous avons dû écarter certains

[266] *Le déluge*, p. 59.
[267] *L'extase matérielle*, p. 239-240.

coordonnants et subordonnants (par exemple *que* et *comme*) de l'analyse à cause de leurs caractères homographiques et polysémiques qui introduiraient des facteurs d'erreur dans l'étude. Le tableau ci-dessous rend compte de la fréquence et de la distribution des effectifs recensés :

	. Et	. Car	. Ou	. Ni	. Donc	. Puis	. Alors
Pv	73	4	27	1	2	29	29
Fi	115	11	16	1	1	63	24
De	1	0	0	0	0	0	0
Ex	11	26	4	12	0	10	16
Fu	80	1	23	0	0	54	19
Gu	81	2	40	1	0	44	46
My	2	0	3	0	0	0	3
Vo	38	1	24	1	1	59	59
Pr	1	3	0	0	0	3	4
Mo	13	1	8	0	0	64	51
In	34	8	11	2	0	30	61
Ic	0	1	1	0	0	0	1
Ar	1	0	0	0	0	3	2
Ds	43	0	14	0	1	76	108
Vi	0	1	1	0	0	4	10
Ro	29	0	9	0	0	32	49
Ch	24	0	6	0	0	48	63
An	2	0	0	0	0	22	12
Rd	6	1	3	1	0	1	7
Re	35	8	0	0	0	9	25
Pi	20	0	6	0	0	17	17
Si	0	1	0	0	0	0	1
On	22	1	4	1	0	39	69
Et	50	3	5	0	1	54	64
Pa	3	0	1	0	0	5	8
Di	26	0	0	0	0	7	8
Qu	51	0	15	0	1	67	27
Po	60	0	15	3	0	25	16
Fe	34	4	3	0	0	11	16
Nu	13	1	3	1	0	5	6
Ha	29	1	13	1	0	21	8
total	897	79	255	25	7	802	829

Tableau n°30 : La fréquence de phrases commençant par un coordonnant.

Nous pouvons constater que la fréquence de ce phénomène semble d'un côté être concentrée sur une partie distincte du corpus, la partie initiale essentiellement, et de l'autre sur les livres de fiction riches en dialogues ou monologues dans tout le

corpus. Les courbes ci-dessous permettent de mieux se rendre compte de la distribution des 2894 situations relevées :

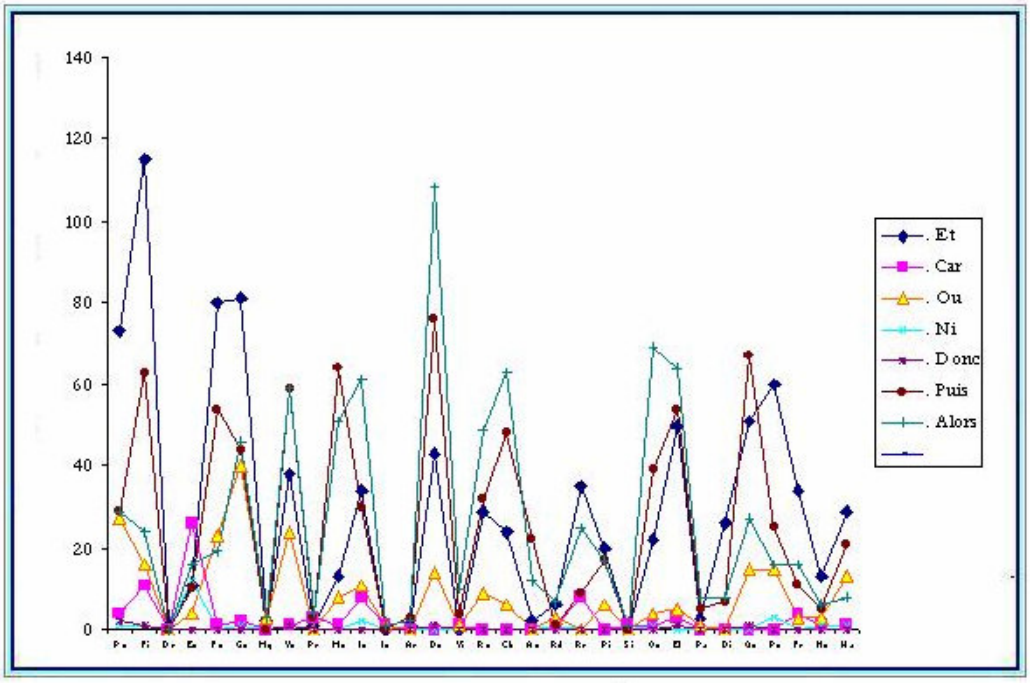


Figure n°67 : Coordonnants en position initiale de la phrase.

Bien que cette analyse sur les phrases de ce type ne puisse prétendre être complète, elle nous fournit une indication importante. Elle permet de constater que les coordonnants en position initiale des la phrase sont souvent répartis sur la première partie de l'œuvre, à l'exception du roman *Le déluge* ainsi que sur les ouvrages romanesques comme *Mondo et autres histoires*, *Désert*, *Le chercheur d'or*, *Onitsha*, *Etoile errante* et *La quarantaine*[268]. Les ouvrages appartenant à un autre genre contiennent nettement moins ce type de phrases.

Si nous prenions ces résultats en considération pour fonder une réflexion sur notre analyse de la longueur de la phrase, il y aurait plusieurs constatations à faire. Ces phrases, où la ponctuation est utilisée pour l'effet de style plutôt que pour délimiter la phrase, pourraient - si l'on prend en considération la notion du

[268] Les valeurs absolues sont ici pondérées en pourcentage du nombre de phrases de chaque texte.

macrosyntagme ou de la phrase comme une entité linguistique plutôt que typographique - modifier l'estimation de la longueur moyenne de la phrase, en la rallongeant considérablement. L'acceptation de cette idée rallongerait les phrases des quatre premiers livres où nous avons observé une longueur moyenne, variant entre 18.12 et 21.89[269]. Toutefois, le nombre recensé étant relativement modeste, il ne modifierait pas à lui seul nos statistiques de façon significative et ne peut que constituer une indication sur un phénomène potentiellement important dans les premières œuvres de Le Clézio.

3.7.3. Les subordonnants.

Nous avons déjà évoqué la difficulté dans le recensement des subordonnants en français. En effet, l'homographie pose de nombreux problèmes, par exemple dans le cas du *que* qui endosse le rôle de conjonction aussi bien celui du pronom relatif. *Comme* et *si* peuvent aussi être des adverbes tandis que *quand* est souvent un interrogatif et ainsi de suite.

Quand nous avons commencé notre recherche, nous avions constitué l'inventaire informatique des subordonnants en établissant des listes de formes, avec l'aide du logiciel Hyperbase, composées de : *comme*, *lorsque*, *parce que*, *pourvu*, *puisque*, *quand*, *que*, *quoique* etc. L'inconvénient de ce système est qu'il ne tient pas compte de l'homographie et de la polysémie que nous venons d'évoquer, et nous nous sommes vue obligée d'écarter certains éléments de la liste.

La nouvelle version lemmatisée d'Hyperbase permet désormais, grâce à l'analyse syntaxique effectuée par le programme, d'établir l'inventaire des subordonnants de façon beaucoup plus fiable et confortable.

[269] Cf. tableau n°27, p. 234.

Le tableau ci-dessous regroupe les fréquences absolues ainsi que les valeurs de l'écart réduit des subordonnants et des coordonnants dans les différents ouvrages de notre corpus :

Subordonnants et Coordonnants

Texte	sub. fréqu.	sub. écart	coord. fréqu.	coord. écart
Pv	1695	3,7	2669	4
Fi	1384	-8	3040	6,3
De	1549	-12	3853	9,6
Ex	1514	-0,7	2781	6,7
Fu	1234	-12,7	2418	-7,3
Gu	1693	-4	3099	2,5
My	135	-2,5	250	-1,1
Vo	2001	1,7	3443	6,8
Pr	80	-5,7	197	-2,6
Mo	1963	7,7	3179	10,7
In	2276	3,6	3596	4
Ic	57	1,6	66	-0,9
Ar	74	3,3	106	2,8
Ds	3262	14,6	4641	9,4
Vi	218	-2,7	386	-1,5
Ro	1733	8,8	2490	5,2
Ch	2341	1,5	3070	-9,5
An	425	0,6	697	1,6
Rd	583	-2,6	882	-4,8
Re	1111	-9,4	2822	10,4
Pi	1495	9,1	1406	-11,3
Si	32	-1,6	74	0,8
On	1150	-3,8	1357	-15,5
Et	2380	8,6	2887	-5,7
Pa	184	0,3	215	-4,3
Di	934	-8,7	2129	3,7
Qu	2423	-1,5	3272	-18,4
Po	1988	14,3	2168	-3,1
Fe	856	-10	2182	6,4
Nu	285	-4,3	512	-3,1
Ha	1174	0,6	1516	-7,7
total	38229		61403	

Tableau n°31 : Inventaire des subordonnants et des coordonnants du corpus.

La distribution des 38529 subordonnants dans notre corpus comporte de nombreuses variations. Le graphique suivant permet de mieux observer les importantes variations :

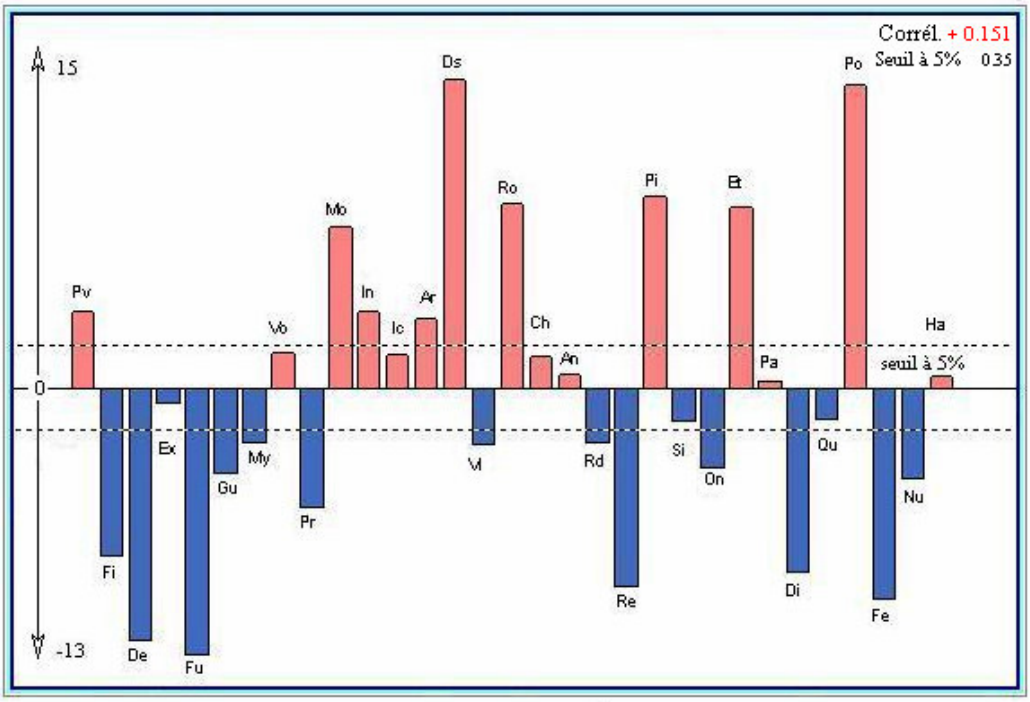


Figure n°68 : Les subordonnants du corpus B (distribution relative).

L'histogramme permet de constater le même mouvement tripartite que nous avons pu constater auparavant, bien que moins distinct cette fois-ci. Nous constatons que, après avoir utilisé la subordination dans son premier livre *Le procès-verbal*, Le Clézio en abandonne l'usage jusqu'au recueil de nouvelles *Mondo et autres histoires*, où nous pouvons encore une fois observer le grand chavirement des chiffres qui correspond à la rupture dans l'œuvre. La deuxième période de l'œuvre exploite fréquemment la subordination, l'usage le plus fréquent se trouve dans les romans de cette période. Enfin, la troisième période commençant par *Voyage à Rodrigues* se caractérise par son extrême hétérogénéité, alternant des valeurs déficitaires et des valeurs nettement excédentaires. Le graphique permet d'observer qu'il n'y a pas, pour une fois, d'opposition générique dans l'utilisation des subordonnants, à part la constatation que les ouvrages ethnologiques et la biographie sont déficitaires. Il semble en effet que l'usage par notre auteur de la subordination dans son ensemble est indépendant du genre littéraire dans lequel il s'exprime et que son utilisation dépend d'autres facteurs.

C'est à l'intérieur de la subordination que nous retrouvons l'opposition du genre littéraire. L'analyse factorielle ci-dessous permet d'observer les liens entre les différents sous-corpus et les subordonnants suivants : *que* (toutes fonctions confondues, sauf celles du relatif) *pourvu que*, *quand*, *lorsque*, *quoique*, *comme*, *puisque* et *parce que* :

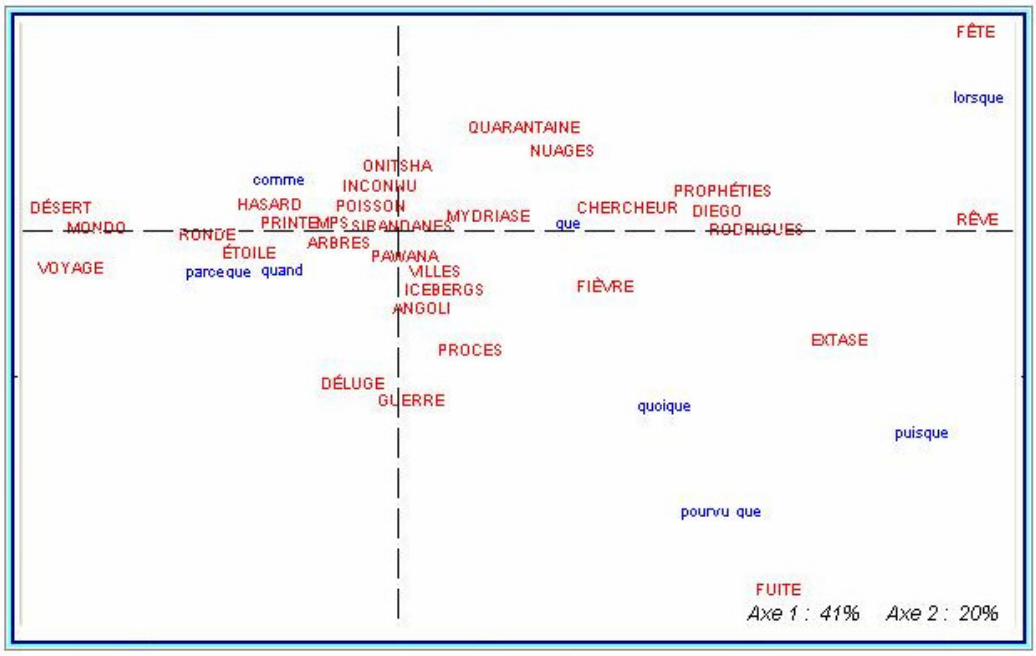


Figure n°69 : Analyse factorielle de la distribution des subordonnants les plus fréquents dans le corpus B.

Le graphique montre en effet que la principale distinction s'établit entre les différents genres, avec *comme*, *parce que* et *quand* près des romans de la deuxième période et à l'opposé, près des essais et des ouvrages ethnologiques, les formes plus soutenues de *puisque*, *lorsque*, *quoique* et *pourvu que*. Une autre opposition importante est celle entre formes propres aux récits suivant la chronologie des événements (temps et cause) et formes nettement argumentatives (justification, concession et condition). Quant au *que*, il se trouve en position assez neutre, au milieu du graphique.

Revenons au graphique précédent[270] qui considère la subordination dans son ensemble. Au premier regard cet histogramme ressemble beaucoup à la courbe illustrant la longueur de la phrase dans le corpus[271], rendant ainsi compte des liens importants qui existent entre la phrase longue et la subordination. Toutefois, la comparaison détaillée des deux graphiques permet la détection de différences assez considérables. Parmi les ouvrages à phrases longues, quelques-uns sont déficitaires en subordonnants. Il s'agit de *Les prophéties de Chilam Balam*, *Voyage à Rodrigues*, *Le rêve mexicain*, *Diego et Frida* et *La fête chantée*. Il semblerait, d'après notre analyse, que la subordination ne soit pas la raison de la longueur de leurs phrases.

Il peut aussi sembler paradoxal de trouver les valeurs les plus excédentaires dans les livres ayant les phrases les plus courtes, notamment dans les recueils de nouvelles, textes brefs par définition, et dans les grands romans incantatoires *Désert*, *Etoile errante* et *Poisson d'or* ainsi que dans le premier roman *Le procès-verbal*. Il s'agirait probablement dans ces cas de la relative richesse de dialogue qui se reflète par l'usage d'un langage près de l'oral, ayant plus souvent recours aux subordonnants.

L'histogramme de la page suivante illustrant la distribution des subordonnants dans le corpus romanesque, le corpus E, met en exergue ce phénomène :

[270] Cf. figure n°68, p. 269.
[271] Cf. figure n°56, p. 235.

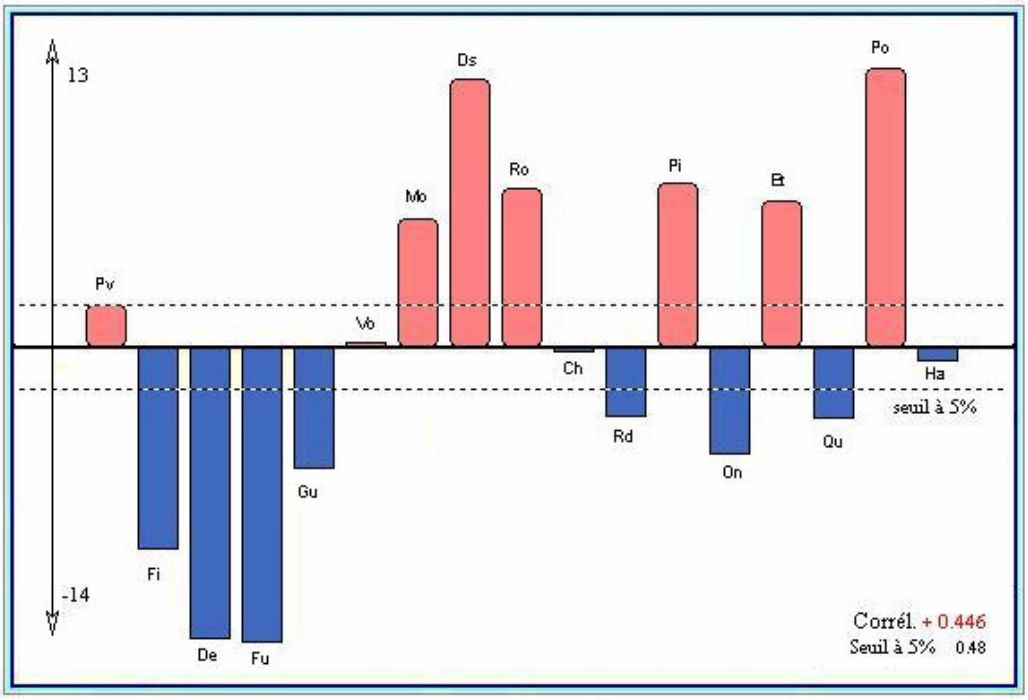


Figure n°70 : Les subordonnants du corpus E (distribution relative).

Nous trouvons en effet des valeurs excédentaires dans les mêmes ouvrages riches en dialogue avec un langage ayant un caractère oral, et les valeurs déficitaires sont à trouver dans les ouvrages du caractère plus narratif.

Un autre facteur déterminant pour la longueur et la complexité de la phrase est celui de la coordination.

Le graphique suivant compare la fréquence des deux catégories dans le corpus B :

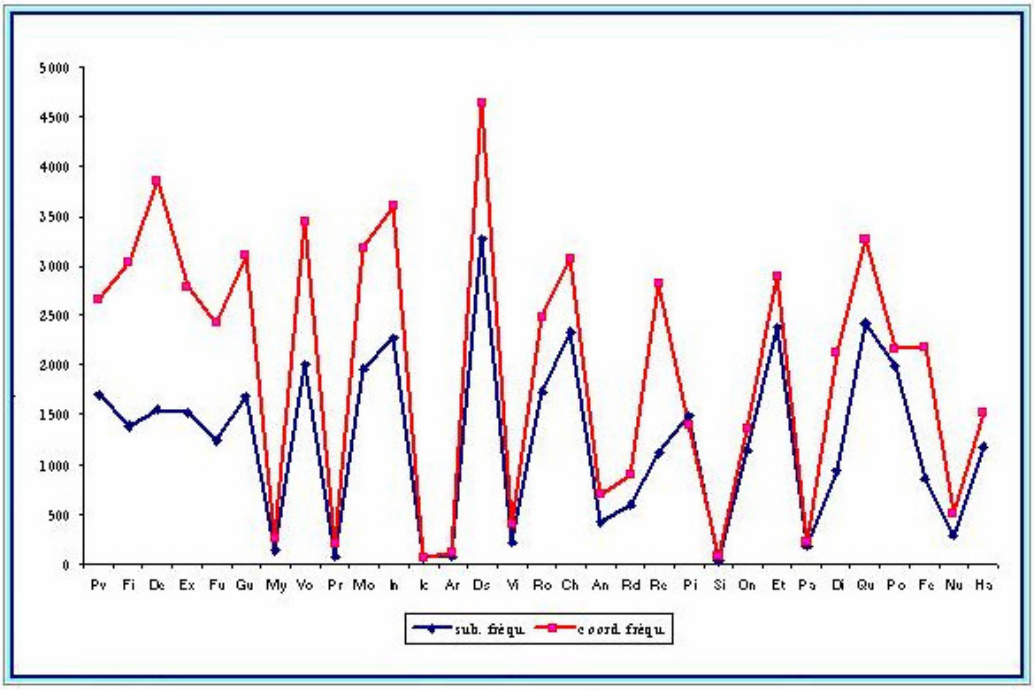


Figure n°71 : Comparaison de la distribution des coordonnants et de subordonnants dans le corpus (fréquences absolues).

Les deux courbes offrent un parallélisme certains. Mais les coordonnants (61403 occurrences) sont ici beaucoup plus fréquents que les subordonnants (38229 occurrences). Dans ses études sur le vocabulaire des auteurs du XIX^{ème} siècle, Étienne Brunet a observé la même relation dans le corpus Hugo, contient 63700 subordonnants et 72201 coordonnants et une relation inverse avec un nombre des subordonnants supérieur à celui ces coordonnants chez Proust avec 46574 subordonnants pour 40547 coordonnants[272].

[272] Rappelons qu'il n'y a pas d'analyse de la phrase dans ces études et l'homographie de *que* n'est pas réalisée, d'où peut provenir un surplus.

3.7.4. Les coordonnants.

La fréquence de coordonnants de notre corpus est effectivement très importante. Le coordonnant le plus fréquent est *et* qui domine massivement avec ses 44.989 occurrences représentant 73.2% de la catégorie.

L'histogramme ci-dessous rend compte de la distribution interne des coordonnants dans notre corpus, prenant en considération les écarts réduits :

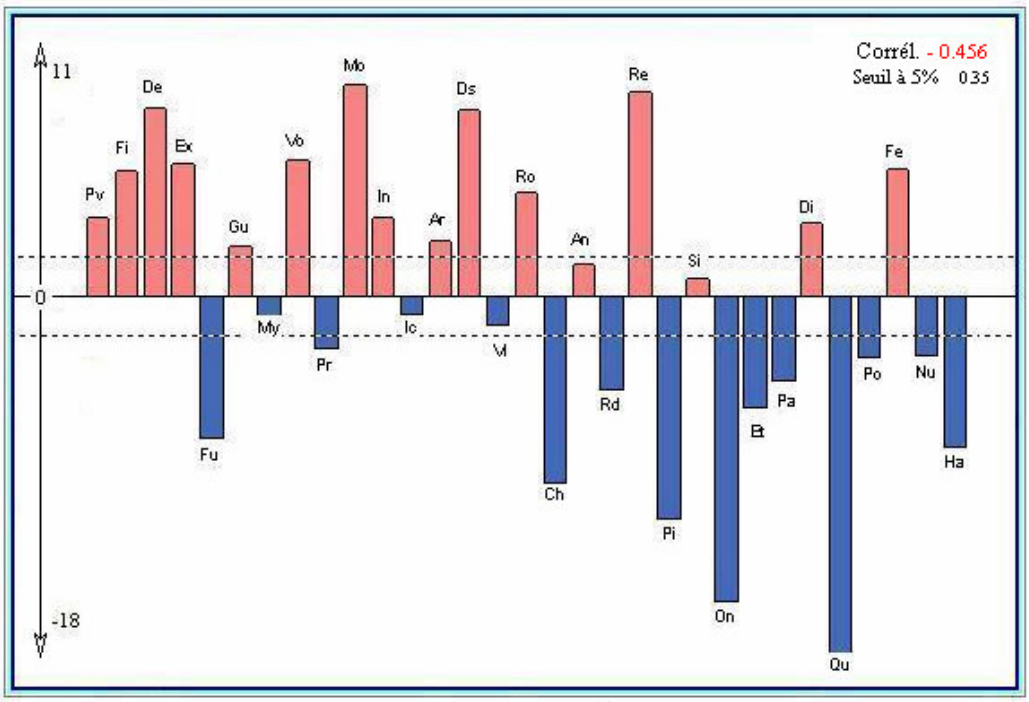


Figure n°72 : Les coordonnants du corpus B (distribution relative).

Le graphique montre que la distribution des coordonnants présente des variations considérables dans le corpus. La tendance dominante cette fois-ci semble être celle de l'évolution chronologique (coefficient de corrélation de -0,456). Après les valeurs excédentaires du début, l'usage des coordonnants décline, à quelques exceptions près, au fur et à mesure que l'œuvre progresse. Les ouvrages les plus excédentaires sont les romans du milieu de l'œuvre ainsi que l'essai *Le rêve mexicain*

tandis que les livres les plus déficitaires sont les romans de la troisième période : *La quarantaine* (écart réduit de -18), *Onitsha*, *Printemps et autres saisons*, *Le chercheur d'or* et *Hasard. Le livre des fuites* de la première période appartient également à ce groupe. Il semble que cet histogramme, tout comme celui illustrant la distribution de subordonnants, soit globalement indépendant au genre littéraire et qu'il n'y a pas d'opposition franche entre les genres. Toutefois, les livres ethnologiques et la biographie *Diego et Frida*, tous excédentaires, résistent mieux au déclin, alors qu'ils étaient déficitaires en subordonnants. Cela indique que la longueur des phrases, qui les caractérise, est une longueur par "empilement" et non par "imbrication".

L'histogramme ci-dessous qui ne tient compte que du corpus exclusivement romanesque permet de cibler l'analyse :

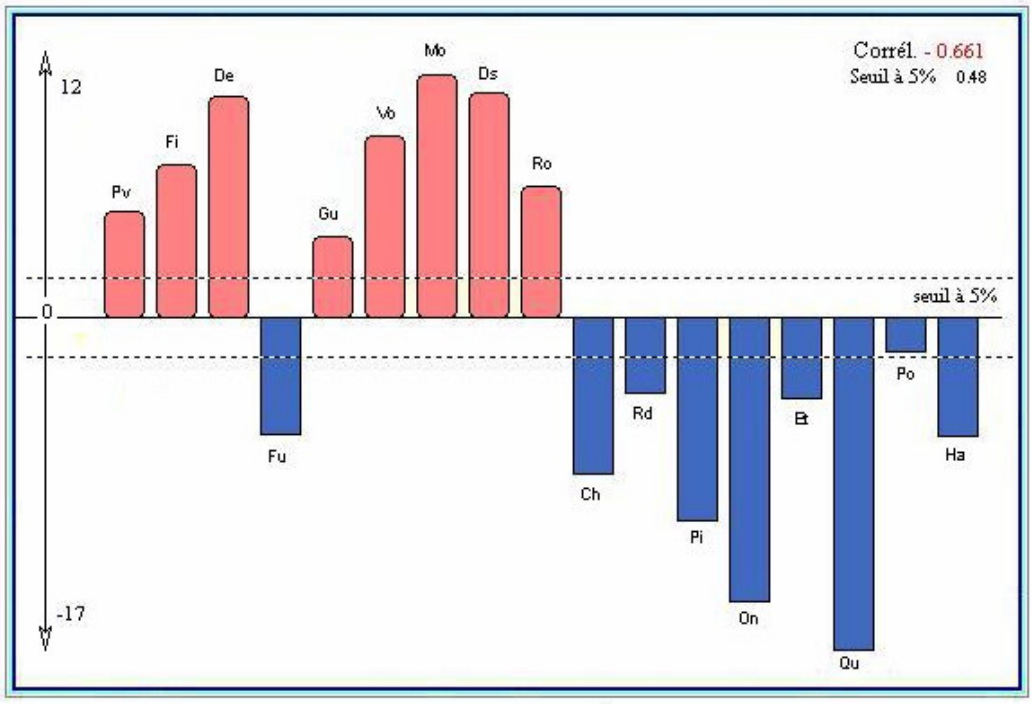


Figure n°73 : La distribution relative de coordonnants dans le corpus E (écarts réduits).

Le coefficient de corrélation est cette fois-ci encore plus significatif que celui obtenu en analysant le corpus en entier, et souligne encore mieux, avec l'exception du roman *Le livre des fuites*, le changement de l'écriture et l'abandon de la

coordination au niveau du roman *Le chercheur d'or*. Cette rupture ne correspond pas à la rupture que nous avons si souvent pu constater (notamment dans l'étude de la structure du vocabulaire) qui intervient vers *Voyages de l'autre côté* ou vers *Mondo et autres histoires*, et il semble que la prédilection de Le Clézio pour la coordination résiste plus longtemps que d'autres facteurs stylistiques au changement important de l'écriture. Ce phénomène se lit également à la comparaison de cet histogramme avec celui de la distribution des subordonnants[273] où la situation globalement est inversée au début et à la fin du corpus, mais avec une intersection au milieu des deux histogrammes, autour de *Mondo et autres histoires*, *Désert* et *La ronde et autres faits divers*, avec des valeurs excédentaires de coordonnants aussi bien que de subordonnants.

Si nous regardons à l'intérieur de la catégorie des coordonnants, il n'y a pas, comme dans la catégorie des subordonnants, d'oppositions intéressantes à constater. L'analyse factorielle des coordonnants les plus fréquents[274] (*car*, *et*, *ni*, *mais*, *donc* et *ou*) ne fait que souligner les déficits importants de *La quarantaine* d'un côté du graphique et pour d'autres raisons d'*Onitsha* à l'autre extrême du tableau, en les écartant des autres ouvrages qui se trouvent agglutinés au milieu du graphique avec les coordonnants :

[273] Cf. figure n°68, p. 269.

[274] Nous avons écarté *or* de cette analyse à cause de sa polysémie. Il est, en effet, beaucoup plus fréquemment question du métal précieux que du coordonnant dans l'œuvre de Le Clézio.

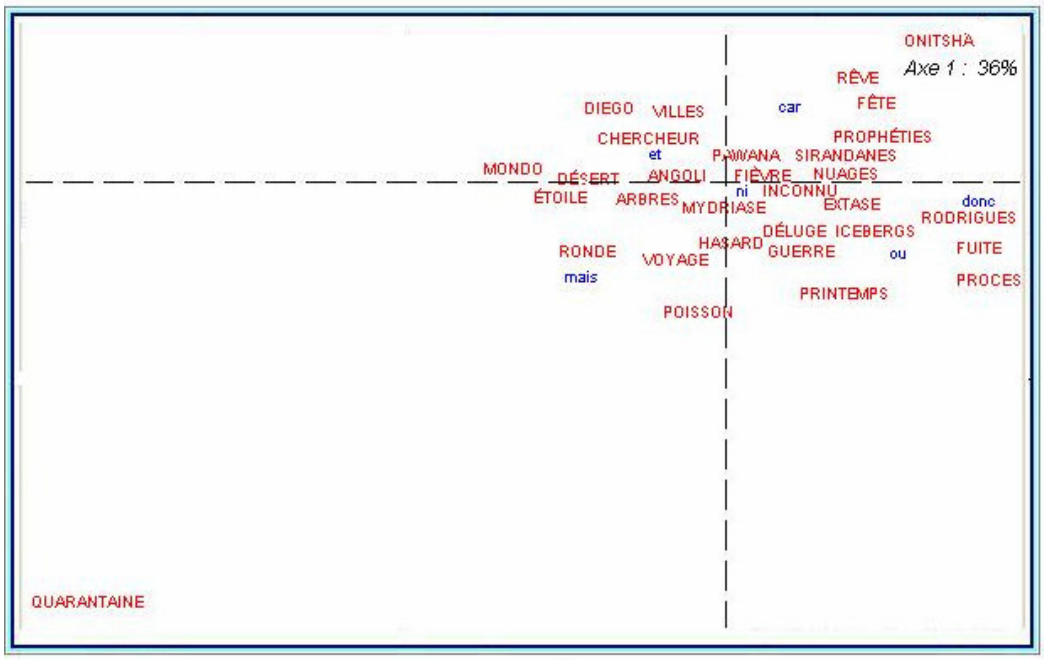


Figure n°74 : Analyse factorielle de la distribution des coordonnants les plus fréquents.

Nous avons déjà dans notre travail observé que le roman de Le Clézio, surtout au début de l'œuvre, contient une forte proportion de phrases-listes, de phrases-inventaires, de phrases qui procèdent par accumulation où la phrase énumérative semble être la phrase par excellence[275] :

> "La foule inconnue prenait des habitudes, des noms : ils s'appelaient Monsieur Cordier, Monsieur Gioffret, Madame Duez, Madame Lemploy, Monsieur José Martin, Madame et Monsieur André Vignaux, Elizabeth, Antoinette, Dick Flanders, Jo, Evelyne, Nicole Nolon."

Dans ce type de phrase, il convient de rappeler l'importance de l'étude de la virgule car ces phrases accumulent les juxtapositions tout en imposant presque toujours – mais pas nécessairement (comme dans l'exemple cité) – un seul coordonnant, à la fin de l'énumération.

[275] *La guerre*, p. 24.

Ces longs réseaux de juxtapositions qui frappent le lecteur ne sont pas nécessairement représentatifs du style de Le Clézio dans l'ensemble de son œuvre. Les immenses réseaux qu'on rencontre parfois dans ses premiers livres traduisent simplement la volonté de l'auteur d'exploiter au maximum l'un des moyens offerts par la langue pour découvrir et appréhender le monde ; les autres structures phrastiques n'en sont pas moins importantes pour la saisie du trait stylistique général qui se définit par l'abondance de chaînes de coordination, quelle que soit la longueur.

Pour mieux comprendre le fonctionnement de ces différents réseaux et afin de mieux pouvoir cerner les procédés techniques qui sont des facteurs déterminants pour la caractérisation du style de notre écrivain, il convient aussi d'étudier la relation entre les deux facteurs d'extension que sont la coordination et la subordination.

3.7.5. La relation entre la coordination et la subordination.

La figure ci-dessous illustre la relation entre les coordonnants et subordonnants de notre corpus, en prenant en considération les fréquences pondérées :

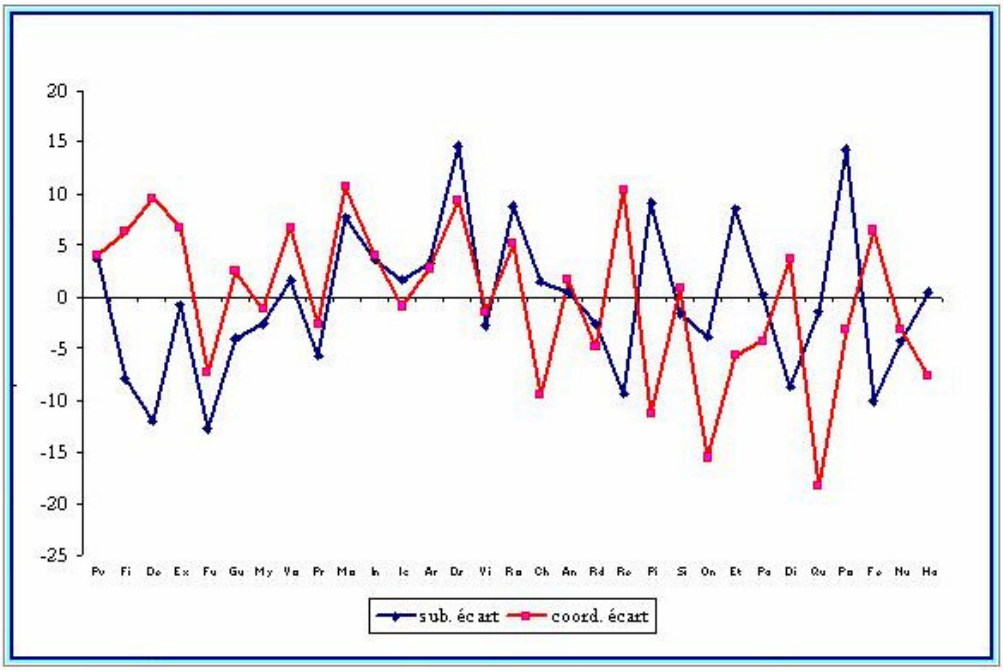


Figure n°75 : Comparaison de la distribution des coordonnants et des subordonnants dans le corpus B (fréquences pondérées).

Le graphique permet d'observer des divergences relativement importantes entre les deux entités dans le corpus. La relation entre les courbes permet la distinction de trois périodes dans l'œuvre, suivant le modèle que nous avons pu observer auparavant. La première période (les quatre premières œuvres) est caractérisée par la divergence des deux courbes, la coordination excédentaire et la subordination déficitaire. L'exemple ci-dessous, tiré du roman *Le déluge* où la divergence est

particulièrement importante, illustre le choix de la coordination au détriment d'une phrase à subordination [276] :

> "Impossible de ne pas succomber un peu au désespoir, et l'armée du souvenir, etc., en profita, lâcha sur nous ces sensations maudites de l'enfance, celles qu'on ne retrouvera jamais plus, ces douceurs, ces paresses, ces intimités d'avant, ces dessins si simples et si goûtés, ces cachettes bien creuses et tièdes, ces trous d'air où le soleil et la pluie étaient mélangés, ces lieux pleins d'objets rouge et or, de délicats animaux, d'anémones de mer et de bernicles, ces perceptions fragiles et muettes, ces odeurs et ces touchers liquides, ces cailloux blancs et carrés, ces univers en forme de dictionnaire, vous savez, ce qu'on appelle les flaques d'eau ; et tout cela revient lentement, essayant de percer la surface de l'être vivant, mais en vain ; et l'on sait que le tourbillon qui avait paru jaillir du néant, était la suprême moquerie, le cri du singe ou du perroquet."

A partir du roman *Le livre des fuites* nous constatons la convergence des courbes et des valeurs globalement excédentaires. La troisième période est marquée par des divergences importantes, les points sont pratiquement tous symétriquement opposés autour de l'axe et les deux courbes qui se croisent et s'entrecroisent ; lorsque la subordination est excédentaire c'est au détriment de la coordination et vice versa. En ce qui concerne le corpus romanesque nous avons pu constater que la subordination est excédentaire tandis que la coordination déficitaire.

L'importance de la coordination par rapport à la subordination est surtout caractéristique de la première période de l'écriture leclézienne et domine les quatre premiers livres *Le procès-verbal*, *La fièvre*, *Le déluge* et *L'extase matérielle* et à la fin de l'œuvre, elle caractérise aussi *Le rêve mexicain*, *Diego et Frida* et *La fête chantée*.

---

[276] *Le déluge*, p. 23.

Conrad Bureau[277] observe, lors de sa comparaison de la phrase de Proust avec celles de Gide et de Le Clézio (s'appuyant uniquement sur le roman *La guerre*, appartenant aussi à "notre" première période) que l'on rencontre en effet des phrases longues aussi bien chez Le Clézio que chez Proust, mais que l'un des facteurs importants de la longueur relative des phrases de Proust est le phénomène de subordination tandis que chez Le Clézio au contraire l'allongement est lié à la coordination. Il écrit à ce propos[278] :

> "Le sort fait par Le Clézio aux réseaux verticaux, et par Proust aux réseaux de subordination, constitue dans chaque cas un surcodage volontaire stylistiquement pertinent qui détermine à la fois la forme et le contenu du texte."

Ce qui distingue la phrase complexe de Le Clézio de celle de Proust, dans son étude, ce n'est pas seulement le nombre plus ou moins élevé de réseaux de coordination, mais aussi la manière de les utiliser. En général, là où la phrase de Proust s'allonge, elle associe réseaux de subordination et réseaux de coordination, de façon à ce qu'il y ait équilibre entre les deux types de réseaux, alors que chez Le Clézio il y a une préférence marquée pour les phrases où les réseaux de coordination se développent aux dépens de l'autre type de réseau.

Le traitement informatique et l'étude des spécificités externes[279], nous permet d'extraire le vocabulaire spécifique positif et négatif de notre corpus, c'est-à-dire les mots particulièrement suremployés et sous-employés par rapport à la base de référence le corpus *Frantext*, et nous apporte ainsi des indications intéressantes concernant le corpus Le Clézio comparé aux autres auteurs de son époque :

---

[277] C. Bureau (1976) p. 151-152.
[278] *Ibid.*, p.152.
[279] Nous reviendrons plus en détail sur les principes de cette analyse dans le chapitre 5.3.2. "Les spécificités externs".

Le coefficient (valeur de l'écart réduit) de *ou* est de +12,38, de *et* de +21,73 tandis que celui de *car* est de – 18,18 et celui de *mais* de –23,94. La valeur de l'écart réduit de *que* est de –64,88 est très parlante surtout lorsque elle comparée à *quand* ou à *comme* qui ont une valeur de l'écart réduit de +38,20 et de +35,76 respectivement.

L'histogramme de la distribution interne du seul mot *que*, largement déficitaire par rapport au corpus *Frantext*, montre un mouvement assez différent de celui de la catégorie de subordonnants :

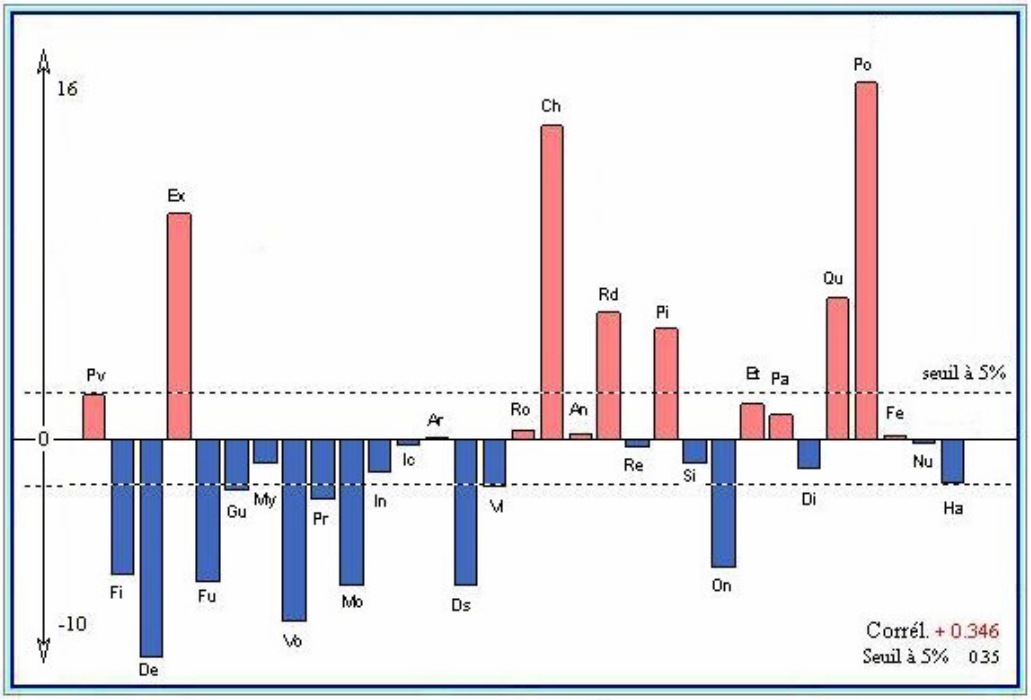


Figure n°76 : La distribution relative de que *dans le corpus A (écarts réduits).*

Nous trouvons dans cet histogramme, à deux exceptions près (*Le procès-verbal* et *L'extase matérielle*), une première période déficitaire suivie par une deuxième période en grande partie excédentaire, toutefois sans que le seuil de signification soit vraiment atteint. Les ouvrages excédentaires sont en réalité les livres riches en dialogues et en oralités comme *Le chercheur d'or*, *La quarantaine*, *Poisson d'or* etc.

Notre étude qui s'appuie sur le corpus quasi complet de l'œuvre de Le Clézio confirme l'analyse de Conrad Bureau et celle de Gunnel Engwall en ce qui concerne la partie de l'œuvre qu'ils ont étudiée. Dans la suite de sa production littéraire et assez tôt dans sa carrière Le Clézio abandonne en grande partie les traits de style particuliers signalés, pour une écriture beaucoup plus conventionnelle où les deux phénomènes d'expansion cohabitent et se complètent dans les mêmes textes à quelques exceptions près.

> "Peut-être plus qu'une opposition, écrit l'écrit Miriam Stendal Boulos[280], ces concepts destinés à caractériser le rythme syntaxique et narratif de Le Clézio expriment surtout une évolution qui est caractéristique de cette écriture. Si maints critiques ont pu remarquer un rythme 'martelé' et la saturation de la phrase dans les premiers romans de Le Clézio, c'est une structure incantatoire qui caractérise les derniers. Tout comme les vécus existentiels sont souvent illustrés par des répartitions volontairement simplifiées des caractéristiques spatiaux chez le Clézio, le rythme illustre le même enjeu : on évoque ainsi la ville, devenue "une machine à extraire l'âme", dans une écriture où dominent le martèlement et la saturation nominale, pour mettre en évidence une situation d'étouffement, et son opposé, un ailleurs originel, où l'homme serait en communion avec la nature et les forces cosmiques, dans un rythme incantatoire."

Ce changement dans l'écriture et plus précisément dans le rythme du récit n'est pas un effet du hasard mais il est le fruit d'une longue réflexion de Le Clézio et une volonté d'aboutir à une maturité, une vérité pure et simple qu'il exprime ainsi déjà dans *L'extase matérielle*[281] *:*

> "Le rythme des villes et des machines est peut-être encore à découvrir ; il est déjà séparé de l'esprit des hommes, et de l'idée de fonction ; il est

[280] M. Stendal Boulos (1999) : p.205.
[281] *L'extase matérielle*, p. 131.

au-dehors. Au-dehors.

Ce n'est pas seulement une affaire collective. C'est aussi la recherche de chacun, isolément, mis en rapport. C'est l'œuvre complète de chaque être vivant, l'œuvre intelligente et instinctive, qui associe, qui éduque, qui ne dompte pas mais libère. C'est peut-être la seule oeuvre vraiment morale, pour la survie, pour la lutte. L'œuvre qui rompt l'indifférence, qui fait naître, vivre et mourir.

Comprendre les rythmes. Parler du monde, aujourd'hui. Renouer avec la terre extatique. Faire de la cacophonie apparente surgir la douce et puissante musique."

3.8. Conclusion.

Nous avons vu, à travers cette étude, que la phrase de Le Clézio - et les unités inférieures qu'elle inclut - fait partie des caractéristiques de son style bien particulier. Jean Onimus s'exprime ainsi à ce propos[282] :

> "Les mots sont une « pâte à modeler » qu'on adapte à ses besoins, à ses rêves ; tantôt des phrases longues, tantôt un style télégraphique ; tantôt une simple énumération, qui est encore le meilleur moyen de comprimer l'information."

En effet, nous avons pu observer, dans les diverses analyses que nous venons d'effectuer, que Le Clézio exploite différentes techniques et différents styles d'expression selon les époques et selon les différents genres littéraires et que les résultats qui en découlent manifestent des variations non négligeables.

Le facteur prédominant de ces divergences semble être celui du genre. Les mots courts dominent l'œuvre romanesque tandis que les mots longs se trouvent dans les ouvrages ethnologiques et dans les essais. C'est également à l'intérieur de ce genre que nous trouvons les phrases les plus longues qui toutefois ne manifestent pas de complexité particulière, préférant la coordination à la subordination.

Les études qui s'appuient uniquement sur le corpus romanesque ont permis d'affiner l'analyse et de souligner la dynamique tripartie de l'œuvre fictionnelle que nous avons déjà pu observer dans notre étude sur la structure lexicale.

Les livres de la première période, souvent qualifiés comme appartenant à "l'école du nouveau roman", sont caractérisés par des mots longs. C'est aussi dans ces

[282] J. Onimus (1994) : p. 167.

livres que l'on trouve essentiellement les signes de ponctuation autres que le point final.

La deuxième période, qui suit la rupture dans l'œuvre de la fin des années 1970, se distingue non seulement par un vocabulaire peu varié – comme nous l'avons vu au chapitre précédent – mais aussi par des mots courts ainsi que des phrases brèves qui toutefois exploitent la subordination aussi bien que la coordination.

C'est à partir du milieu des années 1980 que commence la troisième période, moins distincte que les autres à cause des variations notables dans la distribution des différents paramètres, mais dont la dynamique générale est marquée par une lente progression et une pratique d'écriture qui semble tendre vers la conciliation des deux périodes précédentes, oscillant toujours entre deux extrêmes à cause de la diversité des genres dans cette période.

L'étude de la complexité a surtout mis en relief les particularités du style et du rythme de la première période avec des structures de phrases peu conventionnelles et ses longs réseaux de juxtapositions qui frappent le lecteur[283] :

> "Tout vient de cette bouche, les bruits des moteurs, les rugissements des avions à réaction, les éclats des bombes, les klaxons italiens, les sirènes, les grincements des grues, les raclements des bétonnières, les trains qui passent à 100 kilomètres/heure sur les ponts de fer, les talons des femmes, les cris des oiseaux."

Ces phrases énumératives qui procèdent par accumulation rendent parfois aux livres mêmes un caractère répétitif, à cause du peu de variété syntaxique, de la itération des mêmes structures et de l'abondance de phrases de faible complexité. La technique de Le Clézio d'accumuler les signes, de les entasser, de les additionner à l'intérieur des phrases, qu'il exploite surtout dans la période

[283] *La guerre*, p. 112.

influencée par le "nouveau roman", contribue à lui donner le style et la redondance incantatoire que nous avons déjà évoqués dans notre étude. Conrad Bureau décrit ainsi ce style[284] :

> "Autant Proust est proche d'un Manet ou d'un Turner, autant Le Clézio l'est d'un Vasarelli par exemple qui, par la récurrence d'une même figure – disons le cercle, très utilisé par ce peintre – et l'utilisation de variations très peu marquées dans la couleur de cercles, crée, fabrique pourtant une forme et un contenu nouveaux, là aussi, remarquons-le, avec les moyens les plus simples du domaine pictural."

Pourtant, cette caractéristique de l'écriture de Le Clézio qu'est l'incantation et qui dans la première période de l'œuvre est liée à l'énumération, à la nomination brute des choses, change de caractère au fur et à mesure que l'œuvre progresse. Dans la deuxième et la troisième période de la production l'incantation semble être plutôt liée à la répétition ou à la simplicité d'une structure récurrente.

La redondance et le caractère incantatoire de l'œuvre sont, en effet, un trait de style qui dépend aussi d'un autre facteur : celui de la variété syntaxique qui est fonction du nombre de structures différentes qui apparaissent dans un corpus. L'étude des parties du discours et de la syntaxe nous permettront d'apporter des précisions quant à la structure syntaxique de l'écriture de J.M.G. Le Clézio.

[284] C. Bureau (1976) : p. 174-175.