



HAL
open science

Croiser les corpus calibrés pour faire l'histoire de la langue : le cas de l'anteposition stylistique de l'infinitif et du participe

Pierre Larrivée, Mathieu Goux

► To cite this version:

Pierre Larrivée, Mathieu Goux. Croiser les corpus calibrés pour faire l'histoire de la langue : le cas de l'anteposition stylistique de l'infinitif et du participe. *Corpus*, 2024, La constitution de corpus en diachronie longue, 25, [18 p.]. 10.4000/corpus.8373 . hal-04429739

HAL Id: hal-04429739

<https://hal.science/hal-04429739v1>

Submitted on 31 Jan 2024

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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Croiser les corpus calibrés pour faire l'histoire de la langue : le cas de l'antéposition stylistique de l'infinitif et du participe

Comparing calibrated corpora for the historical study of French: stylistic fronting of infinitives and participles

Pierre Larrivée and Mathieu Goux



Electronic version

URL: <https://journals.openedition.org/corpus/8373>

DOI: [10.4000/corpus.8373](https://doi.org/10.4000/corpus.8373)

ISSN: 1765-3126

Publisher

Bases ; corpus et langage - UMR 6039

Electronic reference

Pierre Larrivée and Mathieu Goux, "Croiser les corpus calibrés pour faire l'histoire de la langue : le cas de l'antéposition stylistique de l'infinitif et du participe", *Corpus* [Online], 25 | 2023, Online since 12 December 2023, connection on 31 January 2024. URL: <http://journals.openedition.org/corpus/8373> ; DOI: <https://doi.org/10.4000/corpus.8373>

This text was automatically generated on January 31, 2024.

The text and other elements (illustrations, imported files) are "All rights reserved", unless otherwise stated.

Croiser les corpus calibrés pour faire l'histoire de la langue : le cas de l'antéposition stylistique de l'infinitif et du participe

Comparing calibrated corpora for the historical study of French: stylistic fronting of infinitives and participles

Pierre Larrivée and Mathieu Goux

1. Introduction

- 1 Le but de cet article est d'illustrer l'utilité de la calibration des corpus selon les types de texte pour faire l'histoire d'une langue. En effet, cette histoire ne peut se faire qu'à partir de textes, mais tous les textes ne sont pas égaux dans le témoignage qu'ils apportent de l'évolution effective. Le but est de réunir des témoignages textuels se rapprochant au mieux de la pratique vernaculaire dont on suppose qu'elle représente la compétence immédiate des locuteurs et serait le lieu où s'initie le changement linguistique. Or, la distance de cette pratique est généralement grande dans la langue littéraire, même quand celle-ci prétend représenter le discours (Guillot *et al.* 2015, Glikman & Mazziotta 2019 ; sur des exemples de discours direct plus conservateur que la narration, voir Pujol 2018 et Imel 2019). C'est pourquoi on est tenté de recourir à des textes non-littéraires, juridiques par exemple.
- 2 La comparaison des textes littéraires et non-littéraires permet ainsi à la fois de suivre une évolution et d'évaluer le témoignage que chaque type donne de cette évolution. C'est ce que fait ce travail, en s'attachant à une construction particulière du français ancien, l'antéposition stylistique. Nous nous concentrons sur l'infinitif antéposé au verbe modal conjugué, comme dans *Autant que faire se peut*, et du participe antéposé à l'auxiliaire, du type *Si dit est* (« s'il est dit »). Le parcours diachronique de chacune de

ces constructions est dessiné dans le corpus. Nous procédons en trois temps. D'abord, nous esquissons les principales propriétés des constructions. Nous dégageons ensuite leur évolution dans les corpus littéraire et légal retenus, pour ensuite explorer des pistes expliquant les divergences entre les deux types de texte. Nous concluons sur les propriétés spécifiques des corpus que permet de documenter notre approche.

2. L'antéposition stylistique

- 3 L'antéposition stylistique est une configuration où un complément, qui se trouve le plus souvent postposé au verbe dans un état de langue donné, s'antépose à lui. Cette configuration est attestée depuis les plus anciens états du français, comme le montre un exemple des *Serments de Strasbourg* également cité par Salvesen (2015 : 311) :

| | | | | | | |
|-----|---|-------|-----------|------------|-------------|---------------------------------------|
| (1) | Cum | om | per dreit | son fradra | saluar dift | (<i>Serment de Strasbourg</i> , 842) |
| | Comme | homme | par droit | son frère | sauver doit | |
| | « comme un homme doit légitimement protéger son frère » | | | | | |

- 4 L'antéposition concerne une vaste gamme de constituants et de compléments (Mathieu 2006, Salvesen 2011, Fischer 2010 : 119, Labelle & Hirschbühler 2014a,b, Labelle 2016). On trouve ici des substantifs, des adjectifs, des adverbes, des participes, des infinitifs, dont certains ont été conservés dans la langue moderne.
- (2) a. **Force** est de constater
 b. **Mal** lui en prit
 c. Comme **bien** vous pensez
 d. **Point trop** n'en faut
 e. Autant que **faire** se peut
 f. Comme **dit** est
- 5 Ce phénomène demeure cependant rare. Dans son survol, Salvesen (2011) relève 411 occurrences de *faire* et *dire* avant *pouvoir*, *devoir* et *vouloir* dans 900 textes littéraires du 9^e au 16^e siècle. Cette rareté est particulièrement manifeste dans les textes en prose (Labelle & Hirschbühler 2014a, Dufter 2018). Elle demeure également énigmatique à bien des égards. Au premier chef, sa motivation est peu claire. L'antéposition stylistique ne semble pas correspondre à un changement de fonction syntaxique. La tradition analytique de la construction équivalente dans les langues germaniques la lie à un contexte avec sujet non-exprimé, que l'élément antéposé servirait à valider (Holmberg 2006). Cependant, comme le montre l'exemple (1), ce n'est pas toujours le cas, spécialement en français ancien avant 1300, et ce ne l'est pratiquement pas dans l'histoire d'une langue romane comme le vénitien (Poletto & Pinzin, en préparation). Les motivations informationnelles, qu'elles soient alléguées pour l'élément antéposé ou pour la proposition entière (pour le français, Rahn 2016, Dufter 2018, Imel 2019 ; au-delà, voir la synthèse de Holmberg 2006), restent équivoques.
- 6 Pour contribuer à la compréhension des motivations de la construction, nous estimons utile de tracer son évolution diachronique, ce à quoi s'emploie la prochaine section.

3. Méthode et résultats

- 7 Pour explorer l'antéposition stylistique dans l'histoire du français, nous considérons sa fréquence au sein de textes de deux types, qui ont en commun de relever de la prose. En effet, les contraintes positionnelles que peut faire peser le vers sur l'ordre des mots ne le désignent pas comme matériel principal d'investigation. Le premier type de prose appartient à l'ensemble littéraire du roman de chevalerie. Comme ces textes ont été très utilisés dans la recherche sur la syntaxe diachronique du français ancien, ils étaient tout désignés pour retenir notre attention. Nous avons sélectionné, via la base de données Frantext¹, la *Queste del saint Graal* datée de 1225-1230, le *Roman de Bérinus* daté de 1370, et le *Jouvencel* de Jean de Bueil de 1461. Ces textes, d'une *scripta* dialectalement peu marquée selon la Base de Français Médiéval², sont annotés en parties du discours et permettent une recherche fondée sur ce paramètre.

Tableau 1. Textes littéraires retenus de Frantext

| Source | Date | Titre | Nombre de propositions | Nombre de mots |
|----------|------|------------------|------------------------|----------------|
| Frantext | 1225 | <i>Queste</i> | 17 058 | 107 677 |
| Frantext | 1370 | <i>Bérinus</i> | 23 825 | 148 347 |
| Frantext | 1461 | <i>Jouvencel</i> | 15 263 | 122 042 |
| | | Total | 56 146 | 378 066 |

- 8 Comme les bases de données ne proposent pas de texte de ce type pour le 16^e siècle, nous avons choisi de retenir à des fins de comparaison un texte de 1577, *La Plaisante et amoureuse histoire du Chevalier Doré et de la pucelle surnommée Cœur d'acier*, dont nous avons, à partir du pdf disponible sur Gallica³, importé les 100 premières pages. Cette sélection représente 26 231 mots graphiques, pour environ 3 700 propositions (en estimant que chaque proposition est composée en moyenne de 7 mots, en accord avec les données des textes du tableau 1).
- 9 Le second type de prose relève du domaine non littéraire des textes légaux. Comme on l'a mentionné, ces textes sont généralement moins conservateurs que les textes littéraires (Goux & Larrivée 2020), ce que cette étude permettra de vérifier. Un premier ensemble est composé d'une série de procès tirés du corpus MICLE (Larrivée & Poletto 2023). Ce corpus réunit des textes légaux du 13^e au 17^e siècles et issus à la même région, la Normandie, avec, pour la période ancienne, des textes de la zone contiguë anglo-normande d'où émane une large proportion des premiers textes en français.

Tableau 2. Procès utilisés du corpus MICLE

| Source | Date | Titre | Nombre de propositions | Nombre de mots |
|--------|-----------|------------------------------------|------------------------|----------------|
| MICLE | 1270-1287 | <i>Anglo-Norman Year Book (I)</i> | 6 615 | 46 661 |
| MICLE | 1292 | <i>Anglo-Norman Year Book (II)</i> | 4 043 | 36 728 |

| | | | | |
|-------|------|---|--------|---------|
| MICLE | 1431 | <i>Minute française du procès de Jeanne d'Arc</i> | 6 241 | 42 494 |
| MICLE | 1460 | <i>Manuel de Flambart</i> | 569 | 7 927 |
| MICLE | 1591 | <i>Fille possédée</i> | 4 543 | 30 991 |
| | | Total | 22 011 | 164 801 |

- 10 Étant donné le faible taux d'usage des constructions, nous avons choisi d'ajouter des coutumiers du corpus ConDé (Larrivée & Goux 2021). Si les textes de ce dernier ne sont pas annotés syntaxiquement, ils le sont en partie du discours, ce qui permet la recherche des antépositions stylistiques comme pour les textes de Frantext.

Tableau 3. Coutumiers utilisés du corpus ConDé

| Source | Date | Titre | Nombre de propositions | Nombre de mots |
|--------|---------------------|---|------------------------|----------------|
| ConDé | Fin 13 ^e | <i>Très ancien Coutumier de Normandie</i> | 2 070 | 19 958 |
| ConDé | 1300 | <i>Grand Coutumier de Normandie</i> | 8 195 | 68 791 |
| ConDé | 1386 | <i>Instrucions et enseignemens</i> | 1 772 | 14 277 |
| ConDé | 1539 | <i>Le Rouillé</i> | 38 162 | 334 717 |
| ConDé | 1578 | <i>Terrien</i> | 34 176 | 541 790 |
| ConDé | 1614 | <i>Bérault</i> | 35 604 | 639 060 |
| | | Total | 119 979 | 1 783 394 |

- 11 Les antépositions de participes et d'infinitifs sont extraites de ces textes à l'aide de différentes requêtes CQL, que nous reproduisons en annexe pour chacune des trois bases de données (Frantext, MICLE et ConDé). Les occurrences recueillies ont été examinées manuellement, pour éliminer les faux-positifs et les erreurs d'étiquetages résiduelles. Le nombre d'attestations des participes et des infinitifs antéposés est présenté pour chaque groupe de texte.
- 12 Nous commençons par l'ensemble littéraire. Le tableau suivant donne le nombre d'occurrences des participes antéposés à un auxiliaire par rapport au nombre de participes postposés et de propositions (calculées selon le nombre de verbes conjugués, et estimées pour le texte de 1577).

Tableau 4. Participes antéposés dans des textes littéraires de Frantext

| Source | Date | Titre | Participes antéposés | Proportions au regard des participes postposés | Proportions au regard des propositions |
|--------|------|-------|----------------------|--|--|
|--------|------|-------|----------------------|--|--|

| | | | | | |
|----------|------|------------------|----|--------------------|------------------------|
| Frantext | 1225 | <i>Queste</i> | 13 | 0,6% (13 / 2 169) | 0,08% (13 / 17 058) |
| Frantext | 1370 | <i>Bérinus</i> | 5 | 0,18% (5 / 2 803) | 0,02% (5 / 23 825) |
| Frantext | 1461 | <i>Jouvencel</i> | 10 | 0,48% (10 / 2 067) | 0,05% (10 / 15 263) |
| Gallica | 1577 | <i>Chevalier</i> | 1 | 0,24% (1 / 424) | 0,03% (1 / env. 3 700) |
| | | Total | 29 | 0,38% (29 / 7 463) | 0,05% (29 / 56 146) |

13 Le nombre d'infinitifs antéposés est fourni dans le tableau 5.

Tableau 5. Infinitifs antéposés dans des textes littéraires de Frantext

| Source | Date | Titre | Infinitifs antéposés | Proportions au regard des infinitifs postposés | Proportions au regard des propositions |
|----------|------|------------------|----------------------|--|--|
| Frantext | 1225 | <i>Queste</i> | 0 | 0% (/ 1 815) | 0% (/ 18 179) |
| Frantext | 1370 | <i>Bérinus</i> | 2 | 0,4% (2 / 492) | 0,008% (2 / 23 825) |
| Frantext | 1461 | <i>Jouvencel</i> | 12 | 1,2% (12 / 994) | 0,08% (12 / 15 263) |
| Gallica | 1577 | <i>Chevalier</i> | 1 | 0,5% (1 / 199) | 0,03% (1 / env. 3 700) |
| | | Total | 15 | 0,43% (15 / 3 500) | 0,02% (15 / 60 967) |

14 Dans cet ensemble, les participes sont deux fois plus nombreux que les infinitifs. Le parcours demeure celui d'une lente baisse, de moins d'une occurrence sur mille propositions, à moins d'une occurrence sur deux mille, voire sur trois mille pour les infinitifs dans le dernier texte. Un taux d'usage situé sous la courbe attendue se retrouve dans le texte du 14^e siècle pour le participe, et celui du 13^e siècle pour l'infinitif.

15 Nous considérons ensuite les textes légaux, et en premier lieu les procès. Le nombre d'occurrences de participes antéposés est donné dans le tableau 6.

Tableau 6. Participes antéposés dans des procès du corpus MICLE

| Source | Date | Titre | Participes antéposés | Proportions au regard des participes postposés | Proportions au regard des propositions |
|--------|-----------|------------------------------------|----------------------|--|--|
| MICLE | 1270-1287 | <i>Anglo-Norman Year Book (I)</i> | 2 | 3,13% (2 / 64) | 0,03% (2 / 6 615) |
| MICLE | 1292 | <i>Anglo-Norman Year Book (II)</i> | 27 | 4,3% (27 / 627) | 0,7% (27 / 4 043) |
| MICLE | 1431 | | 1 | 0,08% (1 / 1 209) | 0,15% (1 / 6 241) |

| | | | | | |
|-------|------|---|----|--------------------|---------------------|
| | | <i>Minute française du procès de Jeanne d'Arc</i> | | | |
| MICLE | 1460 | <i>Manuel de Flambart</i> | 8 | 4,62% (8 / 173) | 1,4% (8 / 569) |
| MICLE | 1591 | <i>Fille possédée</i> | 0 | 0% (/ 1 259) | 0% (/ 2 654) |
| | | Total | 38 | 1,14% (38 / 3 332) | 0,19% (38 / 20 122) |

16 Le tableau 7 fournit les résultats pour les infinitifs antéposés.

Tableau 7. Infinitifs antéposés dans des procès du corpus MICLE

| Source | Date | Titre | Infinitifs antéposés | Proportions au regard des infinitifs postposés | Proportions au regard des propositions |
|--------|-----------|---|----------------------|--|--|
| MICLE | 1270-1287 | <i>Anglo-Norman Year Book (I)</i> | 3 | 0,9% (3 / 336) | 0,05% (3 / 6 615) |
| MICLE | 1292 | <i>Anglo-Norman Year Book (II)</i> | 14 | 3,72% (14 / 376) | 0,4% (14 / 4 043) |
| MICLE | 1431 | <i>Minute française du procès de Jeanne d'Arc</i> | 0 | 0% (/ 378) | 0% (/ 6 241) |
| MICLE | 1460 | <i>Manuel de Flambart</i> | 7 | 12,5% (7 / 56) | 1,2% (7 / 569) |
| MICLE | 1591 | <i>Fille possédée</i> | 0 | 0% (/ 141) | 0% (/ 2654) |
| | | Total | 24 | 1,86% (24 / 1 287) | 0,12% (24 / 20 122) |

17 Les procès donnent deux fois plus de participes que d'infinitifs, ce qui représente trois à quatre fois plus d'occurrences que le corpus littéraire, à la fois globalement et pour la plupart des époques. Si l'on met de côté le premier texte où les nombres sont faibles et le texte de 1460 où ils sont élevés, la courbe de changement suit un déclin régulier de moins d'une occurrence sur cent à une occurrence pour sept cents à aucune pour le participe, et d'une occurrence pour cinq cents à aucune pour l'infinitif.

18 Passons en second lieu aux textes coutumiers. Les taux d'usage des participes sont présentés dans le tableau 8.

Tableau 8. Participes antéposés dans les coutumiers du corpus ConDÉ

| Source | Date | Titre | Participes antéposés | | |
|--------|------|-------|----------------------|--|--|
| | | | | | |

| | | | | Proportions au regard des participes postposés | Proportions au regard des propositions |
|-------|------------------------|--|-----|--|---|
| ConDÉ | Fin 13 ^e | Très ancien Coutumier de Normandie | 1 | 0,21% (1 / 479) | 0,05% (1 / 2 070) |
| ConDÉ | 1300 | Grand Coutumier de Normandie | 6 | 0,3% (6 / 1 912) | 0,07% (6 / 8 195) |
| ConDÉ | 1386 | Instrucions et enseignemens | 10 | 4,63% (10 / 216) | 0,6% (10 / 1 772) |
| ConDÉ | 1539 | Le Rouillé | 97 | 1,54% (97 / 6 284) | 0,25% (97 / 38 162) |
| ConDÉ | 1578 | Terrien | 69 | 0,68% (69 / 10 162) | 0,22% (69 / 34 176) |
| ConDÉ | 1614 | Bérault | 0 | 0% (/ 8 959) | 0% (/ 35 604) |
| | | Total | 183 | 0,65% (183 / 28 012) | 0,16% (183 / 117 909) |

19 Le dernier tableau donne les nombres d'infinitifs antéposés.

Tableau 9. Infinitifs antéposés dans les coutumiers du corpus ConDÉ

| Source | Date | Titre | Infinitifs antéposés | Proportions au regard des infinitifs postposés | Proportions au regard des propositions |
|--------|------------------------|--|-------------------------|--|--|
| ConDÉ | Fin 13 ^e | Très ancien Coutumier de Normandie | 0 | 0% (/ 201) | 0% (/ 2 070) |
| ConDÉ | 1300 | Grand Coutumier de Normandie | 3 | 0,21% (3 / 1 414) | 0,04% (3 / 8 195) |
| ConDÉ | 1386 | Instrucions et enseignemens | 11 | 6,18% (11 / 178) | 0,6% (11 / 1 772) |
| ConDÉ | 1539 | Le Rouillé | 5 | 0,08% (5 / 6 416) | 0,01% (5 / 38 162) |
| ConDÉ | 1578 | Terrien | 71 | 1,83% (71 / 3 882) | 0,2% (71 / 34 176) |
| ConDÉ | 1614 | Bérault | 7 | 0,2% (7 / 3 434) | 0,02% (7 / 35 604) |
| | | Total | 97 | 0,62% (97 / 15 525) | 0,08% (97 / 119 979) |

20 Comme dans les autres corpus, les participes sont deux fois plus nombreux que les infinitifs ; et comme dans les procès, les taux sont plus importants que dans les textes littéraires, de deux à quatre fois. La courbe du changement des participes va d'une

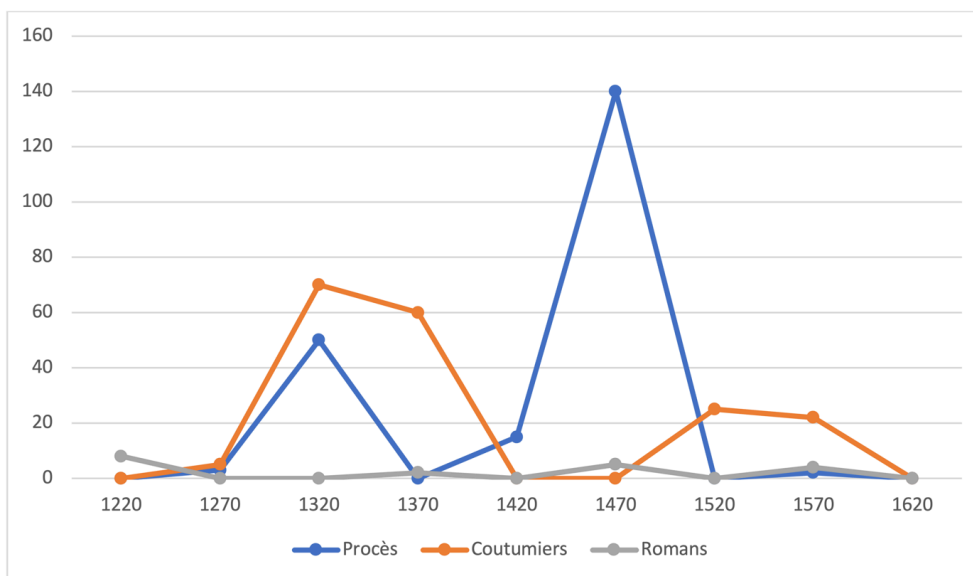
occurrence sur deux cents, à une occurrence sur quatre cents, à rien, si on met de côté les textes du 13^e siècle. Pour les infinitifs, cette courbe passe d'environ une occurrence sur deux cents, à une occurrence sur cinq cents, à une occurrence sur cinq mille, avant de disparaître, si on met de côté le faible taux des deux premiers textes et du 1539.

21 Résumons les grandes tendances se dégageant de cette étude comparative :

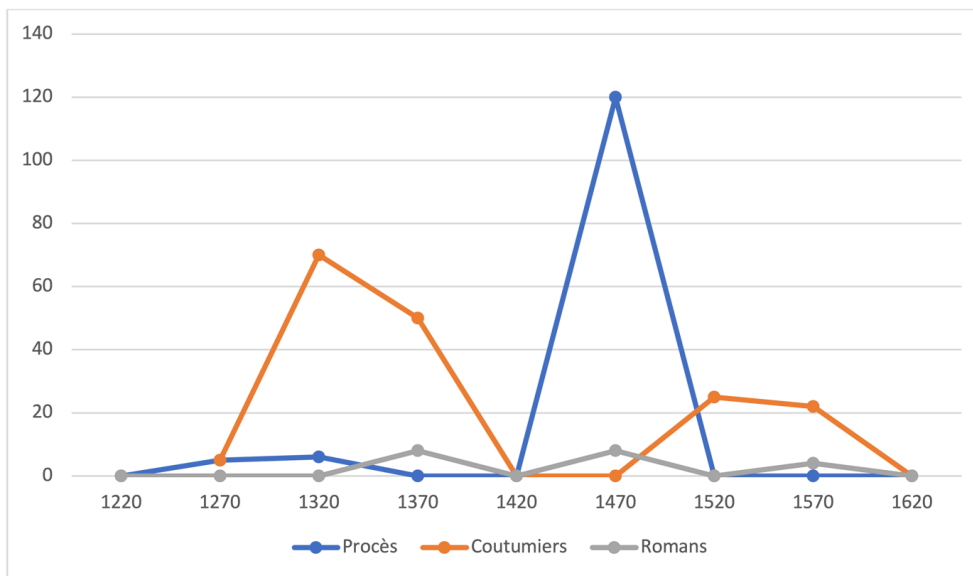
- Les textes littéraires et les textes coutumiers montrent que les taux d'emploi du participe antéposé sont deux fois plus importants que les infinitifs antéposés.
- Les textes légaux donnent des proportions jusqu'à quatre fois plus importantes que les textes littéraires.
- Les tout premiers textes (pour l'infinitif littéraire, le participe des procès, et pour les coutumiers) donnent des taux en décalage avec ceux subséquents de la fin 13^e / début 14^e, indiquant une remontée en moyen français.
- Les faibles taux des constructions expliquent certaines variations locales liées à des textes particuliers (notamment le manuel de Flambart).
- Les constructions à l'étude suivent une courbe de déclin, avec une disparition générale autour de 1620.

22 La remontée des emplois en moyen français et la surreprésentation de la construction sont illustrées par les graphiques 1 et 2, où les proportions sont données sur 10 000 occurrences, et les textes rattachés à la date la plus proche des vingtième et soixante-dixième années du siècle pertinent.

Graphique 1. Évolution des participes antéposés dans les trois types de texte



Graphique 2. Évolution des infinitifs antéposés dans les trois types de texte



- 23 Les raisons de ces deux observations sont explorées dans la prochaine section.

4. Une caractéristique latinisante des textes légaux ?

- 24 Les données sur l'antéposition du participe et de l'infinitif dans deux corpus calibrés sont surprenantes à plusieurs titres. Tout d'abord, on constate un taux d'usage beaucoup plus élevé dans les textes légaux que dans les textes littéraires, ce qui va à l'encontre du constat selon lequel le premier type de texte est généralement moins conservateur que le second. D'autre part, toujours dans les textes légaux, une remontée des taux se manifeste dès les débuts de la période étudiée jusqu'au début du 16^e.
- 25 Ces deux observations interrogent. Qu'est-ce qui amènerait les greffiers des procès et les auteurs de coutumiers à utiliser fréquemment une construction qui semble rare dans la pratique littéraire ? Une influence possible serait celle du latin. La chose a été soulignée pour la survivance de la construction par Ramat (1987 : 150-153) et Berta (2008). En latin, langue à structure OV, l'antéposition du participe et de l'infinitif est l'ordre par défaut. Une recherche rapide dans *De Bello Gallico* issu du *treebank* PROIEL⁴, montre que l'antéposition représente l'état normal des choses. Ce texte donne 163 participes antéposés contre 81 postpositions, et 134 infinitifs antéposés contre 11 postpositions. En ce sens, un examen des procès en latin dans *l'Anglo-Norman Year Book* de 1292 montre que toutes les constructions pertinentes voient le participe et l'infinitif antéposés. Étant donné que les auteurs de textes légaux avaient une connaissance du latin qui n'était plus nécessairement le cas pour les auteurs littéraires, pourrait-on voir dans la présence des constructions antéposées un impact de leur pratique de cette langue ?
- 26 Il ne s'agit pas d'alléguer un calque direct du latin, puisque la construction en français a une distribution différente. Pour n'en citer qu'une seule, la construction française se retrouve plutôt dans les subordinées (3), alors que la latine se retrouve tout aussi bien en principale (4)⁵.

(3) Et pur coe ky *ateynt est* par l'assyse (« Et pour ce qui est atteint par l'assise... », ANYB 1292)

- (4) Acriter in eo loco *pugnatum est* (« En ce lieu-là on combattit durement » [litt., « en ce lieu-là, il fut combattu durement »], *De Bello Gallico*)
- 27 Mais nous pouvons également penser à une influence due au contact des langues dans un milieu professionnel. On retrouve ici la notion, paradoxale, d'emprunt de fréquence : on parle en français « d'anglicisme de fréquence » non pas quand le français emprunte un terme à l'anglais, mais quand le terme français cognat à l'anglais se trouve plus utilisé à cause du contact des langues. Un exemple donné par Claude Poirier est celui de « breuvage », existant en français, mais plus fréquent au Québec, à cause de l'anglais *beverage*, dans un environnement où le contact avec cette langue est régulier (Poirier 1989). Ainsi pourrait-il en être de cette construction du français dans les textes légaux, qui marquerait la familiarité avec le latin de la part des spécialistes du droit⁶.
- 28 Pour explorer l'hypothèse d'une influence du contact avec le latin pour les expressions du français, on peut considérer les expressions particulières qui la représentent. On pourrait supposer que certaines expressions formulaires de la construction seraient propres aux textes légaux, ce qui expliquerait sa prévalence. Nous avons donc relevé dans ces textes les expressions formulaires, définies comme toutes les séquences où le même participe ou infinitif lexical précède le même verbe conjugué lexical (quelle que soit la conjugaison donc), à l'instar de l'exemple (3) *supra* ou des exemples (5) et (6) suivants :
- (5) a. par paiant ou par delivrant *comme dit est* (*Instructions* 1386)
 b. nonobstant *ce que dict est* (*Rouillé* 1539)
 c. et sera ouy *comme dit est* (*Terrien* 1578)
- (6) a. par quey ceo *ferre ne voylent* (« par quoi [ils] ne veulent pas faire cela », ANYB 1292)
 b. ma dicte preuve faicte par la maniere que *faire le voudray* (*Flambart* 1460)
 c. et collationer ledit extrait ou copie, si *faire le veut* (*Terrien* 1578)
- 29 Les résultats quantitatifs sont donnés dans les tableaux suivants.

Tableau 10. Participes antéposés et emplois formulaires dans les procès du corpus MICLE

| Source | Date | Titre | Participes antéposés | Emplois formulaires |
|--------|-----------|------------------------------------|----------------------|---------------------|
| MICLE | 1270-1287 | <i>Anglo-Norman Year Book (I)</i> | 2 | 1 |
| MICLE | 1292 | <i>Anglo-Norman Year Book (II)</i> | 27 | 18 |
| MICLE | 1431 | <i>Jeanne d'Arc</i> | 1 | 1 |
| MICLE | 1460 | <i>Flambart</i> | 8 | 3 |
| MICLE | 1591 | <i>Fille possédée</i> | 0 | 0 |
| | | Total | 38 | 23 (61%) |

Tableau 11. Infinitifs antéposés et emplois formulaires dans les procès du corpus MICLE

| Source | Date | Titre | Infinitifs antéposés | Emplois formulaires |
|--------|------|-------|----------------------|---------------------|
|--------|------|-------|----------------------|---------------------|

| | | | | |
|-------|-----------|------------------------------------|----|----------|
| MICLE | 1270-1287 | <i>Anglo-Norman Year Book (I)</i> | 3 | 0 |
| MICLE | 1292 | <i>Anglo-Norman Year Book (II)</i> | 14 | 8 |
| MICLE | 1431 | <i>Jeanne d'Arc</i> | 0 | 0 |
| MICLE | 1460 | <i>Flambart</i> | 7 | 3 |
| MICLE | 1591 | <i>Fille possédée</i> | 0 | 0 |
| | | Total | 24 | 11 (46%) |

Tableau 12. Participes antéposés et emplois formulaires dans les coutumiers du corpus ConDÉ

| Source | Date | Titre | Participes antéposés | Emplois formulaires |
|--------|---------------------|--------------------|----------------------|---------------------|
| ConDÉ | Fin 13 ^e | <i>TAC</i> | 1 | 0 |
| ConDÉ | 1300 | <i>GC</i> | 6 | 0 |
| ConDÉ | 1386 | <i>Instrucions</i> | 10 | 6 |
| ConDÉ | 1539 | <i>Le Rouillé</i> | 97 | 94 |
| ConDÉ | 1578 | <i>Terrien</i> | 69 | 53 |
| ConDÉ | 1614 | <i>Bérault</i> | 0 | 0 |
| | | Total | 183 | 153 (83%) |

Tableau 13. Infinitifs antéposés et expressions formulaires dans des coutumiers du corpus ConDÉ

| Source | Date | Titre | Infinitifs antéposés | Emplois formulaires |
|--------|---------------------|--------------------|----------------------|---------------------|
| ConDÉ | Fin 13 ^e | <i>TAC</i> | 0 | 0 |
| ConDÉ | 1300 | <i>GC</i> | 3 | 1 |
| ConDÉ | 1386 | <i>Instrucions</i> | 11 | 6 |
| ConDÉ | 1539 | <i>Le Rouillé</i> | 5 | 1 |
| ConDÉ | 1578 | <i>Terrien</i> | 71 | 63 |
| ConDÉ | 1614 | <i>Bérault</i> | 7 | 5 |
| | | Total | 97 | 76 (78,4%) |

- 30 On ne peut que remarquer, et surtout pour les coutumiers, que le retrait des occurrences formulaires change la donne et met les coutumiers au niveau des textes littéraires.
- 31 Les expressions formulaires peuvent se diviser entre celles qui sont communes avec le corpus littéraire :

Tableau 14. Expressions formulaires présentes dans tous les textes

| Expression (n.=) | Textes (n.=) |
|------------------|---|
| Dit est (125) | <i>Jeanne d'Arc</i> (1), <i>instrucions</i> (4), <i>Le Rouillé</i> (94), <i>Terrien</i> (26) |
| Faire peut (59) | <i>GC</i> (1), <i>instrucions</i> (3), <i>Le Rouillé</i> (1), <i>Terrien</i> (50), <i>Bérault</i> (4) |
| Faire veut (9) | <i>ANYB II</i> (2), <i>Flambard</i> (3), <i>Le Rouillé</i> (1), <i>Terrien</i> (3) |
| Fait est (3) | <i>Flambard</i> (3) |
| Faire doit (2) | <i>Instrucions</i> (1), <i>Terrien</i> (10), <i>Bérault</i> (1) |
| Passé est (2) | <i>ANYB II</i> (2) |

- 32 et celles qui ne s'y retrouvent pas :

Tableau 15. Expressions formulaires absentes des textes littéraires

| Expression (n.=) | Textes (n.=) |
|----------------------|------------------------------------|
| Jugé est (12) | <i>GC</i> (1), <i>Terrien</i> (11) |
| Atteint est (8) | <i>ANYB II</i> (8) |
| Trouvé est (7) | <i>ANYB II</i> (7) |
| Contenu est (2) | <i>Terrien</i> (2) |
| Clamer peut (3) | <i>ANYB II</i> (3) |
| Vocher peut (2) | <i>ANYB II</i> (2) |
| Ordonné est (2) | <i>Terrien</i> (15) |
| Agardé est (1) | <i>ANYB II</i> (1) |
| Avérer doit (1) | <i>ANYB II</i> (1) |
| Requis est (1) | <i>Terrien</i> (1) |
| Défendu est (1) | <i>Terrien</i> (1) |
| Descendu lui est (1) | <i>ANYB I</i> (1) |

| | |
|------------------|-----------------|
| Prouver veut (1) | Instrucions (1) |
|------------------|-----------------|

- 33 Ces chiffres montrent que les expressions formulaires se retrouvent plus dans les coutumiers que dans les procès, plus dans les coutumiers tardifs que dans les coutumiers précoces, et plus avec des expressions communes qu'avec des expressions spécifiques. Nous interprétons ces données comme suggérant une influence du latin : les expressions formulaires communes du français sont également bien attestées en latin, les commentateurs de coutumiers sont des praticiens de cette langue à une époque où les auteurs littéraires ne le sont plus nécessairement, et l'usage de ces expressions permet aux auteurs de démontrer cette compétence.

5. Conclusion

- 34 Notre étude a montré que l'antéposition stylistique, plus fréquente avec le participe qu'avec l'infinitif, subit une recrudescence en moyen français, et une sur-représentation dans les textes légaux, ce qui est inattendu puisque ces derniers sont généralement moins conservateurs que les textes littéraires. Cette recrudescence comme la sur-représentation sont cependant dues à l'augmentation spectaculaire des emplois formulaires, en particulier dans les coutumiers. Nous avons proposé que cette augmentation était vraisemblablement un marqueur de l'appartenance professionnelle des auteurs à un domaine où, jusqu'à une époque tardive, la connaissance du latin était bien établie ; or, il s'agit d'une langue dans laquelle ces antépositions constituent l'option syntaxique par défaut⁷.
- 35 Pour conclure, l'usage des corpus calibrés apparaît utile pour départager ce qui relève de la langue vernaculaire où se déroule l'évolution effective, et ce qui relève du type des textes utilisés pour s'en faire une idée. En particulier, elle montre que les textes légaux ne constituent pas une catégorie homogène, et les sous-types de textes légaux, procès et coutumiers, peuvent offrir un témoignage différencié des constructions considérées. La démonstration fournie ici saura, nous le souhaitons, inspirer l'adoption de la calibration des textes pour des travaux diachroniques ultérieurs.

BIBLIOGRAPHY

- Amatuzzi A., Ayres-Bennett W., Gerstenberg A., Schøsler L. & Skupien-Dekens C. (2020). « Changement linguistique et périodisation du français (pré)classique : deux études de cas à partir des corpus du RCFC », *Journal of French Language Studies* 30(3) : 301-326.
- Berta T. (2008). « Sobre la evolución de las construcciones de participio y de infinitivo », in Concepción Company Company et J. G. Moreno de Alba (éd.) *Actas del VII Congreso Internacional de Historia de la Lengua Española (Mérida, Yucatán, 4-8 de septiembre de 2006)*. Madrid : Arco/Libros, vol. 1, 401-411.

- Brethauer I. (2021). « De la formalisation des savoirs à la cartularisation. Le manuel et les archives de Guillaume Flambart (Rouen, 1460-1470) », *Rives méditerranéennes* 62 : 103-128.
- Bybee J. (2002). « Main clauses are innovative, subordinate clauses are conservative. Consequences for the nature of constructions », J. Bybee et M. Noonan (éd.) *Complex Sentences in Grammar and Discourse. Essays in honor of Sandra A. Thompson*. Amsterdam : John Benjamins Publishing Company, 1-17.
- Carlier A. & Combettes B. (2015). « Typologie et caractérisation morphosyntaxique : du latin au français moderne », *Langue française* 187(3) : 15-58.
- Combettes B. (2006). « L'analyse thème/rhème dans une perspective diachronique », *LINX* 55, 75-90.
- Danckaert L. (2017). *The development of latin clause structure. A study of the extended verb phrase*. Oxford : Oxford University Press.
- De Andrade A. (2018). « Aboutness Topics in Old and Middle French : A corpus-based study on the fate of V2 », *Revue canadienne de Linguistique* 63(2) : 194-220.
- Dufter A. (2018). « On participle fronting in Old and Middle French », *Communication orale, Universität München*.
- Fischer S. (2010). *Word Order Change as a Source of Grammaticalisation*. Amsterdam : John Benjamin.
- Franco I. (2017). « Stylistic fronting in Old Italian : A phase-based analysis », *Language* 93(3) : 114-151.
- Glikman J. & Mazziotta N. (2019). « Oral représenté et narration en ancien français. Spécificités syntaxiques dans trois textes de genres distincts », *Linx* 78, <https://doi.org/10.4000/linx.3151>.
- Goux M. & Larrivée P. (2020). « Expression et positif du sujet en ancien français : le rôle de la personne pronominale », *CMLF 2020*, <https://doi.org/10.1051/shsconf/20207803002>.
- Guillot C., Heiden S., Lavrentiev A. & Pincemin B. (2015). « L'oral représenté dans un corpus de français médiéval (9^e-15^e) : approche contrastive et outillée de la variation diasystémique », J. Lindschouw et al. (éd.) *Les variations diasystématiques et leurs interdépendances dans les langues romanes. Actes du colloque DIA II à Copenhague (19-21 nov. 2012)*, *Travaux de Linguistique Romane*, 15-27.
- Holmberg A. (2006). « Stylistic fronting », M. Everaert et H. van Riemsdijk (dir.) *The Blackwell Companion to Syntax*. Oxford : Blackwell, 532-565.
- Imel B. (2019). *Sa nature proveir se volt : A new examination of leftward stylistic displacement in Medieval French through textual domain, information structure, and Oral Représenté*. Thèse de doctorat, University of California Berkeley.
- Labelle M. (2016). « Participle fronting and clause structure in Old and Middle French », in C. Tortora, M. den Dikken, I. L. Montoya et T. O'Neill (éd.) *Romance Languages and Linguistic Theory* 9 : 213-232. Amsterdam : John Benjamins.
- Labelle M. & Hirschbühler P. (2014a). « Déplacement stylistique à gauche de verbes non conjugués en ancien et en moyen français », *Corpus* 13 : 191-219.
- Labelle M. & Hirschbühler P. (2014b). « Y avait-il antéposition stylistique en ancien français ? », *CMLF 2014*, <https://doi.org/10.1051/shsconf/20140801129>.
- Larrivée P. (2019). « To be or not to be informational : Preverbal complements in Medieval French V2 configurations », *Glossa* 4(1) : 1-19.

Larrivé P. & Goux M. (dir.). (2021). *Corpus ConDÉ*, version Bêta 1.0, Caen, CRISCO (EA 4255) et PDN (MRSH) de l'Université de Caen. <https://pdn-lingua.unicaen.fr/coutumiers/conde/accueil.html>.

Larrivé P. & Poletto C. (dir.). (2023). *Corpus MICLE*, Caen (France) / Francfort (Allemagne), CRISCO (EA 4255) / Institut für Romanische Sprachen und Literaturen.

Ledgeway A. (2018). « On the decline of the edge-fronting from Latin to Romance », in A. M. Martins & A. Cardoso (dir.) *Word Order Change*. Oxford : Oxford Academic, 264-278.

Mathieu É. (2006). « Stylistic Fronting in Old French », *Probus* 18 : 219-266.

Octavio de Toledo A. (2016). « Enseñanzas del cambio fracasado : trayectoria y estela de una periphrasis fugaz (infinitive + tener) », *Cuadernos de Linguística* 3(1) : 119-181.

Olivier M. (2022). « Diachronie de la proclise et de l'enclise avec l'infinitif en français médiéval (12^e-15^e s.) », *Studia linguistica Romanica* 8 : 10-35.

Poirier C. (1989). « Le français et les québécoismes », *La Presse*, 25 mai 1989, p. B3 (Section Opinions).

Poletto C. & Pinzin F. (en préparation). « What's left of the V2 high tide : Infinitival Anteposition ».

Pujol i Campeny A. (2018). *Word Order in Old Catalan*. PhD dissertation, Cambridge.

Rahn C. A. (2016). *At the Left Edge : Fronting in Medieval French Embedded Clauses*. Thèse de doctorat, Université de Caen et Universität Konstanz.

Ramat P. (1987). « An example of reanalysis : periphrastic forms in the Romance languages verb system », in P. Ramat (éd.) *Linguistic Typology*. Berlin/New York : Mouton de Gruyter, 141-164.

Salvesen C. M. (2011). « Stylistic Fronting and Remnant movement in Old French », in J. Bern, H. Jacobs et T. Scheer (dir.) *Romance Languages and Linguistic Theory. Selected Papers from 'Going Romance' Nice 2009*. Amsterdam : Benjamins, 323-342.

APPENDIXES

Requêtes CQL

Voici les requêtes CQL utilisées pour la collecte des occurrences d'antéposition :

| Base de données | Requête CQL | |
|-----------------|--------------------|---|
| Frantext | Participe antéposé | <code>([pos="VPP"])([{}0,2])([lemma="être avoir"%cd])</code> |
| | Infinitif antéposé | <code>([pos="VINF"])([{}0,2])([lemma="pouvoir vouloir devoir"%cd])</code> |
| MICLE | Participe antéposé | <code>[udpos!="AUX"][prpos="Ge"][udpos!="PUNCT"]{0,2} [udpos="AUX"][prpos!="Ge"]</code> |
| | Infinitif antéposé | <code>[uppos="MDX AX EX VX"]{0,2}[uppos="MDJ MDX MDPP"]</code> |

| | | |
|-------|--------------------|--|
| ConDÉ | Participe antéposé | [pos!="Vu[cn]"][pos="Ge"][pos!="F.*"]{0,2}[pos="Vu[cn]"][pos!="Ge"] |
| | Infinitif antéposé | [pos="v[vu]n"]{0,2}[lemma="POUVOIR VOULOIR DEVOIR"] |

Nous attirons l'attention, notamment, sur les intervalles $\{0,2\}$ qui nous permettent de trouver les occurrences où un élément interpolé, un pronom par exemple, sépare l'auxiliaire et le complément auxilié. Cette stratégie évite le silence lors de la collecte des occurrences, quitte à ce qu'il faille effectivement trier manuellement le bruit. Les occurrences de postposition ont été obtenues en permutant l'ordre dans lequel les valeurs correspondantes étaient recherchées.

NOTES

1. <https://www.frantext.fr>
2. <http://bfm.ens-lyon.fr>
3. <ark:/12148/bpt6k3135566>
4. https://universaldependencies.org/treebanks/la_proiel/index.html
5. Sur la distinction principale/subordonnée pour l'évolution, voir Bybee (2002).
6. Les taux élevés dans le *Manuel* de Flambart, un nobliau désireux de s'insérer dans la profession juridique où il n'a pas été formé, vont dans le sens d'un marqueur d'appartenance professionnelle (Bretthauer 2021).
7. Cela ne semble cependant pas être le cas dans toutes les langues romanes. Octavio de Toledo (2016 : 146), dans une étude sur l'espagnol, n'observe pas la recrudescence tardive que nous avons relevée, sans que l'on ne puisse considérer que les spécialistes du droit espagnols connaissaient moins bien le latin que leurs homologues français. Une analyse comparative au sein de toutes les langues romanes est cependant encore à mener.

ABSTRACTS

We illustrate the usefulness of generic corpus calibration for the diachronic analysis of the French language. We explore the case of stylistic fronting of the infinitive and the participle. We conduct this study in three corpora consisting respectively of literary texts, Norman custom laws and trials, from the Old French to the Classical French period. The generic calibration reveals that the evolution of these fronting phenomena differs according to the type of text. We also show that legal texts display an over-representation of these constructions, in connection with the practice of the Latin language observed by the legal experts of the time.

Notre article illustre l'utilité de la calibration générique des corpus pour l'analyse diachronique de la langue française. Nous explorons le cas de l'antéposition stylistique de l'infinitif et du participe au sein de trois corpus constitués respectivement de textes littéraires, de coutumiers normands et de procès, de la période de l'ancien français à celle du français classique. Le

calibrage générique montre que les évolutions de ces phénomènes d'antéposition diffèrent selon le type de texte, et que les textes légaux montrent une surreprésentation de ces constructions, en lien avec la pratique de la langue latine qu'observaient les spécialistes du droit.

INDEX

Keywords: corpus linguistics, diachronic linguistics, stylistic fronting, word order, infinitive, participle, text type

Mots-clés: linguistique de corpus, linguistique diachronique, antéposition stylistique, ordre des mots, infinitif, participe, type de texte

AUTHORS

PIERRE LARRIVÉE

CRISCO (UR 4255), Université de Caen Normandie

MATHIEU GOUX

CRISCO (UR 4255), Université de Caen Normandie