



ISSN 2260-1651

ISSN en ligne 2260-4987

Étude textométrique du syntagme nominal dans l'œuvre de Jules Verne

José Gregorio Parada

Escuela de Idiomas Modernos. Universidad de Los Andes, Venezuela
jgregorioparada@gmail.com

Reçu le 12-10-2015 / Évalué le 11-11-2015/ Accepté le 15-04-2016

Résumé

Depuis quelques années nous menons une recherche sur l'œuvre de Jules Verne visant notamment l'influence de la Franc-maçonnerie dans son écriture. Nous avons de même entrepris une étude textométrique d'une partie assez représentative de son œuvre, un corpus qui dépasse les six millions de mots et qui nous a permis une approche différente de sa production littéraire. Dans cet article, nous fixons notre regard sur l'étude de certaines catégories grammaticales, notamment celles concernant le syntagme nominal, c'est-à-dire, le substantif, l'article et l'adjectif. A cet égard, nous analysons la distribution et comportement de ces catégories grammaticales dans un ordre chronologique, leur assiduité selon le genre littéraire ainsi que le comportement du genre grammatical, très particulier d'ailleurs chez Verne. Des histogrammes illustrent le développement théorique de cet article.

Mots-clés : Jules Verne, statistiques lexicales, le syntagme nominal

Estudio textométrico del sintagma nominal en la obra de Jules Verne

Resumen

Desde hace algunos llevamos a cabo una investigación sobre la obra de Jules Verne que apunta especialmente el estudio de la influencia de la Francmasonería en su escritura. También emprendimos un estudio textométrico de una parte bastante representativa de su obra, un corpus que sobrepasa los 6 millones de palabras y que nos ha permitido un acercamiento diferente a su obra literaria. En este artículo fijamos nuestra mirada en el estudio de ciertas categorías gramaticales, especialmente las que conciernen al sintagma nominal, es decir, el sustantivo, el artículo y el adjetivo. A este respecto, analizamos la distribución y comportamiento de estas categorías gramaticales en un orden cronológico, su frecuencia según el género literario así como el comportamiento del género gramatical, muy particular por cierto en Verne. Algunos histogramas y cuadros ilustran el desarrollo teórico de este artículo.

Palabras clave: Jules Verne, estadísticas lexicales, el sintagma nominal

Textometric analysis of noun phrase in Jules Verne's work

Abstract

From some years now, we have conducted some research on Jules Verne's work that especially targets the study of the influence of Freemasonry in his writing. We also undertook a textometric analysis of a fairly representative part of his work, a corpus that exceeds 6 million words and that allowed us a different approach to his writing. In this article we fix our gaze on the study of certain grammatical categories, especially those concerning the noun phrase, ie, the article, the adjective and the noun. In this regard, we analyzed the distribution and behavior of these grammatical categories in a chronological order, their frequency according to the literary genre as well as the grammatical gender behavior, very special in Verne's. Some histograms and tables illustrate the theoretical development of this article.

Keywords: Jules Verne, lexical statistics, the noun phrase

Introduction

Dans cet article nous présentons, dans un premier temps, quelques idées sur la constitution du corpus Jules Verne (1828-1905) et quelques notions sur le logiciel qui nous a permis d'entreprendre notre démarche textométrique. Puis, nous présentons le panorama des catégories grammaticales du corpus, notamment leur fréquence et leur distribution. Pour bien comprendre l'usage fait de ces catégories grammaticales, nous signalons les tendances chez d'autres auteurs de l'époque afin de permettre des comparaisons avec notre auteur cible. Finalement, nous faisons une analyse du syntagme nominal dans l'œuvre de Jules Verne, fil conducteur de cette étude, Pour ceci, nous proposons une étude détaillée du comportement des éléments composant le syntagme nominal.

1. Le corpus, les méthodes et outils informatiques utilisés

Les nouveaux outils informatiques d'aujourd'hui nous permettent d'entreprendre des recherches statistiques conduisant à des résultats assez fiables et, surtout, objectifs. Cette discipline statistique à laquelle nous faisons appel c'est la textométrie.

Pendant les trois dernières décennies, les études textométriques, c'est-à-dire des études statistiques du discours, ont acquis un développement inattendu à tel point de permettre dans l'actualité une approche « millimétrique » du texte (Parada, 2011 : 243). Un bon nombre d'auteurs a été réétudié sous cette nouvelle dimension, notamment par le propre concepteur du programme Hyperbase.

Le traitement systématique et organisé de plusieurs millions de mots à la fois est désormais possible. La qualité des résultats est déterminée par des lois mathématiques et impartialement rigoureuses. Ce fait permet donc de faire des analyses beaucoup plus précises du corpus.

Plus de quatre-vingts textes ont été organisés en ordre chronologique. Trois genres cohabitent dans notre corpus : nouvelles, essais et romans, constituant ces derniers un pourcentage très représentatif de l'œuvre vernienne. Le corpus contient 59 fichiers dont 56 représentent des romans, et les trois restants 18 nouvelles et 9 essais. Ce corpus ne contient pas les œuvres dites « remaniées » par son fils Michel, d'autres écrites en collaboration ou celles parues sous sa signature appartenant intégralement à d'autres auteurs. Ce corpus, constitué par les versions intégrales des textes, est présenté dans l'ordre chronologique d'écriture et mis en contraste avec une norme externe, à l'occurrence avec la base de Frantext par rapport à laquelle le logiciel fait les calculs nécessaires avant de le segmenter automatiquement en formes graphiques, comptabiliser les occurrences, le lemmatiser et établir les résultats sous forme de cadres, histogrammes, analyses arborées et analyses factorielles de correspondances (Parada, 2013).

Le logiciel que nous utilisons, Hyperbase, a été conçu par Etienne Brunet du laboratoire Bases, Corpus et Langage (CNRS-Université de Nice Sophia-Antipolis), associé à l'étiqueteur Cordial (Mayaffre, 2004 : 16), et permet quatre traitements de façon intégrale et simultanée du texte brut : mots traités tels qu'ils ont été écrits ; du texte lemmatisé : les mots sont ramenés à leur canon ("vient" = "venir") ; des codes grammaticaux : les mots reviennent à leur catégorie ou fonction grammaticale ("vient" = verbe à la troisième personne du singulier au présent ; et des structures syntaxiques : le discours est ramené à ses enchaînements syntagmatiques ("le ballon s'éleva" = déterminant+nom+verbe...) (Parada, 2013).

Ce logiciel possède deux grandes fonctions : une dite « documentaire » et l'autre statistique. La fonction documentaire nous donne la possibilité de naviguer autour du texte, de classer ses composants en lemmes et codes et de repérer facilement les passages liés par des traits caractéristiques à la manière d'un moteur de recherche. La seconde fonction d'Hyperbase peut nous permettre, entre autres, de connaître le degré d'utilisation d'un trait linguistique, d'avoir accès à la représentation graphique de la distribution des termes et au calcul de la distance entre textes, en passant par l'analyse factorielle de correspondances et par des analyses arborées (Parada, 2013).

2. La distribution des catégories grammaticales

Le programme Cordial propose onze catégories grammaticales principales pour lesquelles nous avons cherché à connaître la distribution en visualisant les résultats par une analyse factorielle de correspondance. Ces catégories sont les suivantes : verbes, substantifs, adjectifs, déterminants, pronoms, numéraux, interjections, prépositions, adverbes, conjonctions, et délimiteurs (signes de ponctuation).

D'après le résumé des calculs de Cordial inscrits dans le tableau N° 1, la catégorie prédominante est le substantif (21%) suivie par le verbe (14%) et le déterminant en même proportion. Les prépositions et les pronoms sont regroupés autour de 11% et 9%. Les adjectifs et les adverbes, 5% et 6% respectivement. Les conjonctions représentent 4% tandis que les numéraux et les interjections figurent en proportion dérisoire. L'intérêt de Verne penche nettement vers le substantif puis vers le verbe. Son discours doit être alors considéré plutôt nominal face à un discours verbal caractérisant Hugo, par exemple, qui ne méprisait pas non plus le substantif. La tendance est variée : Chateaubriand est favorable à la catégorie nominale ; Zola aux verbes, Proust donnait de l'importance aux mots de relation ; Hugo comme Verne montraient un intérêt pour le substantif et le verbe, la même tendance se retrouve chez Giraudoux (Brunet, 1998 : 171-172).

Catégorie	Effectif	Pourcentage	Catégorie	Effectif	Pourcentage
Verbe	931801	14	Interjection	6332	0
Substantif	1391286	21	Préposition	702106	11
Adjectif	313808	5	Adverbe	409854	6
Déterminant	898079	14	Conjonction	272172	4
Pronom	611569	9	Ponctuation	987338	15
Numéral	57697	1			

Tableau N° 1. La distribution des catégories grammaticales

La distribution des catégories mentionnées précédemment n'est pas uniforme dans le corpus Verne. Le constat, fait sur le schéma résultant de l'analyse factorielle des correspondances (AFC), montre l'éventail de textes mis en rapport avec lesdites catégories. Dans la figure N° 1, les déterminants sont attirés naturellement par les substantifs autour de textes où la description s'impose sur l'action : *Essais, Russes, Fourrures, Hélice, Rayon* et *Bourses*. Autrement dit, ce sont des textes dans lesquels le discours nominal prime sur le discours verbal. En effet, peu importe le texte choisi individuellement, l'écart des substantifs se montre excédentaire. Les déterminants, de leur côté, plus caractéristiques au début de la carrière de Verne se replient à côté des substantifs, et se trouvent en nombre excédentaire principalement dans les textes suivants : *Nouvelles 1, Reculons, Hatteras, Grant, Mers, Autour de la lune, Flottante, Russes, Fourrures, Vapeurs, Essais, Robur,*

Dessus, Hélice, Patrie, Phare et Invasion. Par contre, le texte le plus déficitaire en déterminants et en substantifs est *Chemin de France* que nous retrouvons d'ailleurs à l'extrême gauche de la même figure, dans l'axe correspondant au discours verbal. Ceci nous conduit à diviser donc le schéma du premier facteur de cette AFC en relation avec l'axe vertical : sur la gauche le discours verbal ayant comme noyau le verbe qui attire les pronoms et les adverbes ; sur la droite, le discours nominal autour des substantifs qui attire les déterminants, les adjectifs et les numéraux. Cette tendance vers une association naturelle entre le déterminant et le substantif, d'une part, et le pronom et le verbe, d'une autre, relève de contraintes grammaticales.

Du point de vue du genre, les essais et le premier recueil de nouvelles (Nouv_1), comme nous l'avons signalé dans le paragraphe précédent, sont devenus satellites du discours nominal, se regroupant autour du substantif. Ceci n'est pas le cas pour le deuxième recueil de nouvelles qui gagne « plus d'action » comme le montre sa position vers la gauche plus près de l'adverbe et du verbe, et à une distance considérable du substantif. Les romans de la première étape vernienne se replient notamment du côté du substantif. Les autres montrent une préférence pour le côté gauche, ce qui nous fait penser à une volonté soit de réduire les descriptions et de donner la préférence à l'action, soit d'insérer davantage de dialogues. Les romans de la dernière étape trouvent le juste équilibre entre les deux types de discours se situant au milieu de la distribution verticale, à l'exception d'*Hélice* et *Testament*, déficitaires en verbes et très friands de substantifs et d'adjectifs respectivement.

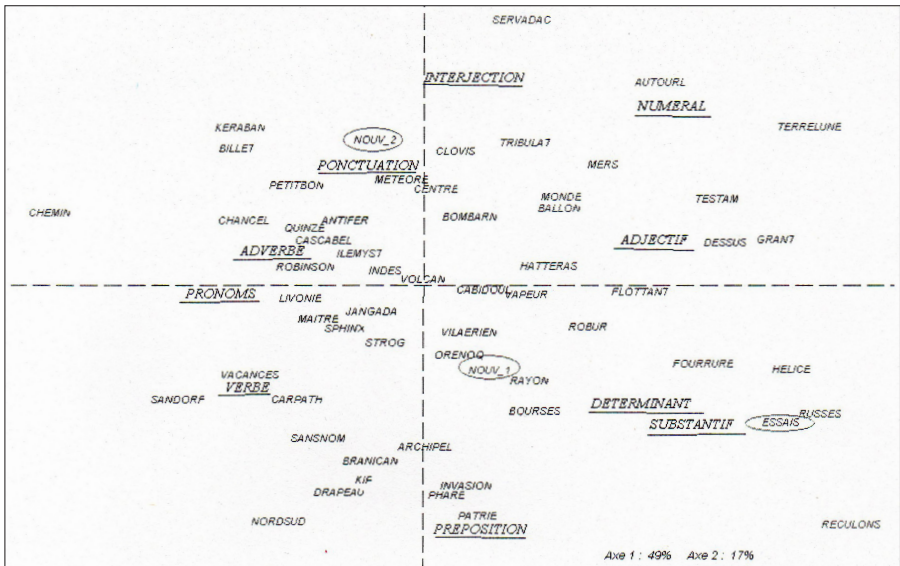


Figure N° 1. Analyse factorielle de la distribution des principales catégories grammaticales. Chaque mot représente un texte, Nouv_1 et 2 évoquent deux recueils de nouvelles. (Méthode Tabet)

Le second facteur met en opposition le haut et le bas du graphique. Le haut qui concerne des textes notamment de la première et deuxième étape de Verne (sur trois étapes au total), se caractérise par une écriture très ornée notamment d'une profusion d'adjectifs, de numéraux, d'interjections et d'adverbes. Descriptions débordantes en adjectifs dans *Grant*, *Ballon*, *Mers* ; calculs exigeant la présence de numéraux (*Autour de la lune*, *De la terre à la lune*) ; nombreuses interjections (*Tribulations*, *Servadac*) ; adverbes en nombre démesuré (*Robinson*, *Ille mystérieuse*, *Cascabel*, *Chancellor*). Le bas de la figure rend compte d'un discours moins orné autour des catégories fondamentales : le verbe et le substantif. Autour du verbe, se situent des romans qui fuient les adjectifs et les substantifs : *Sandorf*, *Vacances*, *Carpathes* et *Sans nom*.

Un nombre non négligeable de textes de la dernière étape se trouve dans le bas du schéma, ce qui nous amène à penser à une simplification du style chez notre auteur vers la fin de sa carrière, due notamment à l'abandon de l'adverbe et de l'adjectif.

3. Le syntagme nominal

L'intérêt de cet article vise notamment trois composants du syntagme nominal : le substantif, l'article et l'adjectif.

Quant aux effectifs, les substantifs atteignent un nombre important (1 391 286) pour dominer vis-à-vis de l'effectif des verbes (931 801). Les proportions de 60% et 40% se maintiennent *grasso modo*, lorsque nous faisons l'addition de catégories qui s'attachent de façon plus ou moins naturelle soit au substantif (adjectif, déterminant et numéral) ; soit au verbe (pronom et adverbe).

3.1. Le substantif

21,12% du corpus est représenté par le substantif dont la distribution n'est pas homogène tout au long de l'œuvre. Voici l'histogramme de sa distribution chronologique.

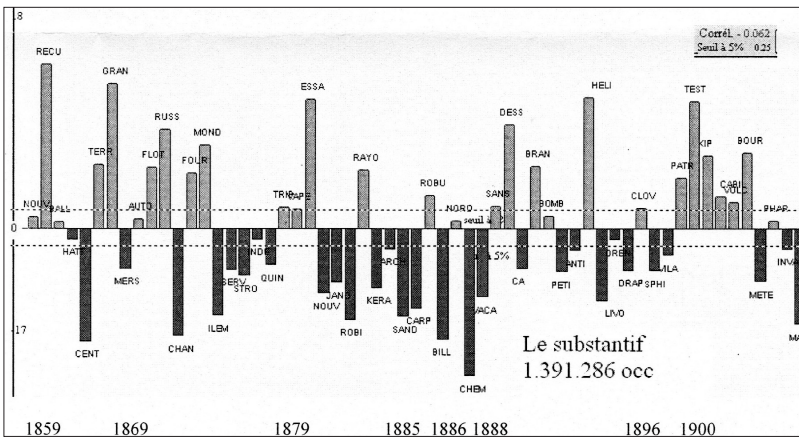


Figure N° 2. La distribution des substantifs. Vers le haut sont signalés les excédents, vers le bas les écarts négatifs.

Deux remarques s'imposent à première vue : les différences générées par le genre et la distribution inégale des substantifs sur la ligne du temps.

Les substantifs ne montrent aucun écart important dans le premier recueil de nouvelles, mais pour le deuxième ils sont déficitaires. Les essais, au contraire, manifestent l'un des écarts positifs les plus significatifs. En effet, d'après E. Brunet (1998 : 304), dans le discours intellectuel et technique il y a une tendance à utiliser les substantifs qui s'adaptent aux définitions abstraites et à une pensée complexe. Pour les romans, la lecture des histogrammes n'est pas simple : un usage important des substantifs est remarquable durant la première période. Cet usage manifeste, par la suite, une sorte de régression jusqu'en 1879 avec *La maison à vapeur*. Des textes sont à signaler pour leur usage important de substantifs : *Reculons*, *Grant*, *Russes* et *Monde*. Après ceci, s'ouvre une parenthèse caractérisée par une utilisation réduite de substantifs jusqu'en 1886 avec *Vacances*. Dans cette parenthèse *Chemin de France* mérite d'être mentionné car il constitue le roman qui atteint l'écart le plus bas, ce qui peut sûrement s'expliquer par le fait que ce texte, entre autres, favorise le dialogue, là où le substantif est déficitaire. Par la suite, les écarts sont plutôt positifs, notamment dans *Dessus*, *Hélice* et *Testament*. Ceci est en contradiction avec la tendance générale selon laquelle la fréquence du substantif tend à diminuer au fur et à mesure, dans le corpus de grande taille.

Dans l'exemple suivant, extrait des *Essais*, la description peut se faire à partir d'une longue liste de substantifs et d'adjectifs accompagnés ou non de verbes :

[...] *Quelles longues traînes, montées sur de petites roulettes de métal qui murmuraient délicieusement sur le sable ! Quels chapeaux, avec lianes enchevêtrées, plantes arborescentes, oiseaux des tropiques, serpents et jaguars en miniature, et dont une forêt du Brésil n'eût donné qu'une idée imparfaite ! Quels chignons, d'un volume si embarrassant et d'un poids si considérable, que ces élégantes étaient forcées de les porter dans une petite hotte en osier, ornée, d'ailleurs, avec un goût irréprochable ! Enfin, quelles polonaises, dont les combinaisons de plis, de rubans et de dentelles, m'eussent semblé moins faciles à reconstituer que la Pologne elle-même ! [...]*

Ces concours ne me déplaisaient pas. Il s'en dégage pour les oreilles et les yeux un utile enseignement. Le fracas strident des machines qui fonctionnent, les hennissements de la vapeur, le bêlement plaintif des moutons parqués dans leur enclos, le caquetage assourdissant des basses-cours, les beuglements de ces grands bœufs qui réclament leur prime, les discours des autorités dont les pompeuses périodes débordent de l'estrade, les applaudissements jetés par la main des lauréats, le doux bruit des baisers que les lèvres officielles déposent

sur les fronts couronnés, les commandements militaires qui retentissent sous les grands arbres, enfin, ce vague murmure qui sort de la foule, tout cela forme un concert bizarre, mais dont j'apprécie vivement le charme. (Une ville idéale, Essais).

À en juger par le nombre de romans concernés, on peut dire que le substantif se montre plutôt déficitaire dans ce genre. Un bon nombre de romans de voyages montre une richesse importante de substantifs en raison de la présentation de nouveaux horizons, à une couleur locale, à l'exotisme qui caractérise les œuvres de Verne. *Grant*, par exemple, explore le 37^e parallèle austral, passant par l'Australie, la Nouvelle Zélande et le Chili, sans dire que le voyage commence en Écosse ; le voyage de *Testament* fait halte partout sur l'espace américain. Par contre, dans *Russes*, c'est le vocabulaire technique de la triangulation géodésique qui accapare la plupart des substantifs ; *Dessus* nous met devant un projet extrêmement complexe : le redressement de l'axe de la Terre !

Une autre caractéristique intéressante chez Verne est la corrélation entre la richesse du vocabulaire et la fréquence manifestée par le substantif. De près, nous constatons que dans les œuvres ayant un vocabulaire étendu, la fréquence des substantifs est majeure (*Hélice*, *Essais*, *Reculons*). Cette remarque n'est peut-être pas indiquée pour *Mers*, roman qui se distingue notamment par sa richesse relative en adjectifs.

Quant à la catégorie des substantifs, grâce à une distribution chronologique, des écarts importants sont à signaler tant pour le masculin que pour le féminin dans la première moitié de la production, le genre masculin se montrant plus affirmé. À partir de *Kériban* (1882) le genre masculin devient moins usité, exception faite de *Robur*, *Dessus*, *Antifer* et *Testament*. Dans la seconde moitié, les textes les plus riches en substantifs féminins sont : *Hélice*, *Patrie* et *Phare*. Mais, en guise de synthèse, nous pouvons dire qu'aucune opposition nette entre l'usage des genres dans l'ensemble du corpus, n'a été relevée, sauf à considérer individuellement les textes : Dans *Patrie* l'écart de substantifs féminins (23,3)¹ est manifestement écrasant par rapport à celui des substantifs masculins (-10,6). *Patrie* reprend, en effet, les thèmes de l'île déserte, de la colonie ; des tempêtes, de l'explosion volcanique, tous caractérisés par le genre féminin. A l'inverse, dans *Antifer*, les écarts se montrent positifs pour le masculin (8,8) et négatifs pour le féminin (-17,1). Les personnages d'*Antifer* sont des hommes et le motif de l'histoire tourne autour d'un trésor, de l'argent, des voyages et des messages codés, autant de mots appartenant au genre masculin.

Le pourcentage de substantifs masculins est de 62 devant 38% de substantifs féminins. Nonobstant, dans les nouvelles, les essais et certains romans, il y a une prédominance des substantifs féminins sur les masculins qui nous fait penser à une forte présence de termes abstraits. Mais, de façon générale, ce sont les nouvelles et les essais qui remportent les écarts positifs de ce type de termes. Des mots tels que « pensée, âme, cœur, esprit, avenir, imagination, amour, génie, bonheur, gloire, idées, intelligence, mémoire, sentiment(s), sagesse, vérité, bonté, désir, espérance et raison » (Magri, 1995 : 102) constituent une liste de termes abstraits que l'on peut facilement trouver dans les textes suivants, tous caractérisés par une prédominance des substantifs féminins sur les masculins : *Nouvelles, Reculons, Centre, Terre, autour, Fourrures, Tribulations, Servadac, Strogoff, Essais, Rayon, Archipel, hélice, Patrie, Phare* et *Maître*. Un texte qui ne rentre pas dans cette liste mais qui présente des écarts significatifs pour un bon nombre des termes cités est *Branican*. L'explication est donnée par Jules Verne lui-même lorsque, en réponse au correspondant du *World* à Paris, Robert Sherard, il explique à propos de Nelly Bly²:

Comme le journaliste lui demandait s'il introduirait Miss Bly dans une histoire, Verne répondit candidement : « Sûrement. Je ne suis pas fameux pour les personnages féminins, mais je vais m'y atteler très bientôt avec un livre intitulé « Lady Franklin » [qui sera en fait Mistress Branican]. Plus tard ; je trouverai certainement un cadre pour cette très jolie image que vous m'avez présentée à Amiens, il n'y a de cela qu'environ soixante-dix jours³». (Lottman, 1996 : 315).

En effet, c'est dans ce roman seulement que Verne réussit à ériger en héros un personnage féminin. « L'exercice a le mérite de définir un héros féminin aux caractéristiques nettement distinctes des héros masculins » (Dekiss, 1999 : 306). Par contre, Paulina Barnett, âme de la colonie dans *Fourrures*, malgré sa diligence, son courage et sa volonté, nous est présentée sous des traits masculins...

Chez Verne, la liste de noms propres masculins prime sur sa correspondante féminine, à l'exception de *Fourrures, Servadac, Rayon, Archipel, Sandorf* et *Chemin*. Quant aux noms communs, la tendance du masculin est toujours écrasante, sauf dans : *Reculons, Centre, Autour de la lune, Hélice, Village aérien, Patrie, Phare* et *Maître*.

Pour le nombre de formes, les substantifs singuliers sont trois fois plus nombreux que les substantifs pluriels et ils prédominent dans tous les textes sauf dans : *De la Terre à la lune, Grant, Mers, Autour, Flottante, Russes, Hélice* et *Invasion*.

3.2. L'article

Sur un total de 546 349 occurrences, 66% correspond à l'article défini masculin, contre 34% pour l'article défini féminin, proportion très opposée à la tendance moyenne manifestée dans *Frantext*, base dans laquelle la préférence pour l'article féminin est de 56% contre 43% pour l'article masculin. Sans doute, une influence notoire chez Verne pour faire varier ostensiblement ces pourcentages, est la forte présence de personnages masculins⁴. Les femmes, nous dit Ghislain De Diesbach, dans l'œuvre vernienne, « occupent une place subalterne ou restent dans les coulisses » (Diesbach, 2000 : 55), l'exemple notoire proposé par cet auteur est la Stilla, dans *Carpathes*, déjà morte lorsque commence le récit.

On a vu à quel point le monde vernien est un monde d'hommes. Il compte peu d'enfants et de vieillards ; surtout des adultes jeunes ou dans la force de l'âge, qui ont une autre caractéristique frappante. Ils sont seuls. Seuls, c'est-à-dire sans femme.

Pour un quart des personnages masculins, l'état civil n'est pas précisé ni déductible des renseignements fournis par l'auteur [...] Au total, plus de 73 personnages masculins dans ce cas, soit 1 sur 6 [...]

Ajoutons les célibataires qui le demeurent - et ils sont nombreux : sur 164 hommes célibataires, seuls 59 se marient. Si les jeunes célibataires (de 18 à 30 ans) se marient majoritairement, la proportion s'inverse ensuite. Sur les 69 qui ont entre 31 et 50 ans, 8 seulement se marient. Et un seul des 13 plus de 50 ans. Le célibat est d'ailleurs parfois la condition requise pour faire partie d'une expédition. Ainsi, les marins qui embarquent sur le bateau du capitaine Hatteras doivent être célibataires (Dusseau, 2005 : 127).

L'article masculin est prépondérant jusqu'à *Servadac* (1874). Après cette année, il manifeste une baisse d'usage au profit de l'article féminin, et en moindre mesure, des indéfinis. Son usage est préférentiel sauf dans les textes suivants : *Reculons*, *Centre*, *Terre*, *Autour*, *Indes*, *Essais*, *Archipel*, *Cascabel*, *Livonie*, *Patrie* et *Phare*. Quant au nombre, même si les articles singuliers sont quatre fois plus nombreux que les pluriels, l'usage montre des écarts positifs assez équilibrés, parfois à la faveur des pluriels, comme dans *Hélice* et *Grant*. Nous apercevons un faible abandon de l'article défini masculin singulier sur la ligne chronologique, tandis que l'abandon des articles indéfinis au singulier est très marqué.

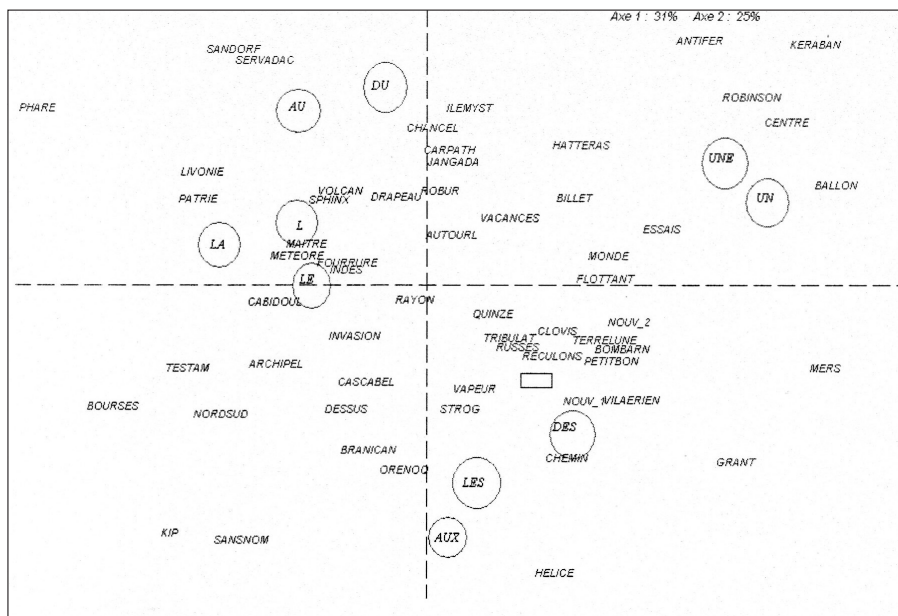


Figure N° 3. Analyse factorielle de la distribution des articles

La figure précédente représente la distribution factorielle de l'article. Le premier facteur montre les textes attirés par l'usage des indéfinis, sur la droite, tandis que sur la gauche se placent ceux attirés par les définis. Les contractés *aux* et *du* ne s'écartent pas trop de l'axe vertical, se montrant plus neutres pour les commentaires⁵. Le second facteur met en relief le nombre. On distingue donc, sur le haut le singulier et, en bas, les textes à fort usage des articles pluriels. L'article le plus proche du centre est naturellement le défini singulier à cause notamment de sa prépondérance dans le corpus. Son correspondant féminin, sur le quadrant supérieur gauche, attire un faible nombre de romans, un peu en correspondance avec sa proportion. Quant à la distribution par genre littéraire, les nouvelles et les essais se placent du côté des indéfinis, mais ces textes ne discriminent aucunement le masculin du féminin.

Par leur fréquence, les articles apparaissent dans le corpus, de manière décroissante, dans l'ordre suivant : *le*, *la*, *un*, *l'*, *les*, *des*, *une*, *au*, *aux* et *du*. En résumé, le défini prime sur l'indéfini et le singulier sur le pluriel. Ceci peut se traduire doublement, en faisant des rapports nets entre articles et substantifs, par une tendance au langage concret, plus scientifique et positiviste que philosophique ; finalement, la primauté du singulier sur le pluriel nous fait comprendre que Verne vise plus l'importance de l'individu que du collectif, observation étayée par l'impressionnante liste d'êtres autonomes qui parcourent ses textes.

3.3. Les adjectifs

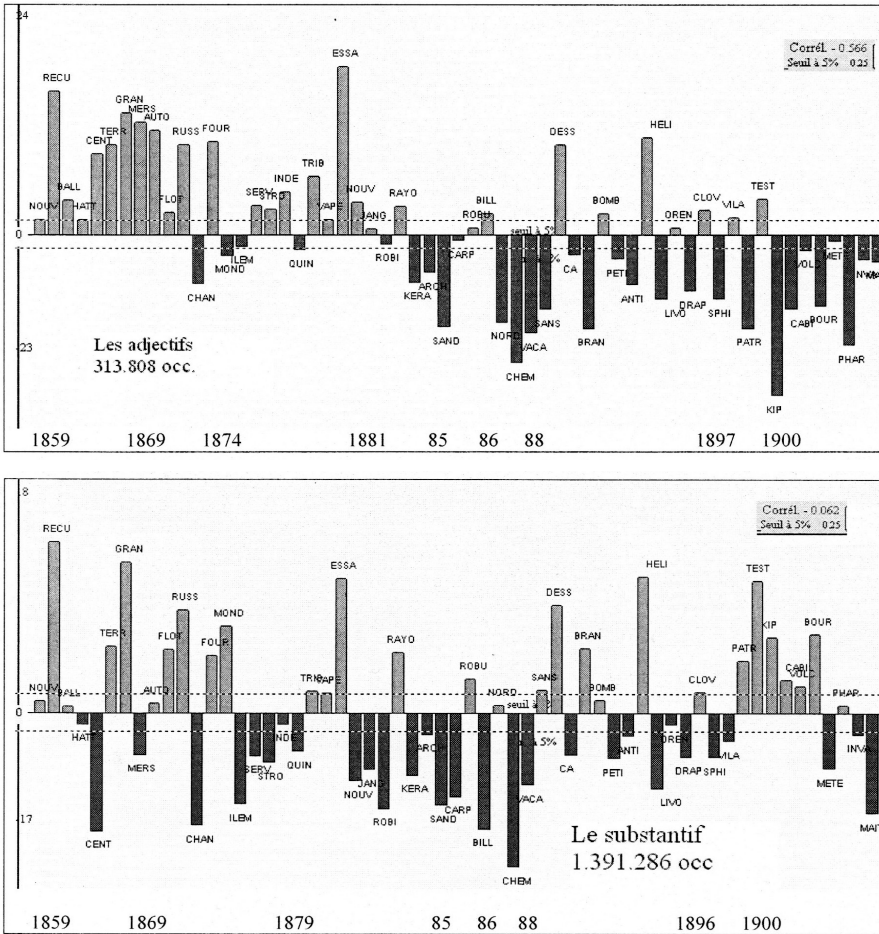


Figure N° 4. Les adjectifs et les substantifs

La catégorie des adjectifs suit de près le comportement de la catégorie des substantifs. En effet, en comparant les deux schémas de la figure 4, nous nous apercevons d'une certaine parenté entre ces deux catégories grammaticales. Les adjectifs se manifestent comme excédentaires dès le début jusqu'à *Rayon* (1881), hormis *Le Chancellor*, histoire d'un naufrage qui se veut proche de la réalité, jusqu'à abandonner tout adjectif gênant ou non nécessaire. Certains textes se voient affectés d'écarts positifs importants qui traitent plus de description que d'action (*Mers*, *Centre*, *Servadac*, *Indes*, *nouvelles 2*). Par la suite, les deux figures maintiennent une solidarité jusqu'à *Testament* (1897). Les bâtons des adjectifs d'Un

billet de loterie et de *Sans dessus dessous* présentent des écarts positifs par rapport à ceux des substantifs des mêmes romans. Par contre, ceux de *Branican* et *Patrie* se manifestent comme très déficitaires. À partir de *Kip* jusqu'à la fin, l'adjectif se montre déficitaire sans garder trop de similarités avec le schéma des substantifs. *Grosso modo*, les titres avec les écarts excédentaires les plus significatifs pour les adjectifs, correspondent à ceux dont l'écart est pareillement excédentaire pour les substantifs, exception faite notamment de certains textes de la première étape.

De façon générale, nous assistons à une baisse plus ou moins graduelle de l'usage de l'adjectif chez Verne. Le nombre total d'adjectifs relevé par le logiciel a été de 313 808, soit 5% du corpus, un peu au-dessous de Hugo (5,64%). Un peu plus de la moitié est constituée par des adjectifs masculins.

Quant au genre, les nouvelles montrent une évolution positive quant à l'usage d'adjectifs, si l'on compare les deux recueils. Les essais, de leur côté, atteignent le point maximal de la liste de textes utilisant significativement les adjectifs, et plus précisément les adjectifs qualificatifs féminins.

Les premiers romans de Verne, comme *Grant*, présentent un vocabulaire très riche, notamment en adjectifs non seulement pour la présentation de faits...

Le 26 juillet 1864, par une forte brise du nord-est, un magnifique yacht évoluait à toute vapeur sur les flots du canal du nord. Le pavillon d'Angleterre battait à sa corne d'artimon ; à l'extrémité du grand mât, un guidon bleu portait les initiales E G, brodées en or et surmontées d'une couronne ducale. Ce yacht se nommait le Duncan ; il appartenait à lord Glenarvan, l'un des seize pairs écossais qui siègent à la chambre haute, et le membre le plus distingué du « Royal-Thames-yacht-club », si célèbre dans tout le Royaume-Uni. (Grant, I partie, Chap I).

mais, bien évidemment, pour les personnages :

Edward Glenarvan avait trente-deux ans ; sa taille était grande, ses traits un peu sévères, son regard d'une douceur infinie, sa personne toute empreinte de la poésie highlandaise. On le savait brave à l'excès, entreprenant, chevaleresque, un Fergus du XIX^e siècle, mais bon par-dessus toute chose, meilleur que saint Martin lui-même, car il eût donné son manteau tout entier aux pauvres gens des hautes terres...

Miss Helena n'appartenait pas à une famille noble, mais elle était écossaise, ce qui valait toutes les noblesses aux yeux de lord Glenarvan ; de cette jeune personne charmante, courageuse, dévouée, le seigneur de Luss avait fait la compagne (Grant, I partie, Chap III).

La liste d'adjectifs très utilisés par Verne inclut des mots tels que : *grand, jeune, bon, petit, long, haut, large, profond, noir, fort, blanc, sombre...*

3.4. Les numéraux

Les chiffres chez Verne sont les bienvenus pour les multiples calculs que connaît son corpus, particulièrement dans les premières années de tutelle de son éditeur Hetzel. Le roman de la science doit compter avec un allié de premier ordre : les mathématiques ; et Verne, avec le sien, son propre cousin, le mathématicien Garcet, professeur de lettres et de mathématiques, ou l'ingénieur A. Badoureau dont les calculs pour *Sans dessus dessous* ont été payés par Verne lui-même, et inclus à la fin du roman. Il n'est pas étonnant de trouver des numéraux excédentaires dans des romans tels que : *Ballon, De la Terre à la lune, Mers, Autour, Monde, Flottante, Russes, Servadac, Robur, Sans dessus, Testament*. Les chiffres y abondent à tel point que le jeu mathématique devient fatigant, utilisé même pour décrire les personnages comme Fogg, un maniaque de la précision :

Phileas Fogg avait quitté sa maison de Saville-row à onze heures et demie, et, après avoir placé cinq cent soixante-quinze fois son pied droit devant son pied gauche et cinq cent soixante-seize fois son pied gauche devant son pied droit, il arriva au Reform-Club, vaste édifice, élevé dans Pall-Mall, qui n'a pas coûté moins de trois millions à bâtir. (Monde, chap. III).

Les calculs constituent l'un des ingrédients principaux de la recette romanesque de Verne:

En effet, fondre un canon égalant un million de fois en volume le canon de vingt-sept, c'eût été un travail au-dessus des forces humaines. On a déjà de sérieuses difficultés pour fabriquer les pièces de quarante-deux centimètres qui lancent des projectiles de sept cent quatre-vingts kilos avec deux cent soixante-quatorze kilogrammes de poudre. (Dessus, chap. XVII).

De la Terre à la lune représente le texte le plus significatif quant à l'usage des numéraux. Balistique et astronomie se divisent les calculs qui pullulent partout dans le roman. *Autour de la lune* et *Sans dessus dessous* sont apparentés naturellement au roman précédent non seulement par la nature du projet et de l'aventure mais aussi par la présence de la plupart des personnages. *Ballon, Mers, Hélice* et *Robur* mettent en relief, grâce aux chiffres, les qualités techniques des machines et les distances que les protagonistes parcourent. Dans *Russes*, il y a les grands calculs de la triangulation géodésique. *Météore* et *Servadac* nous parlent des distances interplanétaires. Trois textes sont pauvres en numéraux : *Carpathes, Nordsud* et *Famille Sans-nom*, le premier étant de caractère fantastique et les deux derniers, historique.

Du point de vue du genre, tantôt les nouvelles, tantôt les essais, présentent des écarts négatifs dans l'usage des adjectifs numériques. Les mots les plus fréquents représentant les numéraux sont, dans l'ordre décroissant : *deux, trois, cent, quatre, cinq, mille, six, huit* et *dix*, tous recherchés sur les lemmes pour éviter les possibles confusions notamment pour l'article *un* et le qualificatif *neuf*. L'AFC révèle que seuls les mots *deux* et *trois* montrent des signes d'indépendance par rapport aux autres. Quant aux chiffres proprement dits, l'ordre est le suivant : 3, 22, 20, 8, 7, 9, 11, 18 et 13. Dans l'AFC, le 13 et le 2 sont les seuls chiffres à se détacher du noyau du schéma, se positionnant le premier à côté de *Volcan* et le deuxième dans les proximités d'*Orénoque*.

3.4.1. Les nombres à quatre chiffres

Les nombres à quatre chiffres, pour la plupart des années, ont été synthétisés dans un tableau répartis au long d'un siècle (1789-1898) et comprennent les fréquences les plus représentatives.

Ici nous ne montrons qu'un échantillon des nombres à quatre chiffres (en bonne mesure ils représentent des années) qui vont entre 1848 et 1898 :

Numéro/date	Nombre d'occurrences	Numéro/date	Nombre D'occurrences
1848	23	1872	18
1849	15	1873	13
1850	20	1874	14
1851	28	1875	25
1852	39	1876	10
1853	33	1877	7
1854	33	1878	12
1855	26	1879	27
1857	58	1880	7
1858	21	1881	16
1859	35	1882	21
1860	49	1883	10
1861	35	1884	11
1862	65	1885	11
1863	23	1886	11
1864	32	1887	11
1865	14	1888	10
1866	18	1889	4
1867	37	1893	4
1868	11	1894	4
1869	10	1896	8
1870	14	1897	6
1871	11	1898	3

Tableau N° 2 Les nombres à quatre chiffres

Les années les plus récurrentes appartiennent aux décennies de 1850 et 1860. Verne travaille sur un schéma assez précis : beaucoup de références historiques pour nourrir le cadre de ses récits : « À cette même place, le Prince-Albert en 1851, l'Advance avec Kane en 1835, furent obstinément pris par les glaces pendant plusieurs semaines » (*Hatteras*, Partie I, chap. XI). Et puis, plus important, une marque d'actualité, par rapport à la date de parution du roman, pour réaffirmer le caractère réel de ses histoires. « Demain, à la marée descendante, le brick le Forward, capitaine, K.Z., second, Richard Shandon, partira de New Princes Docks pour une destination inconnue. Voilà ce que l'on avait pu lire dans le Liverpool Herald du 5 avril 1860 ». (*Hatteras*, Partie I, chap. I). Dans les deux cas, les dates jouent un rôle capital d'ancrage temporaire. Très souvent, le début de l'histoire est marqué par une date très précise : 24 février 1793 (*Le comte de Chantelaine, Nouvelles 1*), 24 mai 1863 (*Centre*), 26 juillet 1864 (*Grant*), 18 mars 1867 (*Flottante*), 27 février 1854 (*Russes*), 17 mars 1859 (*Fourrures*), etc. Ces renseignements nous font croire que la plupart des textes s'attachent fortement à une date qui précède normalement leur parution, bien évidemment quand il ne s'agit pas de textes à coloration historique.

3.4.2. Les chiffres romains

Quant aux chiffres romains, la plupart est employée comme indicateur de chapitres, ce qui fait des dix premiers les plus fréquents. Comme indicateurs des siècles, ils se repartissent de la manière suivante dans l'ordre décroissant pour les formes finissant par 'ème' : XVIII, XIX, XVII, XVI, XIII, XII. Pour les formes finissant par 'e', nous avons repéré les résultats suivants pour les fréquences les plus élevées : XX, XVIII, XVII, XVI, XIX, Ceci n'implique pas que l'auteur ne fait pas référence aux siècles avec d'autres moyens (quinzième, seizième). Certains sont des formes particulières de romans bien précis (XXe pour *Bombarnac*) ; dans d'autres cas les chiffres romains indiquant les siècles se distribuent de façon plus homogène dans le corpus. Les essais montrent une forte tendance aux écarts positifs.

3.4.3. Les ordinaux

Les ordinaux, par leur fréquence, se distribuent dans l'ordre suivant : premier, troisième, deuxième, quatrième, cinquième, septième et sixième. L'AFC des ordinaux révèle que la plupart des ordinaux se neutralisent au centre du schéma. Seuls trois adjectifs numéraux se séparent facilement du reste, à savoir : *premier*, autour duquel se réunissent les romans suivants : *Cascabel*, *Petit bonhomme*, *Volcan*,

Ile mystérieuse, Patrie, Strogoff, Vacances et notamment *Orénoque* ; douzième, qui attire *Kip, Ballon; Centre* et *Hatteras* ; et ; finalement, dixième, proche de *Kériban, Fourrures* et *Hatteras*. S'il y a un texte considérablement excédentaire en adjectifs ordinaux, il s'agit précisément de *Testament d'un excentrique*, roman qui reprend le thème du jeu de l'oie dont les cases correspondent aux différents états qui composent l'union américaine.

En général, les adjectifs numéraux se montrent déficitaires dans les trois genres qui composent le corpus étudié.

Conclusions

Le corpus de Verne se caractérise par une importante présence de substantifs et de verbes, avec la primauté de la classe nominale sur celle du verbe. Il serait risqué nonobstant d'étiqueter son discours d'ouvertement nominal, c'est-à-dire, un discours où primerait la description sur l'action. Il n'est certainement pas, non plus, verbal. Il y a une tendance hugolienne chez Verne étant donné ses préférences en pourcentage pour les verbes et les substantifs. Chronologiquement, la description règne au début de carrière pour céder, par la suite, la place à l'action.

La prédominance des substantifs et naturellement de déterminants masculins sur les féminins, très particulière chez Verne, nous fait croire à une volonté d'un discours concret, très attaché à l'exposition scientifique des faits, et à l'usage d'un langage plus « terre à terre » et moins motivé par les réflexions philosophiques ou littéraires normalement riches en substantifs féminins. La plupart des sujets abordés par l'auteur ne favorisent pas l'abstraction, mais ce fait n'empêche pas la forte présence de substantifs féminins dans certains textes, notamment dans les nouvelles et les essais. Une autre conclusion importante à retenir, par rapport aux substantifs, concerne leur corrélation avec la richesse du vocabulaire : le nombre de substantifs est proportionnel à la richesse lexicale.

Nous avons relevé une prédominance des substantifs singuliers sur les pluriels et des noms propres masculins sur les féminins, et bien évidemment de l'article masculin sur le féminin, et, dans cette catégorie, du défini sur l'indéfini.

Adjectifs et substantifs, évoluent de concert et manifestent un comportement semblable dans les écarts excédentaires ou déficitaires.

Dans le discours scientifique prédominant dans l'ensemble du corpus, les mathématiques occupent une place d'honneur pour les multiples calculs dans lesquels l'auteur se lance pour donner à ses récits un caractère réel ou vraisemblable.

Bibliographie

- Brunet, E. 1998. *Le Vocabulaire de Victor Hugo*. Champion-Slaktine : Paris-Genève.
- Brunet, E. 2009 a. *'Hyperbase Manuel de référence*, Version lemmatisée 8.0.
- Brunet, E., 2009 b. *Comptes d'auteurs*, Tome I. Paris : Honoré Champion.
- Dekiss, J.-P. 1999. *Jules Verne l'enchanteur*, biographie. Paris : Édit. du félin.
- Diesbach, G. de. 2000. *Le tour de Jules Verne en 80 livres*. Paris : Perrin.
- Dusseau, J. 2005. *Jules Verne*. Paris : Perrin.
- Lottman, H. 1996. *Jules Verne*. Paris: Flammarion.
- Magri, V. 1995. *Le discours sur l'Autre à travers quatre récits de voyage en Orient*. Paris : Honoré Champion.
- Mayaffre, D. 2004. *Paroles de président. Jacques Chirac (1995-2003) et le discours présidentiel sous la V^e République*. Paris.
- Parada, J. « Stylistique de la Phrase vernienne sous l'optique de la textométrie ». *Matices en Lenguas Extranjeras*, [S.l.], n. 7, p. 1-21, jan. 2013. ISSN 2011-1177. Disponible en ligne. Consulté le 16 mars 2016.
- Parada, J. 2011. « Panorama d'un corpus millionnaire ». *Hechos y proyecciones del lenguaje*. Pasto, Colombia, n°. 19, p. 241-265.

Notes

1. Le logiciel propose, entre autres, des tableaux avec des résultats numériques signalant les fréquences de chaque catégorie grammaticale et l'écart entre elles. Nous mettons entre parenthèses les écarts les plus significatifs nous permettant de faire des comparaisons entre les textes.
2. Nelly Bly, alias Elizabeth Cochrane, première femme à essayer de battre le record de Phileas Fogg dans un voyage autour du monde. Cette interview a lieu un mois après la visite que cette femme, journaliste américaine et travaillant pour le même journal, rend à Verne à Amiens en 1889.
3. Sans vouloir parler d'une superposition du genre grammatical sur le genre sexuel, on peut constater une prédominance du féminin grammatical dans le peu de romans où les femmes jouent un certain rôle d'importance : *Centre*, *Grant*, *Fourrures*, *Rayon*, *Patrie*.
4. Le mot « homme » dans le corpus apparaît 4 391 fois contre le mot « femme » présent 1 437 fois seulement.
5. *Du* et *des* sont des formes ambiguës en raison de leur double fonction. Nonobstant, leur proportion majeure est occupée par des articles.