



HAL
open science

Caractériser le style d'un auteur par des patrons lexicogrammaticaux : une nouvelle approche en stylistique

Dominique Legallois

► **To cite this version:**

Dominique Legallois. Caractériser le style d'un auteur par des patrons lexicogrammaticaux : une nouvelle approche en stylistique. 3e colloque international de l'Association Internationale de Stylistique : Méthodes stylistiques. Unités et paliers de pertinence ?, Mar 2015, Lyon, France. pp.135-150. hal-03446737

HAL Id: hal-03446737

<https://hal.science/hal-03446737>

Submitted on 24 Nov 2021

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

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

Caractériser le style d'un auteur par des patrons lexicogrammaticaux : une nouvelle approche en stylistique.

Dominique Legallois, Clesthia, Université de la Sorbonne Nouvelle, Paris 3

1. Situation

La détection automatisée des spécificités linguistiques d'auteurs, ou de genres, s'est développée depuis une quinzaine d'années. Qu'ils le revendiquent ou non, les travaux entreprenant une telle tâche s'inscrivent en partie ou totalement dans une stylistique : il s'agit bien de mettre en évidence, puis d'interpréter, des traits de langue statistiquement sur-représentés dans les textes, souvent en comparaison avec d'autres textes. Ces traits de langue constituent alors des singularités langagières, propres à un auteur, à un genre. On appellera « descripteurs » les unités linguistiques mobilisées dans ce type d'analyse ; des corrélations, visant à évaluer la co-distribution de descripteurs, peuvent alors être calculées. Donnons un exemple très simple : le premier tableau ci-dessous indique un certain nombre de descripteurs sur-employés chez Balzac si on compare 5 romans de Balzac à la production de 11 autres écrivains français du 19^e siècle (voir plus bas la présentation du corpus).

Descripteurs ¹	Indice de spécificité ²	Fréquence chez Balzac	Fréquence chez les 12 auteurs (frq totale)
NC	Inf	110237	1575137
le	Inf	56446	810946
de	Inf	39057	552514
NP	Inf	19089	248954
DETPOSS	Inf	15694	185889
en	Inf	7836	93343
VP	Inf	7499	95070
je	Inf	7148	90951
vous	Inf	5960	66305
qui	Inf	5778	73325
NUM	Inf	5015	58103
dire	Inf	4687	52793
pour	Inf	4208	47959
par	Inf	3671	36163

- Tableau 1 : extrait des descripteurs spécifiques chez Balzac (5 romans)

-
- 1 NC = nom commun ; Np = nom propre ; DETPOSS = Déterminant possessif ; VP = verbe au présent ; NUM : déterminant numéral. Les autres formes sont des lemmes (ou des mots invariables). Par exemple *le* vaut aussi bien pour *le, la, les*. *Dire* vaut pour toutes les formes, conjuguées ou non, du verbe.
 - 2 L'indice de spécificité mesure « l'attraction » du descripteur pour le sous-corpus Balzac. L'indice Inf précise une très forte attraction.

Le deuxième tableau montre quelques descripteurs spécifiques chez Zola :

Descripteurs ³	Indice de spécificité	Fréquence chez Zola	Fréquence chez les 12 auteurs (frq totale)
le	Inf	78527	810946
il	Inf	29929	253052
un	Inf	24891	223150
se	Inf	12809	114444
VIMP	Inf	21108	127228
!	Inf	8176	75635
PRES	Inf	5298	41329
on	Inf	4488	38568
INT	Inf	2458	21049
alors	Inf	1040	8126
trop	Inf	715	5578
puis	Inf	1462	9107
très	Inf	1404	6449
ça	Inf	2462	5343

- Tableau 2 : extrait des descripteurs spécifiques chez Zola (5 romans)

On peut calculer, par exemple, quels sont les descripteurs « accompagnant » préférentiellement la forme *ça* chez Zola (dans une fenêtre de 3 formes à gauche et 3 formes à droite de *ça*) - on parle d'attraction du descripteur :

Rang	Fréquence du descripteur dans le voisinage de <i>ça</i>	Indice d'attraction	Descripteur
1	905	21.74159	le
2	493	20.77084	ne
3	471	19.67642	être
4	298	16.16771	pas
5	319	15.94719	que
6	301	15.36449	ce
7	287	14.47471	avoir
8	367	14.16478	il
9	470	14.10885	de
10	220	14.07122	faire
11	196	13.67564	me
12	238	12.73959	et
13	158	11.80000	comme
14	153	11.66299	vous

3 VIMP = verbe à l'imparfait ; PRES= participe présent ; INT = interjection.

15	154	11.64718	dire
----	-----	----------	------

- Tableau 3 : les 15 descripteurs les plus attirés par *ça* dans 5 romans de Zola

On remarque, par exemple, que les formes de la négation (*ne*, *pas*) accompagnent préférentiellement le pronom, ainsi que les verbes *faire* et *dire*.

Les seuls descripteurs mobilisés dans cette tâche de détection sont les formes grammaticales (prépositions, conjonctions, etc.) ou des catégories morpho-syntaxiques (noms communs, infinitifs, adjectifs, etc.) ; les mots lexicaux pleins (à l'exception possible, comme ici, de quelques verbes très fréquents) sont exclus, car ils ne précisent en rien le style d'un auteur (ou quelle que soit la catégorie étudiée, le texte, le genre). Les mots pleins sont évidemment des descripteurs incontournables pour les analyses thématiques⁴, alors que les mots grammaticaux sont des révélateurs de style⁵ (les méthodes d'attribution d'auteur – pour les textes anonymes – n'exploitent généralement que les mots grammaticaux).

Bien que les outils et les méthodes se soient spectaculairement démocratisés, bien que nombre d'analyses aient donné des résultats extrêmement intéressants (par exemple Beauvisage 2001, Malrieu et Rastier 2001), une telle stylistique « outillée » reste encore marginale et surtout, ne semble intéresser que très peu le milieu stylistique « académique » (nulle péjoration de notre part dans ce qualificatif). On peut chercher des raisons dans la complexité des méthodes, la technicité des analyses, les différences « de culture » entre les praticiens des textes, mais on peut également évoquer d'autres types de raisons plus profondes que celles-ci : les analyses statistiques et automatisées portent invariablement sur des descripteurs discrets, c'est-à-dire, des unités sans relations syntagmatiques entre elles. Aussi, la lisibilité contextuelle de ces descripteurs reste floue (le sur-emploi du participe présent chez Zola est-il un phénomène interprétable ?), le retour au texte apparaît assez peu pertinent, puisqu'il s'agit de commenter essentiellement des tableaux d'unités quantifiées, plutôt que des énoncés. Mettre en évidence la spécificité d'un descripteur est bien sûr nécessaire, mais saisir ce descripteur dans son élément fondamental qui est le texte, constitue une démarche relativement abstraite. Là est sans doute la limite de l'approche discrète dont l'accueil par les stylisticiens est plutôt tiède, non sans raison donc.

On veut ici proposer une autre méthode, complémentaire et nullement concurrente de la démarche discrète, qui permette appréhender non plus des unités simples, mais des séquences d'unités, des schémas de phrase, des configurations syntaxiques, ou encore ce que l'on pourrait nommer des « patrons »⁶. Ce type d'unités complexes favorise, comme on le verra à travers l'analyse présentée plus bas, une vision plus précise des caractéristiques stylistiques d'un auteur – sans toutefois prétendre, évidemment, à une quelconque exhaustivité.

L'article s'organise ainsi : la partie prochaine présente le corpus littéraire considéré, ainsi que la méthode employée. La partie qui suivra expose quelques patrons parmi les plus saillants chez quatre auteurs, de façon nécessairement partielle : la méthode des patrons lexico-grammaticaux engendre de nombreux résultats, dont il convient de sélectionner les plus intéressants. La troisième partie discute des applications possibles de la méthode, et des améliorations dont elle pourrait bénéficier.

Il est important de signaler que ce travail ne constitue pas à proprement parler une analyse stylistique : il se situe en deçà d'une telle analyse, au niveau où, dans une démarche (semi) automatisée, le chercheur identifie des patrons statistiquement spécifiques avant de sélectionner ceux qui feront l'objet d'une étude stylistique précise.

4 Cf. par exemple l'étude d'E. Brunet (1985) sur le vocabulaire de Zola.

5 Et même plus que cela pour le psychologue américain J.W. Pennebaker (2011), pour qui ce genre de mots révèle la personnalité des individus. La démonstration de l'auteur est cependant très discutable.

6 Sur la notion de « patrons stylistiques », cf Philippe 2008.

2- Corpus et Méthode

Le corpus est composé de 12 écrivains (19^e siècle) dont 5 romans ont été choisis :

Balzac (*La cousine bette, Les illusions perdues, Le lys dans la vallée, Le colonel Chabert, Le père Goriot*), Dumas (*Les trois mousquetaires, Vingt ans après, les quarante-cinq, La reine Margot, Le comte de Monte-Cristo*), Flaubert (*Madame Bovary, Les trois contes, l'éducation sentimentale, Bouvard et Pécuchet, Salammbô*), Gaboriau (*Le crime d'Orcival, Le dossier 113, Monsieur Lecoq, La Clique dorée, l'affaire Lerouge*), Hugo (*Notre Dame de Paris, Les Misérables, Les travailleurs de la mer, L'homme qui rit, Quatre-vingt-treize*), Huysmans (*Les Sœurs Vatard, À rebours, En route, Là-bas, La Cathédrale*), Maupassant (*Bel Ami, Fort comme la mort, Pierre et Jean, Mont Oriol, Une vie*), Mirbeau (*Journal d'une femme de chambre, Le Jardin des supplices, Le calvaire, La maréchale, L'abbé Jules*), Sand (*La petite Fadette, la mare au diable, Indiana, Consuelo, François le Champi*), Stendhal (*Le Rouge et le Noir, Armance, la Chartreuse de Parme, Lucien Leuwen, Le Vert et le Rose*), Verne (*Voyage au centre de la terre, 20000 lieues sous les mers, Deux ans de vacances, cinq semaines en ballon, De la terre à la lune*), Zola (*Nana, L'Oeuvre, l'Assommoir, la Terre, Germinal*).

La méthode, parfois nommée *méthode des motifs*, est sommairement présentée ici (pour plus de détails, voir Charnois, Legallois et Poibeau à par.) : le corpus est lemmatisé et catégorisé de manière à ce que tout le lexique nominal soit regroupé sous l'étiquette **NC**, tous les verbes (hormis les plus fréquents dont on conserve le lemme) sont regroupés sous l'étiquette **V** (mais on précise le tiroir verbal : **VP, VPS, VIMP**, etc.) – sauf les infinitifs et les participes qui possèdent leur propre étiquette. De même, le lexique adjectival se réduit à **ADJ**, ainsi qu'une grande partie des adverbes (**ADV**). En revanche, pour les autres catégories de mots (les « grammaticaux »), on conserve la forme invariable, ou on réduit la forme variable à un lemme (*il* vaut pour *elle, elles, il, ils*). On obtient ainsi une annotation « multi-niveaux » : niveau spécifique de la forme invariable, niveau intermédiaire du lemme et niveau supérieur de la catégorie morpho-syntaxique. La dernière phrase de *Madame Bovary* est ainsi annotée : *il venir de INF le NC de NC* (*Il vient de recevoir la croix d'honneur*). À partir de cette annotation, on procède à l'identification dans le corpus des segments répétés d'une longueur de 5 à 8 éléments (les ponctuations sont conservés, car considérés comme des descripteurs). Par exemple, pour *Les Trois Mousquetaires*, les 10 segments les plus fréquents (sur un ensemble de 5563 avec une fréquence minimale de 5) sont :

Fréquence du segment dans <i>Les Trois Mousquetaires</i>	Segments
1305	le NC de le NC
482	le NC de DETPOSS NC
360	à le NC de le
327	NC de le NC de
325	à le NC de le NC
221	le NC à le NC
211	le NC de le NC de
207	NC à le NC de
204	le NC de un NC
193	NC de le NC et

- Tableau 4 : extrait des segments les plus fréquents dans les Trois Mousquetaires

Ces segments ne sont pas nécessairement des patrons : ils ne renvoient pas nécessairement, dans les textes, à des séquences cohérentes (des groupes complets interprétables), ils ne sont pas non plus nécessairement statistiquement caractéristiques de l'œuvre. Il convient de recourir à une procédure classique en lexicométrie : le calcul des spécificités (sur la base de la loi hypergéométrique), qui permet de mettre en évidence les segments sur-représentés chez chacun des 12 auteurs (un segment devant apparaître dans au moins 3 œuvres sur 5 pour chaque auteur). Une fois les spécificités calculées, la méthode automatisée s'arrête, pour laisser place à un examen « manuel » qui consiste à déterminer si un segment est véritablement un patron (ou ce que nous appelons dans d'autres travaux, un *motif*). Pour qu'il le soit, le segment doit présenter une régularité d'ordre lexical (un même paradigme lexical est employé dans le patron), et / ou fonctionnel : le patron possède une fonction sémantique, pragmatique, énonciative, voire une fonction d'organisation textuelle. La partie suivante donne de nombreuses illustrations.

3- Résultats

Dans le cadre d'un article, il est impossible de donner des résultats exhaustifs, ni même de donner une synthèse satisfaisante des patrons pour chacun des 12 auteurs. Aussi, avons-nous décidé de présenter un ensemble illustratif de patrons pour seulement quatre romanciers (Balzac, Hugo, Stendhal, et un auteur de romans populaires, Gaboriau). Les patrons sont présentés en gras.

3.1 Balzac

- , **enfin tous le NC** (+ extension)

Ce premier patron, typique chez Balzac, correspond à une formule « résumante » : le groupe nominal subsume, en même temps qu'il qualifie, l'ensemble des référents précédemment cités :

- 1- Rien ne démontrera mieux la singulière puissance que communiquent les vices, et à laquelle on doit les tours de force qu'accomplissent de temps en temps les ambitieux, les voluptueux, **enfin tous les sujets du diable** (*La cousine Bette*).
- 2- Si tu vas chez tes belles madames, je veux que tu effaces ce monstre de De Marsay, le petit Rastignac, les Ajuda-Pinto, les Maxime de Traille, les Vandenesse, **enfin tous les élégants** (*Les illusions perdues*).
- 3- ...je m'aperçus que la femme, autrefois si imposante par ses sublinités, avait dans l'attitude, dans la voix, dans les manières, dans les regards et les idées, la naïve ignorance d'un enfant, les grâces ingénues, l'avidité de mouvement, l'insouciance profonde de ce qui n'est pas son désir ou lui, **enfin toutes les faiblesses qui recommandent l'enfant à la protection** (*Le lys dans la vallée*).

- **il être un NC** (+ sub. relative)

La phrase existentielle en *être* est préférée à celle en *avoir* pour exprimer une propriété d'un référent générique :

- 4- Il est des femmes qui s'éprennent de la grandeur comme d'autres de la petitesse (*Les illusions perdues*)
- 5- Il est des personnes que nous ensevelissons dans la terre, mais il en est de plus particulièrement chéries qui ont eu notre cœur pour linceul (*Le Lys dans la vallée*)

- **plus ADJ que ne le être**

Ce schéma d'inégalité est un dispositif de comparaison exprimant une relation de supériorité par rapport à un référent

- 6- Le militaire, en temps de guerre, n'est-il pas également réservé à des spectacles encore **plus cruels que ne le sont** les nôtres ? (*La cousine Bette*).
- 7- Je croyais à de pures amitiés, à des fraternités volontaires, **plus certaines que ne le sont** les fraternités imposées (*Le lys dans la vallée*).

- **le ADJ NC de le NC**

Le style de Balzac se remarque par l'antéposition d'adjectifs, notamment dans ce patron :

8- mais le maître et le prote, perdus dans **les absorbants travaux de l'intelligence**, se contentaient des ouvrages que leur donnaient leurs derniers clients (*Les illusions perdues*).

9- Pour ces esprits étroits, comme pour ces jeunes gens insoucians, **la sèche misère du père Goriot** et sa stupide attitude étaient incompatibles avec une fortune et une capacité quelconques (*Le père Goriot*).

10- Dévorée par le chagrin d'ignorer le sort de son mari, de ne pouvoir lui faire partager dans cette oasis parisienne, dans la retraite et le silence, le bien-être dont sa famille allait jouir, elle offrait **la suave majesté des ruines** (*La cousine Bette*)

- **PASS par le NC de** (+ le NC, un NC, NP)

Le passif est également une construction favorisée par Balzac ; on le rencontre dans ce tour sans auxiliaire, avec l'expression de l'agent (assez peu souvent humain), et avec une préférence pour la position frontale (début de phrase) en construction détachée :

11- **Ruinés par la catastrophe de Fontainebleau**, les trois frères Fischer servirent en désespérés dans les corps francs de 1815 (*La cousine Bette*)

12- **Préoccupé par les soins d'une ambition dévorante**, il s'était attaché comme secrétaire un ancien avoué ruiné nommé Delbecq (*Le colonel Chabert*)

13- **Harcelé par les regards des femmes, sollicité par des rondeurs blanches**, par des gorges audacieuses qui l'éblouissaient, il se raccrochait à son manuscrit qu'il serrait pour qu'on ne le lui volât point, l'innocent ! (*Les illusions perdues*)

- **un de ce NC ADJ**

Par ce schéma, le référent est dit appartenir à un ensemble possédant une caractéristique hautement distinctive :

14- Votre fille est une de ces beautés effrayantes pour les maris (*La cousine Bette*).

15- La Maison Vauquer est une de ces monstruosité curieuses (*Le père Goriot*).

16- En grand maître, l'avoué regarda le coup de tête de Lucien comme une de ces chances inespérées qui, dans une partie, achèvent de la décider (*Les illusions perdues*).

- **VPS le plus ADJ NC**

Ce patron manifeste une double caractéristique : l'attribution au plus haut degré d'une propriété, et l'antéposition de l'adjectif (dont la position standard est la place postposée).

17- Ce refus **excita le plus violent scandale** parmi les hommes considérables du parti royaliste venus à ce rendez-vous (*Les illusions perdues*).

18- Je subis alors une conversation folle, pendant laquelle il me **fit les plus ridicules confidences**, se plaignant de sa femme, de ses gens, de ses enfants et de la vie...(*Le lys dans la vallée*).

19- Elle **éprouva la plus vive émotion** de sa vie, elle sentit pour la première fois la joie inondant son cœur (*La cousine Bette*).

3.2 Hugo

- **PASS, PASS, PASS, PASS, (+ etc.) ; , ADJ, ADJ, ADJ, ADJ, (+ etc.) ; VP,VP,VP, VP, (+ etc.)**

Cet ensemble de patrons incarne un trait de style indiscutable chez Hugo : la juxtaposition, qui peut être très spectaculaire, car prenant une ampleur considérable :

20- C'est en faisant la bouche en cœur du côté de ma vieille caboche que tu as taillé, coupé, tourné, viré, traîné, limé, scié, charpenté, inventé, écrabouillé, et fait plus de miracles à toi tout seul que tous les saints de le paradis (*Les travailleurs de la mer*)

21- Gauvain avait devant lui l'impossible devenu réel, visible, palpable, inévitable, inexorable (*Quatre-vingt-treize*).

22- L'émeute est une sorte de trombe de l'atmosphère sociale qui se forme brusquement dans de certaines conditions de température, et qui, dans son tournoiement, monte, court, tonne, arrache, rase, écrase, démolit, déracine, entraînant avec elle les grandes natures et les chétives, l'homme fort et l'esprit faible, le tronc d'arbre et le brin de paille (*Les Misérables*).

- **nous le avoir dit ; on se en souvenir ; ce que l'on pouvoir appeler/nommer**

On le sait, l'intervention du narrateur dans l'œuvre est très présente chez Hugo, que ce soit pour le rappel d'un fait, ou pour une proposition métalinguistique :

23- Javert ne reconnut pas Jean Valjean qui, nous l'avons dit, ne se ressemblait plus à lui-même (*Les Misérables*)

24- Il les avait, on s'en souvient, mis à sécher sur les rochers saillants de la petite Douvre (*Les travailleurs de la mer*)

25- Cela avait été, dans l'ombre inexorable, presque ce qu'on pourrait nommer la rencontre des hypocrisies (*Les travailleurs de la mer*).

Un ensemble de trois constructions impersonnelles est sur-employé chez Hugo ; les deux premières constructions peuvent être qualifiées de locatives :

- **il y avoir dans le NC (+ GN)**

26- Il y a dans le désespoir de la femme on ne sait quoi de faible qui est terrible (*Quatre-vingt-treize*).

27- Il y avait dans la femme le fond d'une brute et dans l'homme l'étoffe d'un gueux (*Les Misérables*)

- **il y avoir de le NC dans**

La construction impersonnelle ici est une « machine » à massifier les noms comptables :

28- il y avait de la crinière dans sa perruque (*Quatre-vingt-treize*)

29- Il y a du songe dans le tonnerre (*Les travailleurs de la mer*)

- **il y avoir le NC**

Ici la construction existentielle sert de dispositif pour rapprocher / contraster deux réalités

30- Il y a les fumées paisibles et il y a les fumées scélérates (*Quatre-vingt-treize*)

31- Il y a la bravoure du prêtre comme il y a la bravoure du colonel de dragons (*Les Misérables*)

32- Dans la nuit il y a l'absolu ; il y a le multiple dans les ténèbres (*L'homme qui rit*).

- **le NC avoir un NC, le NC**

Le dispositif informationnel permet, par ce patron, de focaliser sur la qualification de l'objet

« possédé » :

33- La révolution a un ennemi, le vieux monde, et elle est sans pitié pour lui, de même que le chirurgien a un ennemi, la gangrène, et est sans pitié pour elle (*Quatre-vingt-treize*).

34- Le Passé a un synonyme, l'Ignoré (*L'homme qui rit*)

Hugo est l'écrivain de la métaphore ; les énoncés extraits du corpus et répondant aux patrons attributifs sont innombrables. Quelques exemples :

- **, ce être le NC** (+ expansion)

35- La guerre étrangère, c'est une écorchure qu'on a au coude ; la guerre civile, c'est l'ulcère qui vous mange le foie (*Quatre-vingt-treize*).

36- Dédalus, c'est le soubassement ; Orpheus, c'est la muraille ; Hermès, c'est l'édifice (*Notre Dame de Paris*).

37- Mentir, c'est l'absolu du mal (*Les Misérables*)

- **le NC est un NC** (+ expansion) ; **un NC est un NC** (+ expansion)

38- La jeunesse est un plan incliné (*L'Homme qui rit*)

39- Les conjectures sont un plan incliné où l'on va inutilement trop loin (*L'Homme qui rit*)

40- La mélancolie est un crépuscule (*Les travailleurs de la mer*)

41- Une prison est un habit de pierre (*Quatre-vingt-treize*)

- **le NC VP, le NC VP ; le NC être ADJ, le NC être ADJ**

On a affaire ici à des dispositifs permettant de construire des contrastes :

42- Tant que le tyran respire, la liberté étouffe (*Quatre-vingt-treize*)

43- le savant conjecture, l'ignorant consent et tremble (*Les travailleurs de la mer*)

44- le petit est grand, le grand est petit (*Les Misérables*)

45- le ciel est noir, l'océan est blanc (*L'homme qui rit*)

- **quelque chose comme le NC**

Il s'agit d'un patron exprimant une caractérisation imprécise - l'indétermination étant un trait saillant de l'esthétique d'Hugo :

46- Quelque chose comme la mort était venu là (*L'homme qui rit*)

47- C'était quelque chose comme l'assassin qui se vante (*Les travailleurs de la mer*)

3.3 Gaboriau

Les trois premiers patrons correspondent à des constructions hypothétiques (que le roman policier naissant exploite volontiers) ; deux patrons ont la forme canonique des tournures hypothétique en *si*. Le troisième emprunte un dispositif différent : un infinitif thématique, suivi de *ce être se INF*.

- **si** (+ P), **ce être que il** (+ P)

48- S'il se tait, c'est qu'il n'a rien trouvé de plausible (*Le crime d'Orcival*)

49- Or, pensait Lecoq, s'il accepte cette lutte, c'est qu'il entrevoit quelque chance d'en sortir vainqueur (*Monsieur Lecoq*)

- **si (+P), ce en être faire de NP/ DETPOSS NC**

50- Si Trémourel est jugé, c'en est fait de Laurence (*Le crime d'Orcival*)

51- Si vous faites du bruit, dit-il, si vous donnez l'éveil, c'en est fait de nos espérances (*Le dossier 113*)

- (Inf. + expansion), **ce être se INF**

52- Rester, c'était s'exposer à une explication pénible, à des insultes, à une collision peut-être ... (*L'affaire Lerouge*)

53- Laisser voir ses opinions, c'était se créer sans nécessité ni utilité une situation impossible (*Le dossier 113*)

Les deux patrons suivants sont dédiés à l'introduction du discours direct :

- **, et après un NC (+ expansion):**

54- D'un coup d'œil, le vieux brocanteur avait embrassé ces détails, et après un sourire de remerciement à sa sœur : (*La Clique dorée*)

55- Il se redressa un peu interdit, et après un moment de méditation : (*Monsieur Lecoq*)

- **, et ce être de un NC ADJ que (+ P):**

56- De grosses larmes roulaient le long des joues du bonhomme, et c'est d'une voix étranglée qu'il murmura : (*La Clique dorée*)

57- Il se roidit donc contre son attendrissement, et c'est d'une voix brutale qu'il répondit à sa mère : (*Le dossier 113*).

- **ce être avec un NC (+ expansion) que P ; ce être de le NC le plus ADJ que (+P)**

Le sur-emploi de la construction clivée est patent chez Gaboriau :

58- C'est avec une anxiété visible qu'elle attendait une explication (*Le crime d'Orcival*)

59- Cette inquiétude l'empêcha de faire le plus léger mouvement, et c'est avec des précautions infinies, doucement, qu'il se risqua à ouvrir les yeux (*Le dossier 113*)

60- C'est du ton le plus sérieux qu'il dit : (*Monsieur Lecoq*)

- **plus ADJ que un NC ; être devenir plus ADJ que**

Ces deux patrons expriment une relation comparative de supériorité, non sans s'appuyer sur des clichés :

61- Et quand elle ne le voyait pas, elle le devinait, pour ainsi dire, sentant ses regards peser sur elle, plus lourds qu'une chape de plomb (*La Clique dorée*).

62- Daniel était devenu plus blanc que sa chemise (*La Clique dorée*)

63- La jeune fille était devenue plus rouge qu'une pivoine (*Le dossier 113*)

Le goût des constructions existentielles est partagé avec Balzac :

- **il être de ce NC (+ sub. Relative); il être un NC (+ sub. Relative) ; il ne être pas de N (+ sub. Relative).**

64- Mais il est de ces situations où, volontiers, on accuse l'électricité de se traîner plus lente que les escargots (*L'affaire Lerouge*)

- 65- Il est de ces convoitises qui brisent les plus solides probités (*Le dossier 113*)
- 66- Mais il est des destinées qu'on n'évite pas (*Monsieur Lecoq*)
- 67- Il n'est pas de diamant où on ne trouve une tache (*Le crime d'Orcival*)

- **le plus ADJ et le plus ADJ NC**

Deux superlatifs sont coordonnés :

- 68- Aussi, les plus hardis et les plus habiles malfaiteurs reconnaissent-ils volontiers qu'il est à peu près impossible de s'échapper de cette geôle roulante pendant le trajet (*Monsieur Lecoq*)
- 69- ...capables de faire reculer de peur et de dégoût le plus féroce et le plus lâche forçat (*La Clique dorée*)

- **ne être que trop ADJ**

Ce patron marque emphatiquement l'attribut :

- 70- Daniel n'était que trop convaincu (*La Clique dorée*)
- 71- Le poison n'est que trop réel et je pourrais te le nommer sans le retirer de ta poche (*Le crime d'Orcival*)

Enfin, relève du cliché tant il est répandu chez l'auteur :

- **ne pouvoir INF un NC de NC**

- 72- M. Patrigent ne put dissimuler un sourire de satisfaction (*Le dossier 113*)
- 73- Hector devint livide et ne put retenir un cri d'horreur (*Le crime d'Orcival*).

3.4 Stendhal

Un grand nombre de patrons caractéristiques de Stendhal sont relatifs à l'expression du haut degré d'une propriété ou d'un ensemble de propriétés :

- **à un NC aussi ADJ**

- 74- On était étonné à la cour de Parme de ne connaître aucune intrigue **à une femme aussi remarquable** par sa beauté et l'élévation de son esprit (*La Chartreuse de Parme*)
- 75- Après de longues incertitudes, qui eussent pu paraître à un observateur superficiel l'effet de la haine la plus décidée, tant les sentiments qu'une femme se doit à elle-même avaient de peine à céder même **à une volonté aussi ferme**, Mathilde finit par être pour lui une maîtresse aimable (*Le Rouge et le Noir*)

- **tout ce que il y avoir de plus ADJ**

- 76- il fut examiné par tout ce qu'il y a de plus pur, soit pour la naissance, soit du côté des bons principes (*Lucien Leuwen*)
- 77- Pendant cette singulière lecture sa tête et son cœur montés au niveau de tout ce qu'il y a de plus grand, travaillaient à son insu (*Le Rouge et le Noir*)

- **de le NC le plus ADJ ; de le plus ADJ NC ; le NC le plus ADJ et le plus ADJ ; le plus ADJ de NP ; le plus ADJ NC de le NC**

- 78- La perte de ma réputation serait d'autant plus rapide, que j'ai vécu dans l'intimité de plusieurs des femmes les plus accréditées de Paris (*Armance*)

- 79- Après le retour des Français, la comtesse était devenue l'une des femmes les plus brillantes de la cour du prince Eugène, vice-roi d'Italie (*La Chartreuse de Parme*)
- 80- Celui-ci était ravi : il captivait l'attention de la plus jolie fille de Koenigsberg, et qui passait pour fort dédaigneuse (*Le Rose et le Vert*)
- 81- Elle se sentait transportée par la pitié la plus tendre et la plus généreuse (*Armance*)
- 82- Mon père va faire un ministère, il a le plus beau rôle à la Chambre, la femme la plus brillante de Paris semble céder à ma prétendue passion (*Lucien Leuwen*)
- 83- La Raversi, quoique régulièrement laide, avait pour amant le comte Balbi, le plus joli homme de la cour (La Chartreuse de Parme)

- **DETPOSS NC être de INF ; être de ne pas INF**

La construction spécifique est présente dans le discours direct uniquement :

- 84- Mon projet était de parler à Mme Dérien et de ne me présenter chez vous que si l'on me disait que vous ne dormiez pas encore (*Armance*)
- 85- Quant à moi, monsieur, ceci n'est pas une action de guerre et mon rôle est de rester en seconde, et même en troisième ligne (*Lucien Leuwen*)
- 86- De 1806 à 1814, l'Angleterre n'a eu qu'un tort, dit-il, c'est de ne pas agir directement et personnellement sur Napoléon (*Le Rouge et le Noir*)

- **faire un ADJ contraste avec**

Ce patron constitue une sorte de tic d'écriture chez Stendhal, tant il est massif dans les descriptions :

- 87- Sa figure sévère et mécontente fit un étrange contraste avec la douce joie que sa présence chassait (*Le Rouge et le Noir*).
- 88- ...reprit l'abbé d'un sang-froid parfait qui faisait un beau contraste avec l'air passionné qui brillait dans les yeux de son jeune interlocuteur (*Le Rose et le Vert*).

Ces quelques patrons rapidement présentés illustrent, pensons-nous, la pertinence de la méthode : ils émergent d'une analyse automatisée, non supervisée par le chercheur. On ne peut donc nier leur caractère objectif, même si, devant la liste des patrons spécifiques, le chercheur peut toujours opérer un choix : par exemple, s'intéresser à tel ou tel patron, ou à tel ensemble de patrons. Il serait sans doute erroné de considérer chacun de ces patrons comme des stylèmes. Ce qui fait « style », c'est bien plutôt l'ensemble de ces schémas lexico-grammaticaux, dont beaucoup, appréhendés individuellement, ne sont pas interprétables et relèvent certainement de l'idiolecte plutôt que du style.

Voyons maintenant comment améliorer l'efficacité de la méthode.

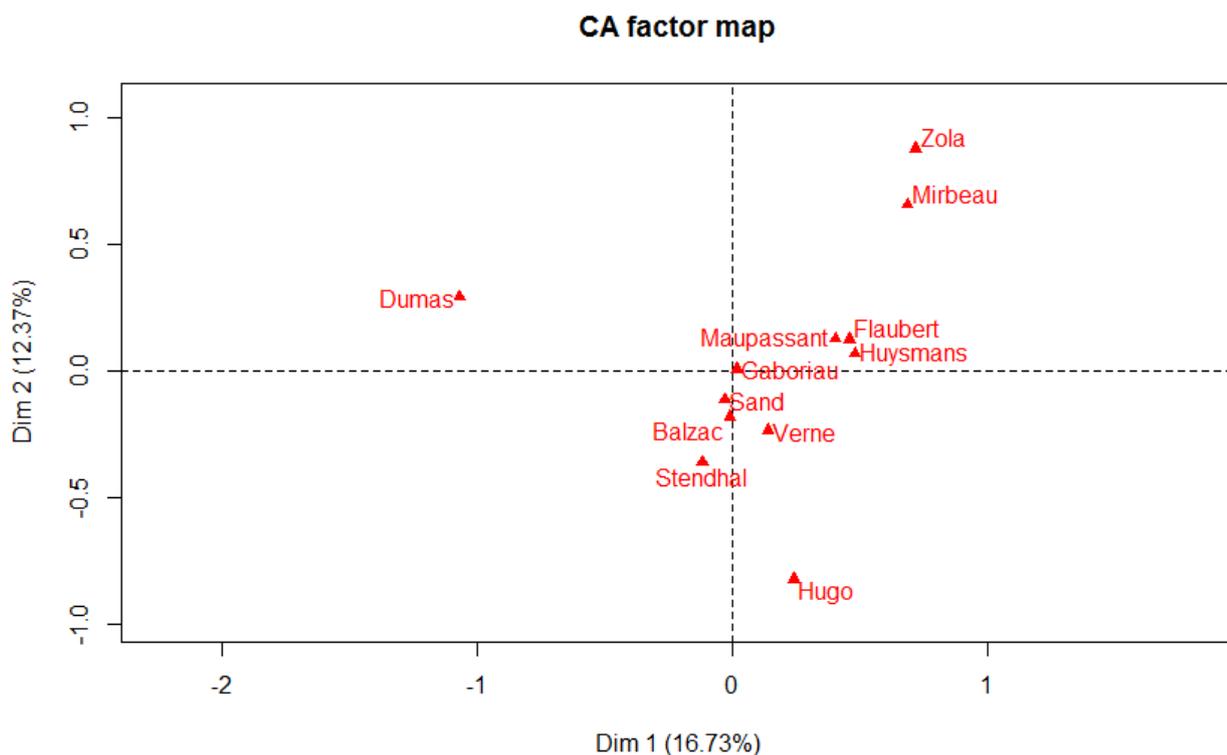
4. Discussion

Les résultats présentés dans la partie précédente sont « bruts » : ce ne sont pas des données mises en perspective dans une problématique (stylistique) particulière. Or, si on veut que la méthode des patrons constitue un véritable apport à la connaissance des styles, il convient d'énumérer les cas où elle permettrait d'établir un véritable dispositif d'observation et d'analyse pour le spécialiste du texte littéraire. Nous donnons ici des pistes, en organisant la réflexion autour de deux points : comment rendre plus informatifs les patrons eux-mêmes ? Quels types d'analyse la méthode permet-elle ?

Il est évident que l'annotation pourrait être enrichie, notamment par des propriétés sémantiques. Les verbes de parole et de perception sont déjà marqués actuellement dans le

dispositif, mais on pourrait considérer d'autres catégories, comme celles des verbes psychologiques (*émouvoir, bouleverser, effrayer*, etc.) ou causatifs (*provoquer, causer, déclencher, permettre*, etc.). De même, les noms d'action, les déverbaux ou encore les noms d'événement pourraient être spécifiquement marqués, à condition d'avoir à disposition des listes fiables. Cet enrichissement permettrait sûrement d'identifier d'autres types de spécificités que celles illustrées ci-dessus. Par ailleurs, manquent encore à la méthode des indications sur les coordonnées textuelles des patrons : certains patrons ont-ils tendance à apparaître en début / fin de roman, chapitre, paragraphe, phrase ? Il s'agit d'une question de « colligation textuelle » (Hoey 2005), encore difficile techniquement à appréhender, mais intéressante, car on peut supposer qu'il existe une relation entre certaines formes et les phases narratives du roman. Par ailleurs, devront être calculées également, les corrélations entre patrons : tel patron a-t-il tendance à être employé dans le même contexte (chapitre, paragraphe) de tel(s) autre(s) patron(s) ? Ce calcul permettrait d'appréhender non plus des unités, mais des ensembles d'unités, donnant lieu, peut-être, à des configurations significatives.

L'application de la méthode est une question essentielle : on peut en effet améliorer l'aspect technique, mais c'est surtout les applications qui peuvent contribuer à sa diffusion. Celles-ci peuvent se décliner de diverses manières. L'analyse à laquelle on pense au vu des résultats présentés plus haut, est la comparaison systématique de styles d'auteurs. L'analyse factorielle ci-dessous a été opérée sur les patrons, pour chaque auteur :



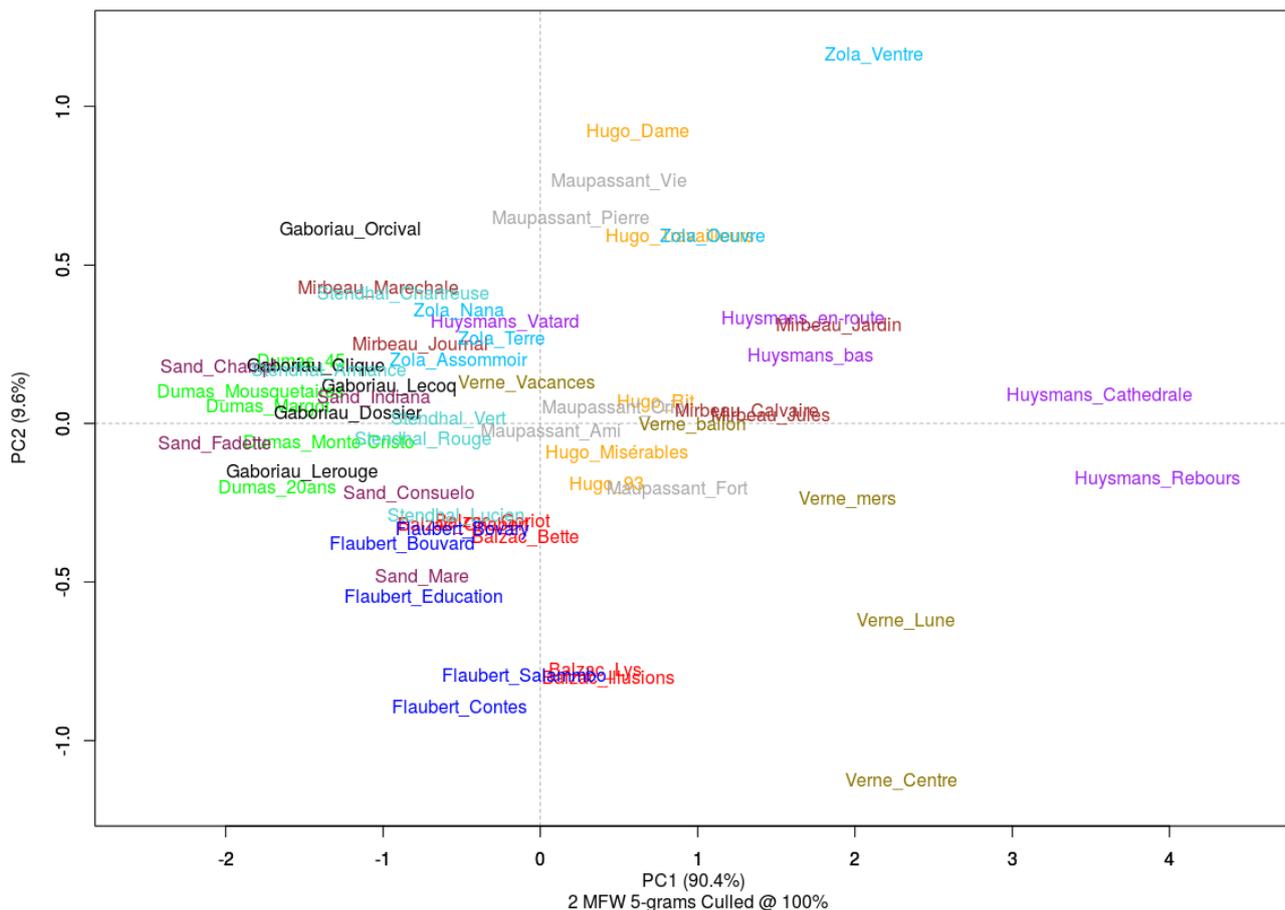
-Figure 1 : analyse factorielle sur les motifs

elle laisse voir une répartition assez cohérente avec la chronologie : dans le plan gauche, les auteurs de la première moitié du 19^e siècle (Dumas, Balzac, Sand, Stendhal), dans le plan de droite, ceux de la 2^e moitié. Les auteurs naturalistes sont regroupés en haut à droite. Dumas est isolé des autres auteurs. La proximité temporelle a ici un poids certain sur la proximité stylistique. D'autres modes d'investigation devront être proposés pour mieux apprécier les différences et les rapprochements ; un de ces modes serait d'examiner si les patrons spécifiques à plusieurs auteurs sont employés de la

même façon par ces écrivains. Dans l'optique d'une stylistique auctoriale, deux autres applications paraissent envisageables :

- étudier l'évolution et la variation stylistique d'un écrivain (par exemple, les romans de jeunesse d'Hugo et ceux de la maturité) ;
- étudier des systèmes sémantiques ou énonciatifs (Monte 2002), par exemple celui de la comparaison, de la manière, du dialogisme, à travers les différents patrons.

Une autre stratégie serait de procéder non pas à une stylistique d'auteurs, mais à une stylistique des textes : autrement dit, comparer les patrons selon les 60 textes. À nouveau, une analyse factorielle a été opérée, donc, sur l'ensemble du corpus :



- Figure 2 : analyse factorielle opérée sur les romans

On constate que la répartition des textes n'est que relativement homogène (les textes d'un auteur ne sont pas systématiquement tous regroupés ensemble). Mirbeau, par exemple, est « dispersé » ; Flaubert, en revanche, est plutôt « cohérent ». Il conviendrait d'examiner pourquoi et comment il y a une telle fluctuation stylistique chez certains auteurs.

Enfin, une autre application est possible, sans doute peu pertinente avec le corpus examiné ici : peut-on mettre en évidence des caractéristiques de sous-genres ? Par exemple, existe-t-il des patrons caractéristiques du roman populaire (par rapport au roman « sérieux ») ? Quels sont les patrons exploités par cette littérature ? Un travail, complétant celui de Magri (2009), sur les différences entre récits fictifs et récits (factuels) de voyage chez les mêmes auteurs, a ainsi été entrepris (Diwersy et Legallois à par.). Bien d'autres sont possibles, dans l'optique d'une stylistique des genres que Larthomas (1964) appelait de ses vœux, mais qui, faute d'outil, a été difficile à

mener jusqu'à maintenant.

On le voit, les applications de la méthode des patrons lexico-grammaticaux sont nombreuses et variées. Une stylistique plus « configurationnelle », plus « globale », pourrait ainsi être systématisée, qui permettrait d'enrichir les études menées dans le domaine.

Références :

BEAUVISAGE Th. (2001), Exploiter des données morphosyntaxiques pour l'étude statistique des genres - Application au roman policier. *Traitement Automatique des Langues*, 43, 45-65.

BRUNET É. (1985), *Le Vocabulaire de Zola. Avec l'index complet et synoptique des 'Rougon-Macquart' d'après les données de l'Institut National de la Langue Française*. Genève-Paris, Slatkine, 1985, 3 volumes

CHARNOIS Th., LEGALLOIS D., POIBEAU Th. (à par.) « *Nicole se tourna vers Franck et posa sa paume ouverte sur son torse, un sourire béat aux lèvres*. La méthode des « motifs » pour repérer clichés et formules dans les romans sentimentaux ». *Lidil*

DIWERSY S. et LEGALLOIS D. (à par.), L'apport de la méthode des motifs aux analyses phraséologiques en Discours. *Colloque international "Approches théoriques et empiriques en phraséologie"*, Tübingen, Stauffenburg.

HOEY M. (2005), *Lexical Priming. A new theory of words and language*. London: Routledge

LARTHOMAS P. (1964). « La notion de genre littéraire en stylistique ». *Le français moderne*, T. XXXII, p. 182-193.

MAGRI V, (2009), *Le Voyage à pas comptés. Le Récit de voyage : mesures, langue et genre de discours*, Paris, H. Champion « Lettres numériques ».

MALRIEU D. et RASTIER F. (2001), Genres et variations morphosyntaxiques. *Traitement Automatique des langues*, vol. 42 (2), 548-577.

MONTE M. (2002) *Mesures et passages. Une approche énonciative de l'œuvre poétique de Philippe Jaccottet*, Paris, Champion, coll. Babeliana.

PHILIPPE G. (2008), « Registres, appareil formels et patrons ». In: Gaudin-Bordes L., Jaubert A. (eds.) *Les Registres : enjeux stylistiques et visées pragmatiques : hommage à Anna Jaubert*. Au coeur des textes 11, Academia-Bruylant, Louvain-la-Neuve, pp. 27-37.

PENNEBAKER J.W. (2011), *The Secret Life of Pronouns: What Our Words Say About Us*. Bloomsbury Press