



PSL 



Spécificités Stylistiques des Prix Littéraires.

Une Étude Quantitative dans la Littérature Française du
XXe siècle.

Pedro Cabrera Ramirez, Jean Barré

1 december 2023

École Normale Supérieure - Lattice

1. Introduction

2. Méthode

Corpus

Caractéristiques textuelles

Modélisation statistique

3. Résultats

Efficacité du modèle

Coefficients Discriminants

4. Analyse

Interprétabilité

Signal de la Littérature Primée

Retour en Lecture Proche

5. Conclusion

Études Littéraires Computationnelles

- « Lecture Distante » (Moretti, 2000) [1]
- Comprenons nous les lignes structurantes de l'histoire littéraire? (Underwood, 2019) [2]
- Canon littéraire - Prestige littéraire

Les prix littéraires comme témoin de littéarité ?

- Peut-on mesurer les différences stylistiques entre la littérature prestigieuse et les autres ?
- Quelles sont les conventions de la qualité littéraire au XXIème siècle ?
- Quelles caractéristiques stylistiques des romans ont pu contribuer à l'attribution de prix littéraires ?

Hypothèses

- Le roman primé est donc considéré comme un genre
- Détection du « Moment grammatical de la littérature française » ? (Philippe, 2002) [3]
- Présence de structures discursives latentes dans l'ensemble des romans primés. Motifs latents ?

Sous-partie du corpus Chapitres (Leblond, 2022)[4]

- Corpus de 3000 romans du XIXe et du XXe siècle
- Étiqueté par genre : roman d'aventures, roman historique, roman sentimental, policier, science-fiction....
- Sélection des romans primés dans les concours suivants :
Renaudot, Femina, Goncourt, Académie Française, Interalliés, Médicis
- 1900-2020 : 1347 romans non-primés & 88 romans primés

Les Motifs stylistiques : Construction

- « patron lexico-grammatical extrait (automatiquement ou non) à partir de textes étiquetés » (Diwersy et Legallois, 2014)[5]
(Legallois et al, 2018)[6]
- En pratique : lemmes de formes invariables (prep, conj, adv), de certains verbes (modaux, auxiliaires) et POS des autres mots (Noms, verb, adj)
- ADJ_NOUN_de_NOUN_le : « Petit déjeuner de l'écolier » ou « Nouveau chapitre de l'histoire ».

Les Motifs stylistiques : Récupération

- Récupération des 1000 ngrams de motifs les plus fréquents, avec une longueur des motifs allant de 1 à 5.
- 4158 motifs correspondant à autant de dimensions dans le tableau final
- Fréquence relative des motifs dans chaque texte

Machine à Vecteur de Support (SVM)

- Chaque texte est vu comme un point de données caractérisé par plusieurs milliers de variables
- Machine à Vecteurs de Support : optimisation de la classification dans un hyper-plan à N dimensions. Dans notre cas, chaque dimension est un motif stylistique
- Nouveaux observables fondés sur les inférences des modèles prédictifs

Détection automatique des prix littéraires : Résultats

Efficacité : 0.77

Baseline : 0.493

	précision	rappel	f1-score	support
Lauréat d'un prix	0.78	0.75	0.76	88
Non lauréat	0.76	0.78	0.77	88

Table 1 – Résultats de l'évaluation du modèle

Comment passer de la prédiction du modèle à la compréhension des critères de classification ?

- Problème : unités peu informatives
- Décalage entre la précision du modèle et la reproduction des résultats par un humain
- Retour à la lecture proche

- Stylométrie
roulante
- Découpage des
textes dans des
morceaux
- Projection des
coefficients dans
le texte
- Objectif :
sélection de
coordonnées
textuelles
paradigmatiques
selon les critères
du modèle

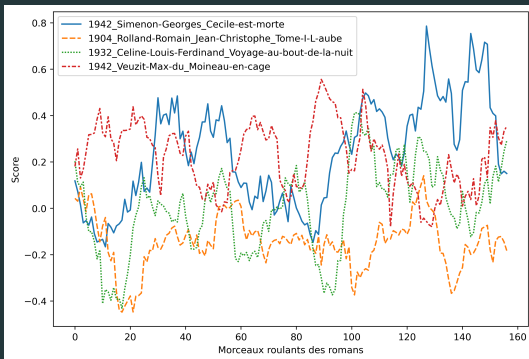


Figure 2 – Signal du roman primé dans 4 romans

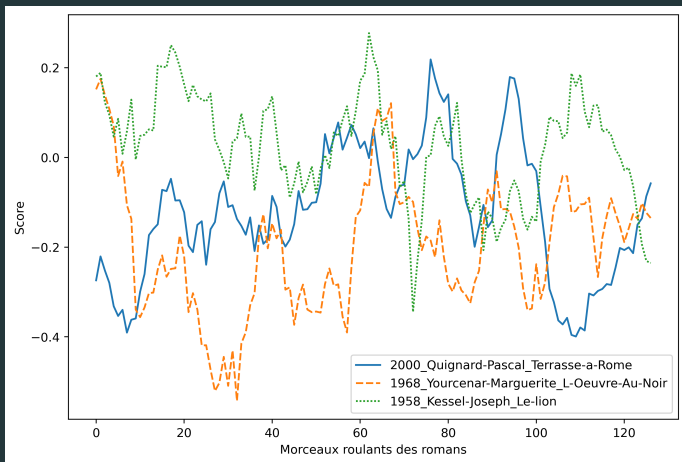


Figure 3 – Signal du roman primé dans 3 romans

Henri-Juste confia l'écolier à son beau-frère, Bartholommé Campanus, chanoine de Saint-Donatien à Bruges. Ce savant usé par la prière et l'étude des bonnes lettres était si doux qu'il semblait déjà vieux. Il apprit à son élève le latin, le peu qu'il savait de grec et d'alchimie, et amusa la curiosité de son écolier pour les sciences à l'aide de l'Histoire naturelle de Pline. Le froid cabinet du chanoine était un refuge où le garçon échappait aux voix des courtiers discutant les draps d'Angleterre, à la plate sagesse d'Henri-Juste, aux caresses des chambrières curieuses de fruit vert. Il s'y libérait de la servitude et de la pauvreté de l'enfance ; ces livres et ce maître le traitaient en homme. Il aimait cette chambre tapissée de volumes, cette plume d'oie, cet encrier de corne, outils d'une connaissance nouvelle, et l'enrichissement qui consiste à apprendre que le rubis vient de l'Inde, que le soufre se marie au mercure, et que la fleur qu'on nomme lilium en latin s'appelle en grec krinon et en hébreu susannah.

Figure 4 – Marguerite Yourcenar. *L'œuvre au noir*. Morceau type de littérature primée.

Une autre carte impudique exécutée dans la même manière (c'est-à-dire qui date d'après 1656) est singulière. Il s'agit de la Tentation de saint Antoine. Le saint ermite se tient assis au-devant de la grotte, le sexe dressé dans sa main. Ses yeux pleurent. Une rocaille sépare le saint d'une femme qui ouvre largement ses jambes, la tête penchée en avant sur sa nuit, qu'elle semble observer, mais qui est indiscernable. Près d'elle un petit diable chie dans un livre qui est ouvert.

Figure 5 – Pascal Quignard. Terrasse à Rome. Morceau type de littérature non-primée.

Le jeune patricien qu'il avait ausculté avec attention, âgé de vingt ans, vigoureux, doté de parties génitales bien conformées, affirmait ne pas pouvoir se marier parce qu'il n'avait jamais éprouvé de sa vie de désir. [...] Marcello Zerra prescrivit des images obscènes qu'Eugenio examinerait une nuit entière en compagnie de deux prostituées florentines, l'une âgée et douce, pour ne pas dire complaisante, l'autre beaucoup plus jeune et vive. Cette tentative non seulement échoua mais provoqua chez Eugenio un dégoût qui alla jusqu'à la nausée, et cette nausée fut si violente qu'elle le fit tomber dans l'angoisse. Les filles de joie de la cité de Florence furent incapables de se mettre d'accord sur le résultat de leurs efforts nocturnes. [...] La putain la plus âgée, craignant de ne pas percevoir la rétribution dont ils étaient convenus pour les deux voyages à l'aller comme au retour, et pour la nuitée entière, prétendit que c'était là une assertion inexacte, que le jeune homme avait eu une érection fugitive et qu'une autre nuit viendrait aisément à bout de ses réticences et autres difficultés qu'elle avait eu le loisir d'observer attentivement. Le jeune Eugenio s'évanouit en entendant la femme de mauvaise vie émettre la proposition d'une nouvelle nuit de plaisirs. [...] Dans le palais familial, interrogées le jour même, toujours par Zerra, les lingères attestèrent qu'elles n'avaient jamais vu la moindre trace de pollution nocturne sur les draps du fils aîné. [...] Eugenio ne parvint jamais à accomplir l'acte sexuel avec son épouse. La jeune femme, toujours intacte, se plaignit aux siens, qui répercutèrent sa détresse.

Figure 6 – Pascal Quignard. Terrasse à Rome.

Statut du roman érotique

	Index	Preds	Proba_A	Proba_nonA
	1975_San-Antonio_Dis-bonjour-a-la-dame	not_awarded	0.240698	0.759302
	1975_San-Antonio_La-vie-privee-de-Walter-Klozett	not_awarded	0.344071	0.655929
	1904_Regnier-Henri-de_Les-Rencontres-de-M-de-B...	not_awarded	0.178458	0.821542
	1908_Barbusse-Henri_L-enfer	awarded	0.896430	0.103570
	1902_Rebell-Hugues_Les-Nuits-chaudes-du-Cap-fr...	not_awarded	0.150884	0.849116
	1956_San-Antonio_Ca-tourne-au-vinaigre	not_awarded	0.119479	0.880521
	1957_San-Antonio_Les-anges-se-font-plumer	not_awarded	0.202517	0.797483
	1954_Reage-Pauline_Histoire-d-O	not_awarded	0.129326	0.870674

Figure 7 – Prédications sur le corpus de romans érotiques

Promesses et limites de l'analyse automatique

- Le modèle montre la possibilité de récupérer un sociolecte du prestige littéraire, mais la précision doit être améliorée
- Les premières explorations suggèrent la mise au ban de certains genres spécifiques
- Décalage entre la performance du modèle et l'interprétabilité
- Besoin d'une approche heuristique pour la compréhension des résultats

Merci de votre attention !



Franco Moretti.

Conjectures on world literature.

New Left Review, 2000.



Ted Underwood.

Distant horizons : digital evidence and literary change.

The University of Chicago Press, 2019.



Gilles Philippe.

Sujet, verbe, complément : le moment grammatical de la littérature française, 1890-1940.

Bibliothèque des idées. Gallimard, 2002.

OCLC : ocm51263704.



Aude Leblond.

Corpus chapitres, December 2022.



Sascha Diwersy and Dominique Legallois.

L'apport de la méthode des motifs aux analyses phraséologiques en discours.

In *Approches théoriques et empiriques en phraséologie*, Nancy, France, 2014.



Dominique Legallois, Thierry Charnois, and Meri Larjavaara.

The balance between quantitative and qualitative literary stylistics : how the method of motifs can help.

In Dominique Legallois, Thierry Charnois, and Meri Larjavaara, editors, *The Grammar of Genres and Styles*, pages 164–193. De Gruyter, 2018-04-09.

Questions ?

N'hésitez pas à nous écrire !

jean.barre@ens.psl.eu, pedro.cabrera.ramirez@ens.psl.eu

Lien des données et du code :

<https://github.com/crazyjeannot/AWARDS>