

L'évolution en français de l'adjectif épithète vers la postposition : réalité syntaxique ou trompe-l'œil lexical ?

Jean-Luc Manguin

CRISCO, CNRS UMR 6170 – Université de Caen, 14032 CAEN Cedex –
manguin@crisco.unicaen.fr

Abstract

This study bears on the temporal variations of the position of attributive adjectives in French, in the category-tagged corpus Frantext, between 1830 and 1959, and in four different literary genders ; after we have noticed the evolution of the position of attributives, we propose four hypothesis to explain the observed variations, then we examine them. At the end, it appears that one of them (based on the variations of use) explains most of the observed effects. As a final point we propose to explore a stylistic hypothesis to go further in the explanation of these variations of use.

Résumé

Cette étude porte sur les variations chronologiques de la position de l'adjectif épithète en français, dans le corpus Frantext catégorisé, entre 1830 et 1959, et ceci dans quatre genres littéraires ; après avoir constaté l'évolution de la position de l'épithète, nous proposons quatre hypothèses pour expliquer les variations observées, puis nous les examinons. Il s'avère en conclusion que l'une d'entre elles (qui concerne les variations d'emploi) explique la majorité des effets observés. Nous proposons au final une piste à caractère stylistique pour expliquer ces variations d'emploi.

Mots-clés : adjectif épithète, variation de position, corpus français écrit.

1. Introduction

La question de la position de l'adjectif épithète en français a déjà été étudiée, engendrant une littérature abondante (voir notamment Reiner, 1968 ; Forsgren, 1978 ; Wilmet, 1981 ; la synthèse de Goes, 1998 : 77-107 ; ou encore le point de vue récent de Abeillé et Godard, 1999) ; cela dit, si la plupart des auteurs s'attachent à expliquer la position de l'épithète dans un cadre synchronique, la question de son évolution chronologique éventuelle depuis le 19^e siècle jusqu'à notre époque reste en revanche relativement peu abordée¹. Et par ailleurs, il est parfois frustrant d'entendre affirmer que « l'adjectif évolue vers la postposition » sans que rien d'autre qu'un comptage d'occurrences ne vienne véritablement expliquer cette constatation, ce qui laisse le lecteur quelque peu sur sa faim (Damourette et Pichon, 1911-1930 : § 528). Grâce à l'arrivée de corpus étiquetés, et de moyens informatiques, le traitement statistique de données textuelles permet d'étudier ce problème non seulement sur un nombre d'occurrences beaucoup plus élevé, mais encore de l'examiner du point de vue du lexique, ce qui n'avait été fait que pour les couples d'unités les plus employées ; l'apport de la lemmatisation est fondamental dans ce cas, et c'est, comme nous le verrons, ce qui permet de dégager un premier niveau de causes pour la période considérée.

¹ Il faut cependant signaler l'étude de Hug (1971) et préciser que l'étude de Glatigny (1967) intègre l'aspect temporel, mais pour un seul auteur (Nerval).

2. Étude des occurrences

2.1. Recueil des données

Nous nous servons du Corpus Frantext catégorisé, dans la période 1830-1959, et dans les quatre genres suivants : Mémoires, Romans, Essais & Traités, et Théâtre ; l'ensemble de ces textes compte 90 millions de mots répartis dans environ 1300 œuvres ; si l'on veut découper cet ensemble en trois tranches à peu près régulières de 30 millions de mots chacune, on aboutit alors à un partage en trois périodes : 1830-1859, 1880-1929 et 1930-1959. Cela dit, un examen préalable du corpus nous révèle deux choses :

- 1) Le corpus Frantext n'est pas homogène en genre tout au long de la période choisie, comme le prouve la proportion de chaque type de texte dans les trois tranches temporelles² :

	1830-79	1880-1929	1930-59
essais + traités	16,2 %	23,4 %	41,2 %
romans	62,0 %	55,5 %	41,2 %
mémoires	16,6 %	16,0 %	11,2 %
théâtre	5,1 %	5,1 %	6,4 %

Ce qui est ici flagrant, c'est l'énorme variation de la proportion des textes techniques (Essais & traités) au cours des trois périodes.

- 2) La proportion moyenne d'épithètes postposées sur toute la période d'étude varie suivant le genre de texte, comme l'avait déjà noté Larsson (1994) ; le tableau ci-dessous donne les valeurs dans notre corpus :

	proportion moyenne
essais + traités	71,6 %
romans	54,8 %
mémoires	57,5 %
théâtre	46,7 %

On se rend compte à partir de là, qu'il est donc nécessaire d'étudier séparément chaque genre littéraire, et de relever dans chacun le nombre d'épithètes anté- ou postposées ; cette étude, faite par tranches décennales³, nous donne les résultats ci-dessous, après régression linéaire des données :

	proportion moyenne	variation estimée	coefficient de détermination ⁴
essais + traités	71,6 %	10,2 %	0,55
romans	54,8 %	6,5 %	0,69
mémoires	57,5 %	6,7 %	0,39
théâtre	46,7 %	3,7 %	0,05

² Le calcul est effectué sur le nombre de mots.

³ Nous donnons en annexe 1 les chiffres détaillés, excepté pour le théâtre.

⁴ Il s'agit ici du coefficient r^2 de la régression linéaire.

Ce tableau synthétique nous indique que la progression vers la postposition est sensible dans tous les genres⁵ ; toutefois, dans le théâtre, les données relevées dans Frantext ne permettent pas de conclure à une progression linéaire, ni même à une progression tout court. Pour les trois périodes « de base », les occurrences pour ce genre sont en effet :

théâtre	1830-1879	1880-1929	1930-1959
œuvres	67	64	71
mots	1477033	1553808	1957934
occurrences anté	12165	13058	15840
occurrences post	9327	13239	13471
proportion post	43,4 %	50,3 %	46,0 %

autrement dit, une évolution « en cloche » qui ne révèle pas de tendance globale.

Dans la suite de cette étude, nous nous limiterons donc aux trois genres non dramatiques.

2.2. Quantification des erreurs

Chaque utilisateur de la base Frantext catégorisée a pu constater que, à l'instar de ses « confrères », l'étiqueteur utilisé pour Frantext commet un certain nombre d'erreurs, que certains ont déjà étudiées en détail dans des cas précis (voir Hug, 2002). Dans notre cas, les erreurs sont, comme nous allons le voir, plutôt rares ; ceci vient du fait que nous imposons la cooccurrence de deux catégories. Ainsi, en antéposition, nous demandons au moteur de recherche de Frantext de nous trouver une structure Adjectif + Nom (ce qui donne la formulation $\&e(g=A)$ $\&e(g=cS)$ dans l'interface de Frantext). Seule la fonction d'épithète peut ainsi être captée, et les éventuelles erreurs ne proviennent que de confusions entre homographes.

Pour la postposition, la requête formulée Nom [+ Adverbe] + Adjectif (soit dans Frantext $\&e(g=cS)$ $\&?(\&e(g=Adv))$ $\&e(g=A)$), capte non seulement la fonction d'épithète, subissant également les erreurs d'étiquetage, mais aussi l'attribut du complément d'objet (qui reste assez peu fréquent) et la fonction attributive dans le cas d'une accumulation provoquant l'ellipse de la copule (voir exemple ci-après).

Pour estimer les erreurs par la méthode d'échantillonnage, nous avons vérifié manuellement un échantillon que nous avons prélevé ainsi : pour chaque tranche décadaire, nous avons conservé les 50 premières occurrences, ce qui nous donne 650 exemples ; cet échantillonnage est effectué séparément pour chaque position adjectivale, car nous supposons, pour les raisons détaillées précédemment, que le taux d'erreurs ne sera pas le même dans les deux cas.

Voici les résultats concernant la partie « Mémoires » du corpus :

	Antéposition	Postposition
Erreurs de catégorie	11	13
Erreurs de fonction		10
Taux d'erreurs	1,69 %	3,54 %
Borne inférieure	1,56 %	3,28 %
Borne supérieure	1,82 %	3,80 %

⁵ La limite de 1959 a été choisie en raison de l'absence de textes du genre Mémoires au-delà de cette date ; néanmoins, une étude avec une date limite ultérieure, dans les genres Roman et Essais & traités, amène une amélioration du coefficient de la régression linéaire.

Pour évaluer les bornes entre lesquelles se situe le taux d'erreur avec une probabilité de 95 %, nous utilisons la méthode classique de Student (on peut se reporter, pour le détail du calcul, à Dodge, 1999 : 257-259).

Comme on le voit, le taux d'erreur est faible, et surtout, si l'on soustrait le nombre d'erreurs estimées à nos résultats, les variations probables (0,13 % pour l'antéposition et 0,26 % pour la postposition, ce qui fait 0,4 % de précision sur la proportion) restent négligeables au regard des variations chronologiques observées.

Pour entrer dans les détails, voici quelques erreurs relevées ; tout d'abord un cas d'antéposition qui n'en est pas un :

*ils se déclaraient hostiles au principe de la nation armée dans lequel ils voyaient l'amorce d'une **armée milicienne** ;*

(Maréchal Joffre, *Mémoires 1910-1917*, 1931, Page 8)

Ensuite, l'inverse, c'est-à-dire une antéposition réelle détectée comme postposition :

*Dans ce pays de Valois, l'été n'est que le **trouble confluent** du printemps et de l'automne.*

(François Mauriac, *Journal III*, 1940, Page 209)

Enfin, voici un cas d'accumulation de plusieurs structures attributives, où l'absence de la copule fait entrer dans nos résultats trois occurrences sans épithète :

*ses connaissances étaient immenses, son ardeur d'étude infatigable, sa **conversation brillante**, son caractère excellent.*

(Benjamin Constant, *Le cahier rouge*, 1830, Page 12)

Pour les Romains, ainsi que pour les Essais & Traités, le taux d'erreur est tout à fait comparable. Ayant ainsi validé les variations que nous avons observées, nous pouvons maintenant tenter de les expliquer.

2.3. Hypothèses de travail

Revenons à nos chiffres, qui révèlent une progression vers la postposition, dans le prolongement sans doute de celle déjà signalée par Damourette et Pichon, 1911-1930, qui voyaient en elle une « évolution de la langue » qui « paraît se continuer de nos jours »⁶. D'après ce que nous connaissons de l'évolution des langues en général, les effets de cette évolution qui peuvent se manifester dans notre étude sont d'ordre syntaxique ou d'ordre lexical.

Au point de vue syntaxique tout d'abord, il est généralement admis (voir Glatigny, 1967 ; Wilmet, 1981) que la position postposée est une caractéristique « moderne » du français, à laquelle certains adjectifs (les plus anciens, qui sont aussi les plus courants) résistent pour diverses raisons qui sont discutées par ces auteurs. On peut alors penser que les unités qui n'ont pas de position préférentielle nette vont, dans le mouvement évolutif, adopter à force d'usage un placement qui deviendra à la longue de plus en plus postposé. D'autre part, comme l'a constaté Glatigny (1967), le doublet {Adjectif, Nom} va généralement s'ordonner suivant les masses croissantes ; cela est bien sûr renforcé par l'adjonction de « satellites » à l'adjectif (adverbe ou complément, par exemple) qui vont accroître la masse du groupe adjectival. Forsgren (1978 : 159-162) avait lui aussi constaté la corrélation entre présence de l'adverbe⁷ et position postposée ; il est alors possible d'imaginer que l'évolution se traduise

⁶ Citation du paragraphe 528 du tome II de l'ouvrage « Des mots à la pensée ».

⁷ À l'exception toutefois de *très*.

aussi par un emploi plus fréquent des adverbes.

Au point de vue lexical ensuite, nous savons d'une part que, selon ces mêmes auteurs, la création de nouveaux adjectifs (par dérivation, généralement) donne des unités adjectivales qui seront généralement postposées pour deux raisons : la dérivation suffixale leur confère une masse syllabique importante, et leur nouveauté ne les pourvoit que d'un sens précis et objectif (Glatigny, 1967 ; Reiner, 1968). D'autre part, si le lexique évolue dans son volume, la fréquence d'emploi de ses éléments varie suivant l'époque ; le fonds ancien des adjectifs, dont la caractéristique essentielle est d'être presque toujours antéposés, pourrait-il être victime du progrès et voir son usage décroître, en même temps que son sens se dilue, qu'il se banalise (Blinkenberg, 1933), et que son poids syntaxique s'affaiblit (Abeillé et Godard, 1999) ?

Ainsi, pour expliquer les variations que nous avons constatées, nous pouvons envisager quatre hypothèses, qui peuvent coexister, mais dont il importe de définir les rôles respectifs ; ces quatre possibilités reflètent chacune la variation d'un seul paramètre, tandis que les autres restent constants :

- a) variation syntaxique : les adjectifs tendent globalement à être plus employés postposés ; par conséquent, si l'on mesure pour chaque adjectif la variation chronologique de la proportion d'occurrences postposées, la moyenne de toutes ces variations doit être positive (mouvement macroscopique qui n'exclut pas des exceptions individuelles).
- b) élargissement lexical : le nombre d'adjectifs augmente, et les nouveaux venus s'emploient plutôt en position postposée. Pour valider cette hypothèse, il faut connaître le nombre d'adjectifs employés dans chaque période.
- c) fréquence d'emploi : le lexique reste constant, mais les adjectifs préférentiellement postposés sont employés plus souvent (ou les antéposés moins souvent).
- d) emploi d'adverbes : certains adverbes induisant la postposition sont plus fréquemment employés.

L'hypothèse d) est rapidement infirmée, car les occurrences impliquant des adverbes, en postposition comme en antéposition, sont en nombre quasi constant, voire en diminution ; on relève par exemple, pour les Romains :

adverbes	1830-79	1880-1929	1930-59
Anté	11112	7363	4479
Post	25727	25411	18247

soit, en pourcentage du total des occurrences :

adverbes	1830-79	1880-1929	1930-59
Anté	4,7 %	3,5 %	3,1 %
Post	10,1 %	9,4 %	9,6 %

Cette hypothèse doit donc être rejetée.

3. Étude des lemmes

La poursuite de notre étude s'est déroulée de la manière suivante : après rapatriement des occurrences fournies par Frantext, nous avons isolé les formes adjectivales, puis nous les

avons regroupées par lemmes au moyen d'un dictionnaire des formes fléchies ; au passage, cette lemmatisation a éliminé quelques formes rares, principalement des ordinaux composés. À l'issue de cette étape, nous avons pu comparer, pour chacun des trois genres, et pour chacune des trois époques, le nombre d'occurrences relevées et le nombre de lemmes impliqués dans les deux cas d'antéposition ou de postposition ; les relevés complets sont donnés dans l'annexe 2. Contrairement aux chiffres des cooccurrences, les données concernant les lemmes ne sont pas sensibles aux erreurs catégorielles ou fonctionnelles discutées précédemment ; en effet, les unités impliquées dans ces erreurs apparaissent également dans des occurrences « vraies », et n'ont donc pas à être éliminées.

Dans ces tableaux, nous avons fait figurer en bas le résultat du test du χ^2 (autrement dit la probabilité d'indépendance des deux caractères, qui sont ici la situation temporelle et la position de l'adjectif) ; nous remarquons d'abord que l'élargissement du lexique adjectival associé à la postposition n'apparaît pas dans les Mémoires, et que dans les deux autres genres, il va de pair avec un élargissement plus sensible du lexique associé à l'antéposition, au point que la proportion d'adjectifs qui sont employés postposés tend à diminuer (comme dans les Essais & traités). Cette variation est significative, si l'on regarde les résultats des tests statistiques. Parallèlement à cela, l'emploi de l'antéposition diminue (la densité des adjectifs antéposés chute grosso modo de 15 %) tandis que l'emploi de la postposition varie de manière diverse : pas de variation pour les Mémoires, assez peu pour les Romans (+ 5 %) et beaucoup pour les Essais & traités (+ 26 %). Ces variations d'emploi sont elles aussi, très significatives.

On pourrait maintenant conclure avec une quasi-certitude que l'hypothèse B (lexique) n'est pas prépondérante, et que c'est l'hypothèse C (emploi) qui prime ; néanmoins, une certaine rigueur nous oblige à prendre avec précaution des statistiques qui incluent des adjectifs n'apparaissant qu'une fois tous les 20 ou 30 ans ; pour nous affranchir de ces critiques, nous avons repris les mêmes relevés en ne conservant que les adjectifs observés au moins 20 fois dans chaque tranche, dans l'une ou l'autre des positions⁸ (pour les romans, ce seuil est fixé à 50, en raison de la taille de cette partie du corpus). Le nombre de lemmes est ainsi nettement réduit, et par conséquent, le nombre d'occurrences aussi (d'environ 20 %) ; nous donnons en annexe 3 les relevés correspondants. Cette fois, comme en témoigne le test du χ^2 , le lexique ne varie pas (sauf dans les Essais & traités, mais nous y reviendrons) ; par contre, on note toujours une baisse de l'emploi antéposé, et même parfois une augmentation de l'emploi postposé (+ 10 % dans les Essais et traités). Ceci confirme donc l'hypothèse C, selon laquelle l'emploi antéposé décroît, tandis que l'emploi postposé reste stable ou augmente dans certains textes.

Le lexique antéposé impliqué dans les Essais et traités s'élargit dans des proportions significatives, et ce de manière assez singulière ; en effet, les adjectifs qui apparaissent comme antéposés dans la dernière période sans être utilisés dans cette position durant les deux premières, sont essentiellement postposés ; ils accèdent après 1930 à l'usage antéposé, assez rarement toutefois⁹. Ces adjectifs sont (pour les moins rares) : *primordial* (6 occurrences), *providentiel* (5), *précaire* (5), *graduel* (5), *gratuit* (4). En voici quelques exemples :

et, comme premier milieu social proprement humain, la famille avec sa vie propre, écono-

⁸ En d'autres termes, cela veut dire que l'adjectif est employé au moins 60 fois en antéposition OU 60 fois en postposition.

⁹ Il faut aussi signaler que certains adjectifs — comme *équivalent*, employé 5 fois en antéposition avant 1880 dans ce genre de textes — ne se rencontrent plus dans cette position par la suite. Il y a donc un double mouvement dans le lexique antéposé, qui résulte en un enrichissement ; par contre, le lexique postposé est remarquablement stable.

*mique et spirituelle, et avec ses **primordiales** exigences de droit naturel dont la législation civile fixe les modalités, mais qu'elle ne crée point.*

(Jacques Maritain, *Humanisme intégral*, 1936, Pages 89-90)

*Aujourd'hui que nous avons dominé l'effroi nu des présocratiques, la fausse et **précaire** sérénité de l'âge aristotélicien,*

(Emmanuel Mounier, *Traité du caractère*, 1946, Page 120)

*le royaume franc de Jérusalem, bénéficiant de la division politico-confessionnelle entre le khalifat fâtimide du Caire et les royaumes turcs de la Syrie intérieure, favorisé en Syrie même par le **providentiel** émiettement turco-arabe, jouant à son gré de l'anarchie musulmane, apparaissait comme l'arbitre de l'orient.*

(René Grousset, *L'épopée des croisades*, 1939, Pages 208-209).

*Il n'y a aucune raison a priori de penser que cette **graduelle** réduction de l'organique à l'inorganique doive un jour trouver sa limite. Le mécanisme a donc actuellement une position extrêmement solide,*

(Jean Rostand, *La vie et ses problèmes*, 1939, Pages 154-155)

*regarde le méchant bien en face et il l'excuse, non pas bien qu'il soit méchant, mais parce qu'il est méchant notre frère, c'est-à-dire par scandaleuse, absurde et **gratuite** charité. Le pardon est la totale franchise,*

(Vladimir Jankélévitch, *Le je-ne-sais-quoi et le presque-rien*, 1957, Page 171)

Cette constatation nous ramène à l'hypothèse A (syntaxe) que rien pour le moment n'est venu infirmer ; nous avons en effet constaté que les lexiques associés à l'antéposition ou à la postposition varient en effectif. Dans le cas des essais et traités, nous avons même précisé que ce sont des adjectifs essentiellement postposés qui accèdent à l'antéposition¹⁰. Cela laisse sous-entendre que des variations syntaxiques entrent en jeu (même si dans le cas présent, elles vont à l'encontre de la tendance qui résulte des variations d'emploi).

Pour les adjectifs courants déjà sélectionnés, nous avons alors mesuré individuellement la variation de proportion postposée ; ensuite, pour chaque genre, et sur l'ensemble des variations mesurées, nous avons calculé les paramètres statistiques élémentaires (moyenne, médiane et écart-type) ; ces calculs donnent :

	romans	essais & traités	mémoires
moyenne	1,04 %	0,99 %	-2,22 %
médiane	0,00 %	0,00 %	-0,56 %
Écart-type	8,56 %	8,61 %	10,42 %

Cela montre que globalement, la variation syntaxique est très faible sur la période considérée ; néanmoins, les variations individuelles ne sont pas inexistantes, comme le prouvent les valeurs des écarts-types. Une étude plus détaillée, que nous ne pouvons malheureusement insérer dans le cadre de cet article, montre en outre que ce sont les adjectifs dont la proportion moyenne de postposition est proche de 50 % qui varient le plus. A contrario, les adjectifs fortement postposés ou fortement antéposés ne varient quasiment pas.

¹⁰ On peut néanmoins se demander si le nombre de textes présent dans la base Frantext pour ce genre et cette période ne joue pas un rôle.

En quelque sorte, les variations syntaxiques individuelles vont globalement se compenser, malgré la variabilité extrême de certains adjectifs, comme par exemple pour le Roman, *fatal* (qui passe de 52 à 88 % d'occurrences postposées) ou à l'inverse *bref* (de 80 à 33 %) ou encore *extraordinaire* (de 99 à 67 %).

4. Conclusion

En conclusion, et même si nous avons en fin de compte retenu essentiellement l'hypothèse C (variations d'emploi), nous pouvons dire que les variations syntaxiques recèlent peut-être l'explication au niveau suivant ; en effet, il s'avère que les diminutions d'emploi sont souvent subies par des adjectifs très courants et antéposés, comme *grand*, *jeune*, *bon*, *beau*, ou encore *vieux* ; le fait qu'un adjectif comme *extraordinaire*, emphatique de *grand*, voie dans la même période ses occurrences antéposées multipliées par 20 peut suggérer que des emplois d'adjectifs courants ont pu être transférés vers des synonymes plus précis, mais pas toujours antéposables (alors que les épithètes les plus usuelles sont très souvent antéposées notamment du fait de leur brièveté, voir Guiraud, 1960 ; Glatigny, 1967 ; Muller, 1979), ou bien des synonymes antéposables, mais avec un effet de désémantisation (voir Blinkenberg, 1933 ; Wilmet, 1981 ; Goes, 1998 : 90-97).

D'autre part, il serait intéressant d'explorer plus en détail les textes théâtraux, pour tenter d'expliquer tout d'abord pourquoi on y rencontre trois fois moins d'épithètes que dans les autres genres, et ensuite pourquoi la proportion d'occurrences postposées, qui ne semble pas augmenter, y est aussi nettement plus faible (voir, là aussi, Muller, 1979, sur la lexicométrie des textes théâtraux).

Références

- Abeillé A. et Godard D. (1999). La place de l'adjectif épithète en français : le poids des mots. *Recherches linguistiques*, vol. (28) : 9-31.
- Blinkenberg A. (1933). *L'Ordre des mots en français moderne*. Levin & Munksgaard.
- Damourette J et Pichon E. (1911-1930). *Des mots à la pensée*, t. (II). D'Artrey.
- Dodge Y. (1999). *Premiers pas en statistiques*. Springer.
- Frantext, base textuelle catégorisée* (1999). CNRS, ATILF (Analyse et traitement informatique de la langue française), UMR CNRS-Université Nancy2, <http://www.inalf.fr/atilf>.
- Forsgren M. (1978). *La Place de l'adjectif épithète en français contemporain*. Studia Romanica Upsaliensia.
- Glatigny M. (1967). La place des adjectifs épithètes dans deux œuvres de Nerval. *Le Français moderne*, vol. (35) : 201-220.
- Goes J. (1998). *L'Adjectif entre nom et verbe*. Duculot.
- Guiraud P. (1960). *Problèmes et méthodes de la statistique linguistique*. PUF.
- Hug M. (1971). L'adjectif épithète et le complément du nom dans la langue des journalistes. *Études de linguistique appliquée*, vol. (I) : 58-102
- Hug M. (2002). Désambiguïsation automatique d'homographes verbe/nom. In *Actes des JADT 2002*.
- Larsson B. (1994). *La place et le sens des adjectifs épithètes de valorisation positive*. Lund University Press.
- Muller Ch. (1979). *Langue française et linguistique quantitative*. Slatkine.
- Reiner E. (1968). *La Place de l'adjectif épithète en français. Théories traditionnelles et essai de solution*. Braumüller.
- Wilmet M. (1981). La place de l'épithète qualificative en français contemporain. *Revue de linguistique romane*, vol. (177-178) : 17-73.

Annexe 1. Données relevées pour les cooccurrences

Romans	1830-39	1840-49	1850-59	1860-69	1870-79	1880-89	1890-99	1900-09	1910-19	1920-29	1930-39	1940-49	1950-59
nb œuvres	30	93	22	22	27	46	27	30	37	94	66	44	37
nb mots	3045900	6701356	2043168	3140193	2925349	3972500	2241758	2268951	2820948	5671203	5855916	3649738	3170036
ante	42881	91518	24910	38150	36493	52046	31917	26365	34516	65812	68908	41766	34568
post	45243	92391	26640	45766	43479	66894	40472	37444	42743	83580	88851	54105	48083
total	88124	183909	51550	83916	79972	118940	72389	63809	77259	149392	157759	95871	82651
densité	0,0289	0,0274	0,0252	0,0267	0,0273	0,0299	0,0323	0,0281	0,0274	0,0263	0,0269	0,0263	0,0261
post/ante	51,3 %	50,2 %	51,7 %	54,5 %	54,4 %	56,2 %	55,9 %	58,7 %	55,3 %	55,9 %	56,3 %	56,4 %	58,2 %
Essais & traités	1830-39	1840-49	1850-59	1860-69	1870-79	1880-89	1890-99	1900-09	1910-19	1920-29	1930-39	1940-49	1950-59
nb œuvres	15	21	14	11	7	12	28	28	21	44	57	67	94
nb mots	887034	1457685	1066316	814161	451423	745567	1394982	1178964	902699	2949381	4041009	4116951	4500310
ante	11899	18771	14367	10850	4388	10450	16706	13745	10432	33900	48369	44007	52392
post	19135	33354	30865	21242	13904	19983	43247	36863	27160	79876	125037	134943	145073
total	31034	52125	45232	32092	18292	30433	59953	50608	37592	113776	173406	178950	197465
densité	0,0350	0,0358	0,0424	0,0394	0,0405	0,0408	0,0430	0,0429	0,0416	0,0386	0,0429	0,0435	0,0439
post/ante	61,7 %	64,0 %	68,2 %	66,2 %	76,0 %	65,7 %	72,1 %	72,8 %	72,2 %	70,2 %	72,1 %	75,4 %	73,5 %
Mémoires	1830-39	1840-49	1850-59	1860-69	1870-79	1880-89	1890-99	1900-09	1910-19	1920-29	1930-39	1940-49	1950-59
nb œuvres	7	8	4	9	2	1	10	11	6	16	12	8	8
nb mots	419677	1569687	665681	1548194	587676	93691	1549029	696190	895928	1676890	1581605	786357	1074318
ante	5604	21360	9500	19560	7982	1470	19746	9405	10382	17890	19805	8800	11408
post	6484	24924	11236	27412	10002	1993	24383	12218	12765	30521	25836	11201	21277
total	12088	46284	20736	46972	17984	3463	44129	21623	23147	48411	45641	20001	32685
densité	0,0288	0,0295	0,0312	0,0303	0,0306	0,0370	0,0285	0,0311	0,0258	0,0289	0,0289	0,0254	0,0304
post/ante	53,6 %	53,9 %	54,2 %	58,4 %	55,6 %	57,6 %	55,3 %	56,5 %	55,1 %	63,0 %	56,6 %	56,0 %	65,1 %

Annexe 2. Cooccurrences et lemmes employés

Romans		1830-79	1880-1929	1930-59	lemmes	1830-79	1880-1929	1930-59
occurrences								
nb mots		17855966	16975360	12675690	nb mots	17855966	16975360	12675690
anté		226941	200733	138057	anté	1887	2375	2150
post		246570	263134	184757	post	4948	5643	5571
densité a (p. mille)		12,71	11,82	10,89	proportion P	72,4 %	70,4 %	72,2 %
densité p (p. mille)		13,81	15,5	14,58				
proportion P		52,1 %	56,7 %	57,2 %				
test		0			test	0,01		

Essais & traités		1830-79	1880-1929	1930-59	lemmes	1830-79	1880-1929	1930-59
occurrences								
nb mots		4676619	7171593	12658270	nb mots	4676619	7171593	12658270
anté		54849	76057	130141	anté	1156	1705	2144
post		113810	198527	388553	post	3954	5465	6714
densité a (p. mille)		11,73	10,61	10,28	proportion P	77,4 %	76,2 %	75,8 %
densité p (p. mille)		24,34	27,68	30,7				
proportion P		67,5 %	72,3 %	74,9 %				
test		0			test	0,103		

Mémoires		1830-79	1880-1929	1930-59	lemmes	1830-79	1880-1929	1930-59
occurrences								
nb mots		4790915	4911728	3442280	nb mots	4790915	4911728	3442280
anté		60990	54965	36741	anté	1309	1588	1301
post		77348	78356	55828	post	3953	4304	3679
densité a (p. mille)		12,73	11,19	10,67	proportion P	75,1 %	73,0 %	73,9 %
densité p (p. mille)		16,14	15,95	16,22				
proportion P		55,9 %	58,8 %	60,3 %				
test		0			test	0,044		

Annexe 3. Cooccurrences et lemmes employés (lemmes les plus courants)

Romans

occurrences	1830-79	1880-1929	1930-59	lemmes	1830-79	1880-1929	1930-59
nb mots	17855966	16975360	12675690	nb mots	17855966	16975360	12675690
anté	213840	188452	130430	anté	555	569	569
post	184350	192294	129367	post	622	620	620
densité a (p. mille)	11,98	11,10	10,29	proportion P	52,8 %	52,1 %	52,1 %
densité p (p. mille)	10,32	11,33	10,21				
proportion P	46,3 %	50,5 %	49,8 %				
test	0			test	0,925		

Essais & traités

occurrences	1830-79	1880-1929	1930-59	lemmes	1830-79	1880-1929	1930-59
nb mots	4676619	7171593	12658270	nb mots	4676619	7171593	12658270
anté	52151	71056	122937	anté	419	584	669
post	95268	148898	280392	post	804	806	808
densité a (p. mille)	11,15	9,91	9,71	proportion P	65,7 %	58,0 %	54,7 %
densité p (p. mille)	20,37	20,76	22,15				
proportion P	64,6 %	67,7 %	69,5 %				
test	0			test	0		

Mémoires

occurrences	1830-79	1880-1929	1930-59	lemmes	1830-79	1880-1929	1930-59
nb mots	4790915	4911728	3442280	nb mots	4790915	4911728	3442280
anté	56318	49448	33092	anté	345	372	351
post	52003	50740	37959	post	422	421	421
densité a (p. mille)	11,76	10,07	9,61	proportion P	55,0 %	53,1 %	54,5 %
densité p (p. mille)	10,85	10,33	11,03				
proportion P	48,0 %	50,6 %	53,4 %				
test	0			test	0,728		