

Faculté de philosophie, arts et lettres

Étude sémantico-historique du terme *seruus* entre le 9^e et le 13^e siècle

Essai de diplomatique numérique

Auteur : LEROY Noé

Promoteur(s) : BERTRAND Paul et PERREAUX Nicolas

Lecteur(s) : DE VALERIOLA Sébastien

Année académique 2021-2022

Intitulé du master et de la finalité : Master en Communication de l'Histoire

Table des matières

Introduction	1
1 État du sens et de l'emploi de <i>seruus</i> dans l'historiographie	8
1.1 Définition de <i>seruus</i> : exemple de l'étymologie et de la lexicographie . . .	8
1.1.1 Étude étymologique	8
1.1.2 Lexicographies médiévales	9
1.1.3 Lexicographies modernes	12
1.1.4 Dictionnaires latins contemporains	15
1.2 Historiographie de l'emploi de <i>seruus</i> : les courants, les penseurs.	17
1.2.1 L'école historique de la deuxième moitié du 19e siècle et du début du 20e siècle	17
1.2.2 La révolution Marc Bloch et l'essai de synthèse de Charles Verlinden	19
1.2.3 L'entreprise néo-marxiste	21
1.2.4 Les écoles allemande et italienne	23
1.2.5 L'historiographie récente	25
1.3 Approche complémentaire : la sémantique historique	26
1.3.1 La sémantique historique : une piste pertinente?	27
1.3.2 La sémantique diachronique	27
1.3.3 Définition de la méthode : qu'est ce que la lexicométrie?	30
2 Analyse synonymique de <i>seruus</i> au sein de son champ lexical	33
2.1 Le choix de <i>seruus</i>	33
2.2 Comparaison des cas	36
2.2.1 Démarche méthodologique	36
2.2.2 Interprétation des résultats	38
2.3 Les schémas de pesage traditionnels pour les termes : <i>Term Frequency and Inverse Term Frequency</i>	40
2.3.1 Méthodologie	42
2.3.2 Présentation des données et interprétation	45
2.4 Les indices de similarités	49
2.4.1 Indice de Dice	51
2.4.2 Indice de Jaccard	55
2.4.3 Similarité Cosinus	59
2.4.4 Corrélation de Pearson	60
2.4.5 Fonction de similarité personnelle	61
2.4.6 Conclusion	63

3	Définition du terme seruus	66
3.1	QLF	68
3.1.1	Étude générale dans les <i>CEMA</i>	69
3.1.2	Comparaison régionale	76
3.2	VBE	78
3.3	SUB	82
3.4	Le rôle central de la Bible dans le langage médiéval	88
3.4.1	Ancien Testament	91
3.4.2	Nouveau Testament	92
4	Au coeur de l'acte diplomatique : les types de documents et la formule d'humilité	93
4.1	Les différences sémantiques au travers de la typologie : chartes, bulles, lettres et notices.	93
4.2	Histoire et signification de <i>Seruus seruorum dei</i>	94
4.2.1	Histoire de <i>seruus seruorum Dei</i>	96
4.2.2	L'importance du titre d'évêque de Rome	98
	Conclusion	104
4.3	L'inutilité étymologique et les pistes de la lexicographie	104
4.4	Quels résultats pour les outils numériques	105
4.4.1	La différenciation sémantique : une piste caduque?	106
4.4.2	La définition par cooccurrence	107
4.4.3	Une dichotomie sémantique dans la Bible	108
4.5	Ce que révèle la formule d'humilité	108
4.6	Les résultats face à l'historiographie	109
4.7	Les pistes d'amélioration et de prolongement	110
	Table des figures	113
	Bibliographie	115

Remerciements

*”Une difficulté n’en est plus une,
à partir du moment où vous en souriez, où vous l’affrontez.”*

Baden Powell.

Un mémoire représente son lot d’émotions, de rebondissements, d’imprévus, de stress et de découragement, mais il est surtout une source de défi, de dépassement, de maîtrise de soi, et au bout du compte, d’une fierté immense. Les heures passées à écrire, les dizaines à compiler les informations et les lectures, les centaines à coder font ressentir, *in fine*, une satisfaction pleine, une sensation de mérite incomparable qu’il serait aisé de s’attribuer. Mais même si le gros de l’oeuvre peut m’être imputé, tant d’autres ont contribué à le façonner, le perfectionner. Qu’il s’agisse de conseils, d’idées, de pistes, ou encore du soutien et des encouragements, toutes ces personnes méritent elles aussi d’être reconnues et de trouver place parmi cet ouvrage qui est, à ce jour, la plus grande réalisation de ma vie.

Dans ses *Pensées pour moi-même*, Marc-Aurèle débute ses premières pages en citant les choses pour lesquelles il se sent redevable, les qualités qu’il a développées au contact des gens qui lui sont importants. J’aimerais rendre un même hommage à tous ceux qui m’ont épaulé au cours de ces deux années de labeur.

Tout d’abord, je me dois de remercier les deux principaux acteurs de ce mémoire, ces maîtres à penser qui m’ont guidé dans mes choix, rassuré dans mes doutes et conseillé dans mes hésitations, et sans qui ce projet n’aurait certainement jamais vu le jour : mes promoteurs, Monsieur Paul Bertrand et Monsieur Nicolas Perreaux. Au premier, je dois l’apaisement ressenti tout au long du processus et l’assurance solide de l’utilité de mon travail. Ce fut une chance de pouvoir compter sur lui et sur ses nombreux conseils. Au second, je dois la passion ardente pour mon sujet de recherche et un soutien moral constant. Les circonstances de fin d’année ont été particulières, mais je n’ai jamais douté de la sincérité de son engagement à me guider. Je n’aurais pu souhaiter de meilleur soutien académique.

Je tiens ensuite à remercier la personne qui est la plus chère à mon coeur : ma maman. Mon parcours académique ayant été ce qu’il a été, je ne pourrai jamais exprimer avec des mots la gratitude que je ressens envers celle qui m’a soutenu tout du long, que je pensais si souvent décevoir mais qui rayonnait de fierté à chaque accomplissement. Il m’aura fallu huit longues années pour arriver au bout, et plus les années passaient, plus je bataillais pour la rendre fière. J’espère qu’il en sera ainsi longtemps, merci maman.

Il m’importe aussi de remercier mon grand-père qui nous a malheureusement quittés en avril 2020, au début de la pandémie. À toi, je dois tellement. Ton départ m’a énormément

coûté, j'en ai beaucoup souffert. Lors de notre dernière rencontre, lorsque tu étais encore lucide, tu m'as soufflé que tu étais fier de moi, et ce malgré les déceptions constantes que j'ai pu t'infliger. Tu voulais que je réussisse mes études, que je m'épanouisse dans ce que je fais. Quelle tristesse que tu ne puisses être témoin de ma réussite, de mon épanouissement. Sache en tout cas, que pendant ces deux années, tu fus ma plus grande source d'inspiration. Ce travail est mon hommage à la grande personne que tu as été pour moi.

Je suis aussi très reconnaissant du soutien inconditionnel de tous les autres membres de ma famille. Qu'il s'agisse de ma grand-mère et de ses encouragements constants, son rire et sa bonne humeur communicative, de mon père et de son infaillibilité à me conduire au train pour que je puisse me rendre à la bibliothèque, à mon grand-frère Timothée, lui que j'admire tant et qui me rend mon admiration, à Vivien, Tristan et Leah, que je sais être heureux pour moi, même si tous les jours n'étaient pas faciles, et que nous formons une tribu bruyante et atypique, je vous remercie sincèrement d'avoir été là tout au long de mon parcours.

Comment ne pas penser aussi à celui qui frôla l'oubli de soi pendant les derniers jours d'intense labeur afin de m'aider à finaliser la mise en page du mémoire. Je ne puis le remercier assez et ne pourrais rêver d'un meilleur ami. Merci donc à Damien Marcotty, à ces heures passées à me rassurer dans mes instants de doute, à tous ces souvenirs partagés, à la chaleur du foyer qu'il m'a offert pendant mes phases de rédaction. Merci aussi à Marie-Catherine, sa maman, qui m'a mis dans les meilleures conditions pour faciliter ma concentration, à sa soeur, Séverine, ma partenaire de travail et le public contraint de mon humour douteux.

À tous mes camarades historiens (et en particulier à Giuseppe Camarda, Émile Overtus et François Vagman) qui ont partagé la même galère, je dis également merci. Merci aussi à tous mes amis rencontrés à l'École Nationale des Chartes pour avoir cru en moi dans ses moments où cela m'était difficile. Pour les citer brièvement, merci à Malamatenia Vlachou, Zoé Suart, Cédric Pilate, Simon de Schaetzen, Victor Francken, Madeleine Schmit, Abdellatif Lahlou, Chiara Loriaux et tous les autres.

Je n'oublie pas non plus mes relecteurs et amis : Thomas Cleerebaut et Hadrian Nevzadi. Merci d'avoir pris le temps de lire, corriger et commenter. Et je sais qu'il l'aurait fait s'il avait eu le temps, merci à Julien De Vos qui m'a donné confiance en mes capacités. Parmi les historiens, il ne me reste qu'à remercier mon ancien camarade de promotion et proche ami Stratos Sergakis. Enfin, merci aux premiers à m'avoir offert leur aide, Manoëlle De Viron, Antoine Sadzot et Franck Fadeur, sans qui beaucoup de mes problèmes de code n'auraient jamais trouvé de solution.

J'en oublie sûrement quelques-uns, mais il m'est impossible d'être exhaustif ici. Dans tous les cas, à vous tous je dois d'être arrivé si loin. Grâce à vous tous, j'ai pu affronter tous ces obstacles et conserver le sourire tout du long, malgré tout.

Résumé

Souvent, dans l'historiographie, le terme *seruus* a été employé pour appuyer des thèses qui accentuaient la servitude des plus humbles au Moyen Âge. On parle de servage. Mais cette perception est biaisée par la volonté, depuis la Renaissance, de voir en cette période un repoussoir, une période obscure qu'il faut absolument dépasser. L'appellation *seruus* n'était cependant pas l'unique apanage des asservis, mais bien aussi le titre des serviteurs de Dieu (tant les fidèles que les hommes d'Église). Difficile de croire que dans les documents diplomatiques médiévaux, un scribe ait pu songer à des sens si distincts. Les techniques numériques et l'anthropologie historique donnent des clefs d'analyse pour deviner le ciment de cette polarité sémantique. Les analyses cooccurentielles de *seruus* dans les documents diplomatiques appuient cette thèse de pluralité sémantique, à laquelle la théorie de pensée analogique d'Anita Guerreau-Jalabert donne sens. La traduction n'est pas ici à l'ordre du jour, car les définitions des lexicographies médiévales et modernes sont peu représentatives d'une réalité mouvante et fluctuante. Mais il n'est pas faux de dire que la conception de la servitude pour les *serui* passe par la projection de la relation *seruus/deus* à leur propre *dominus*.

Introduction

"L'historien peut-il ne pas se préoccuper d'herméneutique?"¹. Peut-on envisager d'étudier les textes sans se préoccuper du sens des mots qui le composent? Les lexicographes médiévaux, modernes et contemporains se sont attelés à les traduire, mais l'entreprise n'est que l'expression d'une réalité *hinc et nunc*. L'opposition entre *verba* et *realia* est obsolète, il est temps de reconnaître l'intraduisibilité des mots latins et de se pencher sur leurs occurrences dans le texte et leur assemblage (notion phare pour Alain Guerreau) articulé².

L'ambition de ce mémoire est de répondre à cet appel et de réviser la conception biaisée/erronnée de la notion de *seruus* ainsi que les tentatives de définitions limitatives qui ont jalonné l'histoire médiévale, moderne et contemporaine. Il ne s'agit pas là de se risquer à le définir à nouveau, mais plutôt d'exprimer les limites des méthodes traditionnelles en ce qui concerne ce lemme spécifique, et de tenter de les compléter au moyen de techniques numériques novatrices.

Pour commencer donc, il sera question de reprendre ces techniques jugées insatisfaisantes (étymologie et lexicographie), pour évaluer le niveau des connaissances et des préjugés des historiens. Il ne s'agit pas là d'une démonstration d'exhaustivité. L'ambition n'est pas de tomber dans le même travers que les lexicographes eux-mêmes, mais plutôt

¹ GUERREAU, A., *L'avenir d'un passé incertain*, p.104.

² *Idem*.

de se construire une idée de leur représentation du mot, de manière diachronique. Cette étape est importante pour la phase suivante : le résumé de l'historiographie au sujet de la servitude au Moyen Âge. Avec la synthèse des définitions de *seruus*, il sera plus aisé d'aborder les concepts qui reposent sur ces sens approximatifs. Enfin, le premier chapitre se clôturera avec une brève présentation des techniques plus récentes, à l'instar de la sémantique historique et de la lexicométrie.

Ce n'est que par la suite que le corpus de base sera décortiqué. Initialement, le projet portait sur une comparaison régionale entre le territoire de la Belgique actuelle (au travers de la base de données des *Diplomata Belgica*) et de la Bourgogne médiévale (les *Chartae Burgundiae Medii Aevii*), mais *in fine*, grâce à la publication des *Chartae Europaeae Medii Aevi*, qui inclut les *CBMA*, le corpus a été reconsidéré, élargi même. Les techniques de traitement de données appliquées aux *CEMA* sont de l'ordre de la distinction de *seruus* avec les autres lemmes de son champ sémantique (par des comparaisons de cas, des similarités de lexiques de cooccurrences ou des comparaisons statistiques plus basiques) et de l'approche sémantique par cooccurrences (au sein des *CEMA* et en comparaison avec les cooccurents présents dans la Vulgate latine).

Enfin, le lemme *seruus* sera passé au crible d'une démarche double qui mixe l'analyse diplomatique des formules et la projection de concepts anthropologiques (notion de matrice binaire d'analogie dans les conceptions médiévales). Cela permettra ainsi de l'analyser, nous l'espérons, de la manière la plus transversale et exhaustive possible. Une fois que cela sera fait, il conviendra de reprendre certains concepts proposés dans l'historiographie et de les mettre en parallèle avec les résultats des recherches.

Détermination du sujet et brève évolution du projet

Pourquoi avoir choisi le terme *seruus*? L'intention initiale pour le mémoire était de traiter de la question de la servitude médiévale au travers d'un corpus vaste. Le champ sémantique de la servitude étant riche, il fallait opérer un choix. De tous les termes identifiés dans l'ouvrage d'Alice Rio *Slavery after Rome*³, *seruus* apparaissant comme le plus englobant et surtout le plus manifesté.

Dans un second temps, il fallait circonscrire la temporalité la plus pertinente. Lors du séminaire d'accompagnement de première année de Master, c'est au travers des *CBMA* que les balises ont été initialement posées. De prime abord, les siècles les plus fournis en occurrences sont les 10^e-14^e siècles. Et bien que le 9^e siècle ne se démarque pas similairement, contextualiser l'essor de l'écrit à cette époque, lors de la Renaissance carolingienne,

3 RIO, A., *Slavery after Rome (500-1100)*, 2017.

présentait un intérêt certain. Il permet en effet de fournir une continuité logique avec la sémantique antérieure liée au thème de l’esclavage. Une étude diachronique s’étalant sur 6 siècles pouvant sembler démesurée, la décision de ne pas traiter du 14^e siècle fut considérée et adoptée.

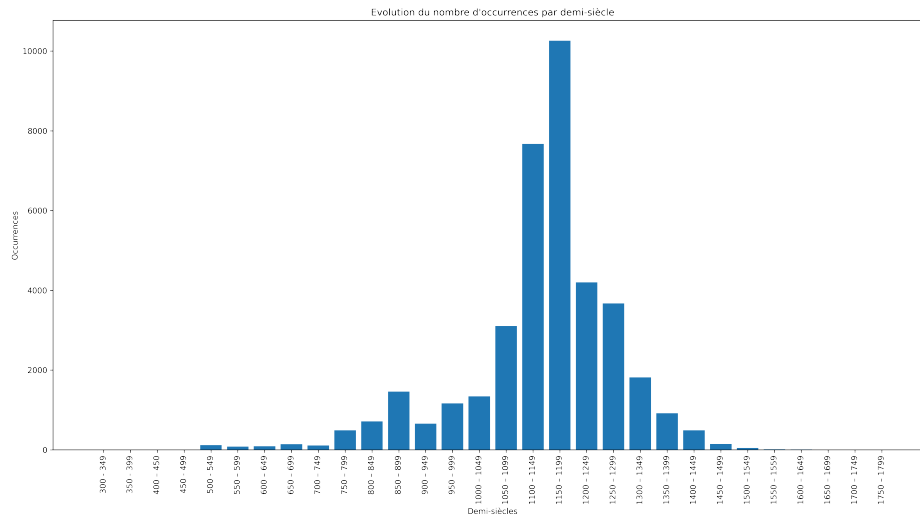


Figure 1 – *Évolution des occurrences de seruus par demi-siècle dans les CEMA.*

Le changement de corpus ne fut pas synonyme de changement de cap, puisqu’après une analyse similaire dans les *CEMA*, la distribution des occurrences de *seruus* rejoignait notre première intuition (Figure.1).

Le passage de la première base de données à la seconde a signifié plusieurs choses. D’abord, le corpus initial, composé d’une trentaine de milliers d’actes diplomatiques numérisés et lemmatisés, se voyait démultiplié par un facteur 10. Il est désormais question d’un corpus de plus de 75 millions de mots. Plus d’actes pour plus d’occurrences, plus d’occurrences pour une meilleure représentativité. Néanmoins, l’avantage des *CBMA*, c’est qu’elles ne concernaient qu’une zone géographique unifiée et étaient moins enclines aux variations sémantiques régionales. Mais surtout, le passage d’un outil d’exploitation comme TXM⁴ à l’API des *CEMA* a nécessité une adaptation rapide, puisque la décision de changement est survenue au cours du mois de février. Tout le travail et les compétences accumulés jusque-là n’étaient donc plus d’actualité et ont du être laissés de côté (*hierarchical clustering*, formation à la textométrie,...) et renouvelés.

Afin de pouvoir exploiter au mieux cette base de données brute, il a donc été nécessaire de se former à la programmation, plus particulièrement au langage Python. Grâce à

4 Heiden Serge, Maguë Jean-Philippe, & Pincemin Bénédicte. (2010). TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement. In JADT 2010 : 10th International Conference on the Statistical Analysis of Textual Data (pp. 12 p.). Rome, Italie.

un Erasmus à l'École Nationale des Chartes, les bases ont été engrangées mais la formation à la librairie *Pandas* fut très majoritairement développée en autodidacte. Sans elle, aucune donnée n'aurait pu être extraite⁵ et ce mémoire aurait été complètement vain.

Il a également été décidé de réaliser un *Github* dans lequel seront rassemblés tous les codes produits à l'occasion de nos recherches, ainsi que tous les schémas et graphiques générés qui n'ont pas trouvé place dans la version papier. L'ambition était également de proposer librement l'accès aux données par souci d'objectivité afin d'assurer une possible reproductibilité des calculs, mais aussi, dans une même mesure, de promouvoir un modèle de conservation.

Car il ne s'agit pas ici d'un mémoire commun, s'inscrivant dans une tradition historiographique classique. Il n'y a aucun ouvrage duquel s'inspirer, aucun auteur à avoir abordé le problème de cette manière. Une éventuelle filiation méthodologique peut se retrouver dans quelques travaux, dont aucun ne traite de la question de *seruus*. C'est grâce à l'appropriation partielle des travaux de Nicolas Perreaux, Alain Guerreau, Anita Guerreau-Jalabert ou encore Eliana Magnani qu'il a été possible de penser une méthode⁶.

En tout, ce n'est pas moins de 900 fonds d'archives que l'on retrouve dans les CEMA, et parmi eux, près de 570 ne comportent qu'une vingtaine de chartes. Ces multiples provenances provoquent un déséquilibre au sein du corpus, surtout d'un point de vue des métadonnées. En effet, ces dernières ont été générées d'après des critères et des intentions différentes. Elles ne se superposent pas. Mais le projet a pour démarche d'uniformiser cela en simulant grâce à l'intelligence artificielle les métadonnées manquantes. La page de présentation du projet fait écho à cette intention et exprime même leurs *desiderata* : « Une mise à disposition de tous les actes lemmatisés sous différents formats et permettant le traitement par différents programmes (R, TXM, CQP-Web), tout d'abord, mais aussi de « diffuser une bibliographie des éditions diplomatiques européennes ».

Une dernière précision peut-être. Les chartes sont les sources médiévales les plus nombreuses qui nous soient parvenues. Il est donc très logique de s'y pencher pour étudier la dynamique sociale à travers l'écrit. Malgré tout, la formalisation des actes, bien que révélatrice par moment, conscrit la sémantique dans une interprétation limitée. Pour bien faire, il aurait sûrement fallu intégrer d'autres types de documents et procéder à une étude transversale (documents hagiographiques, *Patrologia latina*,...), ce qui n'était pas envisageable pour des raisons pratiques évidentes.

5 Nous tenons à remercier Chahan Vidal-Gorène d'avoir permis de participer à son cours de Python, même après la fin de l'Erasmus

6 Il n'y a jamais eu un seul ouvrage, pour chacun, plusieurs de leurs travaux ont à chaque fois été consultés (Cf: la bibliographie)

Présentation de la base de données : les *CEMA*

Le projet est né au cours des années 2008-2009 à l'initiative de l'historien Nicolas Perreaux dans le cadre d'une thèse qu'il consacrait à l'étude statistique sur le champ sémantique de l'eau. Il avait à cette fin commencé à collecter un corpus abondant de plusieurs milliers de chartes numérisées au travers de bases de données préexistantes (*Chartae Galliae*, *CBMA*, *Scripta*, *Cartulaire d'Île-de-France*, *Deeds Project*, *dMGH*, *Codice Diplomatico della Lombardia Medievale*) pour lesquelles les métadonnées étaient inégales et hétérogènes. Après quelques étapes de formatage et de lemmatisation, elles étaient enfin prêtes à l'emploi.

L'API de référence pour cette base de données est *NoSketchEngine*, qui nécessite une connaissance basique du langage en *CQL*. Les services que peut rendre cette base de données sont nombreux et sont pour les médiévistes une source d'une richesse inégalable, surtout en diplomatique. Qu'il s'agisse d'études en sémantique ou des dynamiques de l'écrit à échelle européenne ou régionale, les champs des possibles sont très vastes.⁷

Cette entreprise colossale recouvre un très vaste territoire géographique et est subdivisée en conséquence. Il est possible de décortiquer les informations par des prismes multiples touchant tant aux régions qu'aux diocèses, aux pays qu'aux institutions locales, à la typologie qu'aux demi-siècles. Dans le cas de notre mémoire, l'étude séculaire s'impose, et étant donné que l'objectif final relève d'une évolution large, les demi-siècles seront même préférés.

Ce mémoire se repose donc sur une base de données de documents diplomatiques dont le traitement passe par des outils numériques et statistiques. Ce genre d'approche est encore marginal et souffre de manque de visibilité. Néanmoins, il est certain que la discipline historique gagnerait beaucoup à exploiter un corpus de cet acabit.

Une brève histoire de la diplomatique et de son évolution paradigmatique

La diplomatique classique telle qu'elle a été pensée à l'époque de Mabillon connaît depuis quelques décennies des changements paradigmatiques majeurs avec l'essor des techniques numériques. Des premières traces de contestation des documents diplomatiques attestées dès le 15^e siècle (Nicolas de Cues et Lorenzo Valla démontrent la fausseté de la Donation de Constantin⁸) à aujourd'hui, la discipline a connu une évolution constante dont la numérisation est désormais le point d'orgue. Depuis la création officielle de la diplo-

⁷ PERREAUX, N., *Possibilities, challenges and limits of a European charters corpus (CEMA)*.

⁸ GUYOTJEANNIN, O., *Diplomatique médiévale*, p.17

matique dont le but, la méthode et l'ambition scientifique ont été forgées dans la seconde moitié du 17^e siècle, à aujourd'hui, il a été question d'autres extensions, d'autres imports thématiques. Dans un article de 1996, Olivier Guyotjeannin fait état de l'élargissement du champ d'étude de la discipline depuis ses débuts jusqu'à l'aube du 21^e siècle⁹ qu'il jalonne comme suit. Dans la seconde moitié du 18^e siècle, la « diplomatie pratique » s'est attachée à vulgariser le corpus de connaissances auprès des médiévistes. Au 19^e siècle, elle se spécialise dans le raffinement des méthodes et le champ diplomatique se structure en parties distinctes, en branches spécialisées. Depuis les années 50, la diplomatie est une nouvelle fois partie explorer et conquérir d'autres champs comme par exemple celui du document diplomatique à proprement parler. Le « document » est « monument »¹⁰.

Ce dont les dernières années témoignent à leur tour, c'est de l'émergence de bases de données considérables et multiples telles que les Chartes originales de l'Artem¹¹, les *Diplomata Belgica*¹², le *Deeds Project*¹³ ou encore les *Chartae Burgundiae Medii Aevi*¹⁴. Chacune de ces bases de données accumule des centaines voire des milliers d'actes détaillés par des métadonnées diversement approfondies. Les CEMA sont à leur suite le résultat d'un long processus d'agglomération visant à centraliser toutes ces données. De tels outils permettent de consulter beaucoup de mots et d'améliorer les informations qu'ils trouvent dans les éditions (datation, authenticité, établissement du texte), même s'ils ne sont pas exempts de reproches pour autant, omettant en grande partie les commentaires diplomatiques qui accompagnent originellement l'édition d'un texte, ajoutant parfois même des fautes aux éditions qu'elles utilisent¹⁵. Pour citer l'historien Benoît-Michel Tock, « les nouvelles possibilités offertes par les techniques modernes ont donné à la diplomatie un aspect que l'on pourrait qualifier de magique. L'on accède facilement à un nombre de documents encore impensable il y a quelques années de cela, l'on dispose d'outils renouvelés pour leur étude... Pour autant, l'on peut encore améliorer la chose, en affinant les instruments ainsi mis à notre disposition, et en profitant de ces nouvelles possibilités pour ne pas hésiter à donner des éditions et des outils différenciés, avec des niveaux d'élaboration correspondant aux situations et aux besoins »¹⁶.

9 GUYOTJEANNIN, O., *La diplomatie médiévale et l'élargissement de son champ*, dans *De la diplomatie médiévale à la diplomatie contemporaine (actes du colloque organisé par l'École nationale des chartes et la Bentley historical Library de l'université de Ann-Arbor [Michigan, États-Unis]*, Paris, 1993.

10 *Ibid*, p.14.

11 <http://telma.irht.cnrs.fr/outils/originaux/colophon/>

12 https://www.diplomata-belgica.be/colophon_fr.html

13 http://deeds_library_utoronto_ca/

14 <http://www.cbma-project.eu/>

15 MICHEL-TOCK, B., *L'étude du vocabulaire et la datation des actes : l'apport des bases de données informatisées*, dans GERVERS, M., *Dating Undated Medieval Charters*, p.82.

16 Notice introductive de Benoit Michel-Tock dans AMBRSIO, A., BARRET, S. et VOGELER, G., *Digital diplomatics : the computer as a tool for the diplomatist ?*, p.303.

Une telle profusion de bases de données est néanmoins contrecarrée par un engouement peu partagé par la communauté des médiévistes, peu encline à quitter les sentiers des méthodes traditionnelles, comme le regrette Benoît-Michel Tock¹⁷. Le décalage entre la création de corpus de grande envergure et l'exploitation de ces derniers est donc avérée et déploré¹⁸.

Mais l'apport de l'informatique ne s'arrête pas uniquement à la concaténation documentaire, elle s'illustre aussi partiellement dans l'exploitation des données. Des outils de transcription numérique ont commencé à voir le jour depuis une dizaine d'années, à commencer par des programmes en HTR (*Handwritten text recognition*). Ceux-ci emploient les méthodes liées aux intelligences artificielles et rendent possible une production relativement précise de textes issus d'images de manuscrits ou de documents diplomatiques¹⁹. L'intérêt du numérique est donc présent de l'extraction textuelle au rassemblement en base de données, mais sort aussi de ces carcans, allant jusqu'à permettre un traitement de ces données. Des programmes comme le projet textométrie (et plus récemment les ») offrent des avantages et d'excellentes pistes dans l'analyse de données, et, comme le rappelle Alain Guerreau, « même si les développements récents [...] visent la facilité d'installation et de manipulation [des bases de données], c'est une lourde erreur de croire que l'on puisse vraiment obtenir le meilleur résultat de tous ces outils sans jamais se salir les mains ! [...] les codes ou langages informatiques les plus courants (regexp, html, perl ou python, R) s'apprennent [...] très rapidement, mais ne deviennent vraiment utiles que dans la mesure où l'on en acquiert l'expérience »²⁰. Autrement dit, au-delà même de tous ces programmes qui existent et qui facilitent la recherche et l'exploitation de ces larges corpus, pour en tirer le meilleur, il est nécessaire de se retrousser les manches et de s'initier à certains langages de programmation.

17 « Mais curieusement, celles [les bases de données de textes diplomatiques] qui existent ne sont pas encore très utilisées. Disons-le franchement : elles sont sous-utilisées, et n'ont pas, ou pas encore révolutionné la diplomatique comme elles auraient dû le faire » dans B.-M., Tock, *L'apport des bases de données*, p.283.

18 PERREAUX, N., *De l'accumulation à l'exploitation ? Expériences et propositions pour l'indexation et l'utilisation des bases de données diplomatiques*, dans AMBRSIO, A., BARRET, S. et VOGELER, G., *Digital diplomacy : the computer as a tool for the diplomatist ?*, p.191.

19 Dans le domaine, des programmes comme Transkribus, Tesseract ou plus récemment, E-Scriptorium sont incontournables.

20 GUERREAU, A., *Textes anciens en série*, dans *Bulletin du centre d'études médiévales d'Auxerre*, p.51.

1 État du sens et de l'emploi de *seruus* dans l'historiographie

Depuis longtemps déjà, on relève un certain malaise parmi les historiens lorsque se pose la question de la traduction du terme *seruus*. Selon qu'ils faisaient de l'histoire économique ou de l'histoire sociale, ils pouvaient aboutir à des conclusions qui ne se superposaient pas, voire même se contredisaient²¹. En outre, comme le souligne Vincent Corriol, « le qualificatif de *servus*, pêché dans le vocabulaire juridique romain, est trop mal adapté pour qu'on puisse couramment l'employer »²². La réalité à laquelle renvoie le terme pose donc des soucis d'interprétation et demeure source de débats. Si les analyses historiques ne permettent pas d'aboutir à une définition consensuelle, même partiellement, alors le problème a besoin d'être pris à la racine, autrement dit, à son étymologie.

1.1 Définition de *seruus* : exemple de l'étymologie et de la lexicographie

L'étymologie est une piste qui, en théorie, est modérément admissible mais qui correspond à une entrée en matière logique. Bien que son sens radical puisse être absent de son sens effectif, il donne un point de départ, une balise temporelle, une limite à l'intervalle d'analyse. Elle composera donc la base de cette étude diachronique de la sémantique de *seruus*. C'est plutôt au travers de la comparaison des lexiques et dictionnaires médiévaux que son sens initial verra le jour. Malgré un échantillon lacunaire assumé, cette démarche a le mérite d'offrir une perspective évolutive dans la conception des médiévaux eux-mêmes. En outre, les dictionnaires modernes donneront une perspective neuve à la définition, tout comme les dictionnaires contemporains. Les premiers pour le regard biaisé face à cette époque de laquelle ils cherchent absolument à se distancer, et les seconds par le regard érudit, plus critique, et par leur méthode.

1.1.1 Étude étymologique

Littéralement, *seruus* provient du grec et signifie gardien. Mais cette origine a peu à peu été oubliée. Si bien que finalement, seul le sens d'esclave est passé dans les dérivés tels que *servio* ou *servitus*²³. Telle était la perception que l'on en avait à la fin du 19e siècle en tout cas : une vision très parcellaire, faisant fi de la transition entre le sens étymologique et la réalité sociale sous-jacente. Moins d'un siècle plus tard, la tendance

21 MAGNOU-NORTIER, É., *Servus - servitium : une enquête à poursuivre*, dans *Media in Francia. Recueil de mélanges offert à Karl Ferdinand Werner à l'occasion de son 65^e anniversaire par ses amis et collègues français*, p.269.

22 CORRIOL, V., *Les serfs de Saint-Claude*, p.21-47. (consulté le 12/02/2022 : <https://books.openedition.org/pur/126435?lang=fr#tocfrom1n2>)

23 BRÉAL, M. et BAILLY, A., *Dictionnaire étymologique latin 1918*, p.344.

s'inverse. Dans le Dictionnaire étymologique de la langue latine, la notice relative à *seruus* débute comme suit : « esclave ; adjectif s'opposant à *liber*, se dit des hommes, *seruus homo*, et des choses, en particulier des biens soumis à une servitude »²⁴. Autrement dit, prime alors l'attribution de la servilité sur l'origine sémantique à proprement parler. Mais l'auteur évoque néanmoins les racines linguistiques multiples qui tendent à expliquer le sens de *seruus* (avesta, latin, grec) et le rapprochent de ce sens de gardien. Toutefois, « le fait précis qui éclaircirait le passage de *seruus* du sens de « gardien » à celui « d'esclave », seul attesté en fait, [reste] inconnu »²⁵.

En ce qui concerne ses origines latines, le rapprochement avec le verbe *servare* semble intuitif. Certaines sources romaines plus tardives (*Institutiones Iustiani*) font ce rapprochement « because the slaves were believed to be originally prisoners of war whose lives had been spared »²⁶. Cela pourrait expliquer le rapprochement entre les deux sens. Sont esclaves ceux qui sont gardés en vie, ceux qui ont été épargnés. Cependant, cette hypothèse est contredite, ou au moins nuancée par le linguiste Émile Benveniste. Lui aussi se retrouve piégé dans la dichotomie entre l'étymologie *stricto sensu* et la condition sociale et légale de l'esclave. Autrement dit, il réfute le lignage avec *servare* sous le prétexte d'une contradiction originelle basée sur cette observation sociale, mais ne nie pas pour autant qu'une relation entre les deux termes existe²⁷.

Tout cela pour dire qu'au-delà de l'incertitude historique, l'étymologie elle aussi pêche par son incomplétude. Ce que les manuels montrent, c'est que la définition du terme fluctue et qu'aux origines le mot détenait un sens aujourd'hui perdu. Cette brève présentation étymologique permet de mettre en avant une méthode de définition sémantique qui s'inspire d'une situation sociale à une certaine époque pour donner du sens aux mots mêmes qui la qualifient. En d'autres mots, c'est un serpent qui se mord la queue. Nous pouvons comprendre d'où vient le mot, nous savons ce qu'il en est advenu, mais nous ne sommes pas capables d'expliquer le passage du premier sens au second.

1.1.2 Lexicographies médiévales

Un autre angle d'attaque dans cette lourde tâche de définition du mot *seruus* consiste à parcourir diachroniquement les travaux de lexicographie, l'intérêt étant de voir comment le terme est considéré à une époque donnée. Pour ce faire, il est nécessaire de structurer notre démarche. Avant tout, il aurait convenu de traiter de la présence du lemme auprès des grammairiens latins, pionniers dans la discipline, mais pour des raisons de temps et

24 ERNOUT, A. et MEILLET, A., *Dictionnaire étymologique de la langue latine*, p.660.

25 *Idem*.

26 PARVULESCU, A., *Lat. Servus*, dans *Indogermanische Forschungen*, p.190.

27 *Idem*.

surtout des raisons de chronologie, cela ne sera pas développé dans ce travail. L'essentiel des lexiques, glossaires et autres dictionnaires datent donc *a minima* de la période médiévale. Nous avons rassemblé un échantillon arbitraire de lexicographies latines et de dictionnaire français, rassemblés d'après les lectures des articles de Laura Bondi²⁸ et de Béatrice Turpin²⁹. Travailler sur des dictionnaires français anciens permet d'évaluer la conception postérieure du terme *seruus* et d'évaluer les connotations qui lui sont restées ou bien qui ont évolué. Enfin, dans un dernier temps, nous allons également rapidement passer en revue les différents lexiques de latin médiéval de l'époque moderne à nos jours, toujours dans cette perspective d'analyse diachronique de la sémantique.

Il est difficile de ne pas débiter avec l'une des figures latines du Haut Moyen Âge les plus centrales, à savoir Isidore de Séville. Homme de foi et surtout homme d'Église, l'auteur fut élevé au rang d'évêque de la ville de Séville en 601 à la suite de son frère, et resta à la tête de l'évêché jusque sa mort en 636³⁰. Tout au long de sa vie, il fut l'auteur de 17 œuvres dont la plus majeure, ses *Etymologies*, « was a compendium of classical and Christian knowledge arranged according to the etymologies of words and became a standard point of reference in the Middle Ages »³¹. Sa centralité et sa pertinence justifient sa place comme première œuvre analysée. Voici comment il définit le terme *seruus* :

« Quant aux esclaves (*serui*), ils tirèrent leur nom de ce que ceux qui, selon les lois de la guerre, pouvaient être mis à mort par les vainqueurs, quand on les préservait, devenaient des esclaves. De préserver (*seruare*) évidemment est venu le nom de serf (*seruus*). Le nom de servante (*ancilla*) vient de soutien. »³²

Cette définition du 7^e siècle fait écho avec l'explication précédemment mentionnée d'Adrian Parvulescu, à savoir qu'un esclave était une victime indirecte de la guerre, un survivant asservi. En outre, dans son œuvre, il fait également mention de tous les termes qui renvoient à la question de l'esclavage en expliquant les formes et les nuances³³. La condition servile simple telle qu'elle est évoquée ici renvoie donc à une conception primaire de la servitude. Là où ce texte est intéressant, c'est dans sa seconde partie. Car dès lors qu'il peut être mis à mort par les vainqueurs mais qu'il est finalement épargné, le *seruus* est « préservé ».

28 BIONDI, L., *Sur quelques mots et leur parcours dans la lexicographie latine du Moyen Âge et de l'Humanisme*, dans *Histoire Épistémologie Langage*, t.39, fasc.2, 2017.

29 TURPIN, B., *L'esclavage vu par les dictionnaires français*, dans DUMONT, L. et FREMIN, M., *Esclavage, Esclavages*, UCP, p.31-58

30 FEAR, A. et WOOD, J., *Isidore of Seville and his reception in the Early Middle Ages : transmitting and transforming knowledge*, p.16.

31 *Idem*.

32 REYDELLET, M., *Isidore de Séville : Étymologies*, t.IX, p.178.

33 BONDUE, D., *De servus à sclavus : la fin de l'esclavage antique (371-918)*, p.27.

Un siècle plus tard, le *Liber Glossarum* se fait l'héritier d'Isidore. Diffusé initialement dans l'entourage de Charlemagne³⁴, le manuscrit s'illustre comme « un maillon fondamental de la lexicographie occidentale »³⁵ et « peut sans aucun doute être placé au commencement de la longue lignée dans laquelle s'insèrent les instruments ultérieurs de la lexicographie médiévale »³⁶. Un récent projet porté par le *European Research Council* en 2011 propose une première édition critique de ce manuscrit du 8e siècle. En outre, leur site permet d'interroger directement les notices lexicographiques, préalablement ocrées, dans lesquelles on peut retrouver la définition suivante en ce qui concerne le terme *seruus* :

« uocabulum in latina lingua inde creditur dictum quod hii qui iam iure belli possent occidi, a uictoribus conseruabantur et serui fiebant, a seruando igitur serui sunt appellati. »³⁷

Globalement, on ne distingue pas d'évolution dans le sens du mot. Bien que très brève, l'explication évoque le même type de servitude qu'Isidore, à savoir la perte de liberté pour raison de guerre. Ceux qui sont défaits en viennent à servir les vainqueurs. Selon le *Glossarum*, le statut de *seruus* est relatif à la notion de changement de statut (*serui fiebant*). Les définitions d'autres manuscrits plus tardifs vont également dans ce sens, qu'il s'agisse de l'*Elementarium doctrinae rudimentum* de Papias, le grand lexicographe du XIe siècle³⁸, des *Magnae Derivationes* de Osbern de Gloucester³⁹, grammairien et exégète du milieu du 12e siècle (ici l'auteur inclut tout de même la notion de mise à mort sous couvert de la loi, ce qui est une nouveauté) ou encore du *Catholicon* de Jean de Gênes⁴⁰, ouvrage central qui fut réédité maintes fois jusque la Renaissance⁴¹. Ce dernier, néanmoins, nuance quelque peu le propos. Il reconnaît que les *serui* le devenaient par la guerre, lorsqu'ils

34 GRONDEUX, A. et CINATO, F. (éd), *Liber Glossarum digital*, Paris, 2016. (<http://liber-glossarum.huma-num.fr>)

35 GRONDEUX, A., *Le "Liber Glossarum" (VIIIe siècle) : prélogomènes à une nouvelle édition*, dans *Bulletin Du Cange*, n°69, 2011, p.23.

36 BIONDI, L., *Sur quelques mots et leur parcours dans la lexicographie latine du Moyen Âge et de l'Humanisme*, dans *Histoire Épistémologie Langage*, t.39, fasc.2, 2017, p.156

37 <http://liber-glossarum.huma-num.fr/exist/apps/libgloss/entry-SE612>

38 On retrouve dans son texte la notice suivante au sujet de *seruus* : *Seruus uero ex conitione seruitutis. | Item seruus in bello captus atque seruatur | dicitur famulus uero ex propria familia ortus | mancipium ex hostibus*. Cela provient du manuscrit latin conservé à la Bibliothèque nationale de France, département des manuscrits à la côte Latin 9341. (<https://gallica.bnf.fr/ark:/12148/btv1b10037103q/f250.item>).

39 CECCHINI, E., *Uguccione da Pisa : Derivationes*, 2004, p.1083. La notice reprend ces mots : *unde hic servus -vi, quia olim in bello capti, cum iure possent occidi, servabantur et fiebant servi, et ideo servus dicitur a servando propter antiquam consuetudinem, licet et quidam dicantur modo servi qui sic servati non fuerunt ; servus ergo dicitur ex condicione servitutis. | Item servus in bello captus et servatus, famulus ex propria familia servorum ortus, mauod autem dicitur servus quasi serens vana vel quasi servans vultum domini sui vel quasi serens vesrutias ethimologia est, non compositio*.

40 Ms-978 à la Bibliothèque nationale de France. Bibliothèque de l'Arsenal (<https://gallica.bnf.fr/ark:/12148/btv1b105154441.image#>)

41 JACQUIN, G., *L'oeuvre de Robert Estienne et les dictionnaires de la Renaissance*, dans *Revue 303 arts recherches créations. Patrimoine linguistique*, n°120, 2012, p.28

étaient pris en otage et il l'explique non pas par un sens qui lui serait contemporain, mais davantage par l'usage ancien lié à la coutume. De ce fait, il s'inscrit bien dans la tradition d'Isidore et des autres auteurs. L'évolution qu'il soulève concerne non pas un statut, mais plutôt une description. *Seruus* se dit de celui qui sème, qui engendre des problèmes inutiles, sans raison, ou qui sert sous le regard de son maître.⁴²

Entre le 7^e et le 13^e siècle, la définition *stricto sensu* telle que présentée par les différents auteurs susmentionnés trouve ses racines dans la guerre et n'évolue que très peu, par sursauts. Est initialement *seruus* le prisonnier de guerre, l'homme épargné par le vainqueur. Mais d'autres réalités sociales voient le jour dans les lexiques. De leur analyse nous pouvons tirer deux observations. D'abord, il est à noter que les définitions ne renvoient qu'à un statut et non pas à une condition. Pas d'indice donc sur le type de servilité qui les contraint. Mais c'est principalement la stagnation sémantique qui saute aux yeux et interpelle. Sur ces cinq siècles, le mot *seruus* ne donne pas l'impression de connaître de véritable évolution. En outre, les définitions renvoient beaucoup à l'étymologique bien que la réalité sociale soit bien différente. Qu'en est-il vraiment ?

Tout comme pour l'étymologie, l'étude sémantique par la lexicographie n'est pas la solution. Pire encore, elle s'illustre, selon Alain Guerreau comme une véritable aporie⁴³. Et tout comme pour l'étymologie, ce passage par une étude traditionnelle de la sémantique était nécessaire pour mesurer la distance à prendre avec grâce aux nouvelles technologies.

1.1.3 Lexicographies modernes

Au-delà des définitions médiévales, certains dictionnaires modernes reprennent aussi le mot latin pour expliquer et encadrer le sens de l'esclavage. C'est l'ordonnance de Villers-Cotterêts de 1539 qui consacre le français comme langue officielle de l'administration et permet ainsi l'émergence de dictionnaires plus spécifiques, normalisants, qui marquent le point de départ de notre analyse. Nous allons donc, à la lecture de l'article « L'esclavage vu par les dictionnaires français » de Béatrice Turpin, essayer de mettre en exergue la tradition du sens médiéval dans les dictionnaires modernes de langue française. À cette fin, nous allons tourner notre attention sur deux ouvrages qui marquent la transition, à savoir le « Dictionnaire français-latin » de Robert Estienne et « Le Trésor de la langue française » de Jean Nicot⁴⁴.

Dans le premier, nous trouvons, dans la notice relative au mot esclave, un rattachement à deux étymons latins, à savoir *servus* et *sclavus*. Voici comment Robert Estienne le

42 *Quod autem dicitur seruus quasi serens uana uel qui seruans uultum domini sui ,uel quasi serens uersutias ethimologia est.*

43 GUERREAU, A., *L'avenir d'un passé incertain*, p.104.

44 TURPIN, B., *L'esclavage vu par les dictionnaires français*, p.4

présente :

« ESCLAU ou ESCLAUE . Quali sclauus, servus emptitius. Une esclave prise du pays de Carie – Esse captam è Caria. Esclave prisonnier de guerre – Mancipium vel mancipium. SERF , SERUUS . Serf né en nostre maison de nostre serue & esclave – Verna, vernae Serf fugitif – Fugitiuus seruus. »⁴⁵

Tout de suite, nous pouvons constater que le sens de prisonnier de guerre est conservé et attribué non pas à *seruus* mais à *mancipium*, un terme sémantiquement très proche. L'esclave reste donc considéré comme un butin de guerre. Mais cela reflète-t-il une réalité d'actualité, ou est-ce juste là l'héritage d'une définition anachronique ? Il apparaît qu'« entre la conception antique ou médiévale de l'ennemi vaincu réduit à être exterminé, devenir esclave ou servir de monnaie d'échange, et celle issue des conférences de paix de La Haye de 1899 et 1903, les guerres de la fin du XVI^e siècle marquent sans doute une avancée, du point de vue de l'éthique plus que du droit. »⁴⁶. Il est donc difficile de savoir si le sens originel est toujours pertinent pour cette époque. D'autant plus que le mot *seruus* est traduit en français par serf. Mais les deux mots sont-ils absolument compatibles ? En tout cas, nous pouvons noter que l'hérédité de statut semble attestée.

Par la suite, le « Trésor de la langue française » définit l'esclave comme « celui et celle qui sont serfs ou par captivité forcée, comme par la guerre [...] comme celui qui né de pere et mere esclave ». À cela, il ajoute une brève définition de *seruus*, à savoir qu'il est le « nom attribué aux forçaires des galeres faicts serfs de peine »⁴⁷. Dans l'optique moderne, « esclavage et servage sont ainsi présentés comme équivalents sémantiques : l'esclave, c'est le serf captif »⁴⁸. On reprend là des éléments originaux fondant le statut de l'esclave chez les Latins, la captivité forcée et la captivité née. À cette définition, l'auteur rajoute également l'acception de serf « serviteur ». Autrement dit, apparaît enfin une trace de potentielle polysémie.

À quelles conclusions tout cela nous amène-t-il ? L'étude de l'étymologie nous a révélé que le sens originel de gardien avait été rapidement perdu pour ne désigner *in fine* que les personnes de condition servile. Par la suite, tous les témoignages médiévaux s'alignent et renvoient à l'origine de leur servitude, à savoir la captivité de guerre. Aucune information ne renseigne directement dans les dictionnaires sur des éléments attribuables au servage avant le 16^e siècle et l'apparition des lexiques français. En interrogeant les bases de données Gallica et Google Books au moyen de leur API de recherche sémantique (Google Ngram viewer (Figure.2) et Gallicagram (Figure.3)), cette tendance s'illustre d'autant

⁴⁵ *Ibid*, p.5.

⁴⁶ LE GOFF, H., *Droit de la guerre et droits des prisonniers de guerre au XVI^e siècle : le cas de la Ligue en Bretagne (1587-1598)*, dans *Annales de Bretagne et des Pays de l'Ouest*, n°124/2, 2017, p.26

⁴⁷ TURPIN, B., *L'esclavage vu par les dictionnaires français*, p.6

⁴⁸ *Ibid*, p.7

plus. Aucune mention des termes serf, servage ou autres formes dérivées (*serue*) n'est décelable avant le 16^e siècle alors que l'essor de langue française a débuté au 12^e siècle⁴⁹. Dans une notice d'un dictionnaire franco-anglais datant de 1580, le mot de servage est traduit par "to carye one into captiuity", tandis que la servitude est désignée comme "la subjection de ceux qui sont les Serfs à quelque seigneur"⁵⁰. Le servage est donc, au 16^e siècle, considéré comme une forme de mise en captivité, l'assertion d'une relation de dépendance et d'appartenance.

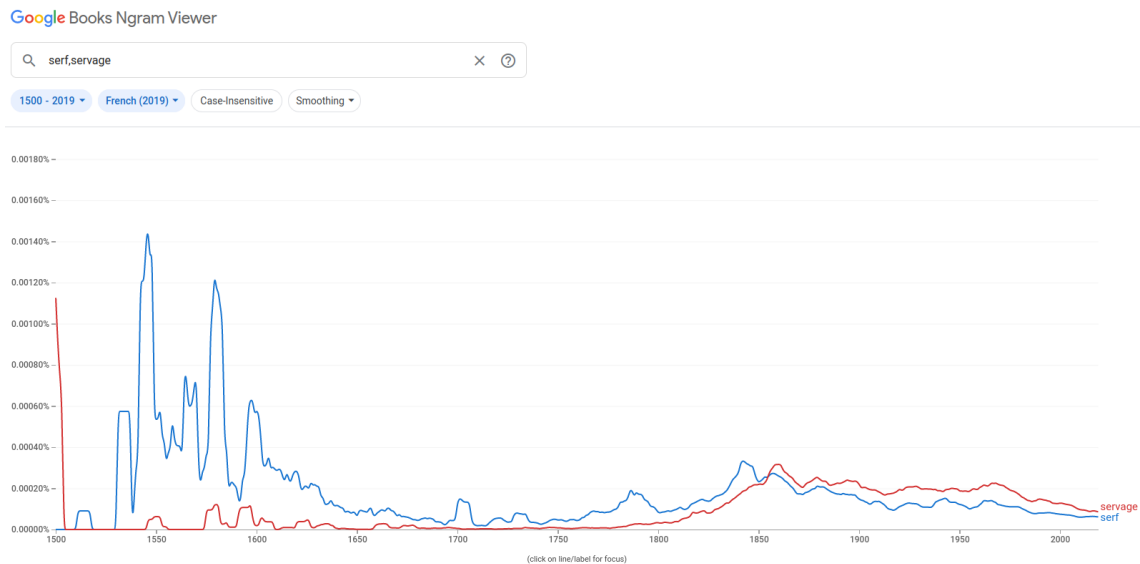


Figure 2 – Évolution de la proportion de serf et servage dans Google Ngram Viewer

49 BRUNNER, T., *Le passage aux langues vernaculaires dans les actes de la pratique en Occident*, dans *De Boeck Supérieur : Le Moyen Âge*, 2009, t.115, p.30.

50 DE SAINLIENS, C., *The Treasure of the French Tongue : Teaching the ways to Varie All Sortes of Verbes*

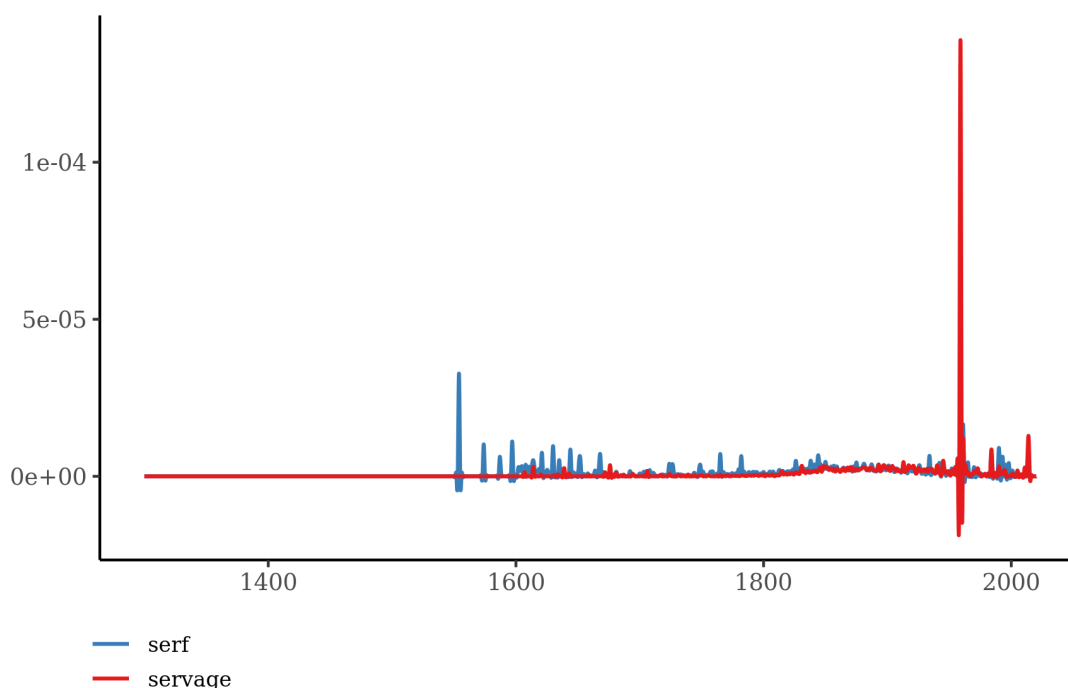


Figure 3 – Évolution de la proportion de serf et servage dans Gallicagram

Il convient toutefois de rester prudent sur ces constats car l'échantillon analysé ne donne pas vue sur l'évolution entre le 13^e et le 16^e siècle. Il est également à noter que le rapprochement direct entre le mot *seruus* et l'esclavage n'apparaît que lors de la période moderne. Nos dictionnaires français sont plus prompts à établir ce rapprochement sans nuance que leurs prédécesseurs médiévaux. En outre, ce n'est que tardivement que le mot *seruus* tend à proposer d'autres interprétations, même si celles-ci restent familière du champ sémantique de la servitude. Finalement, nous n'exposons ici que des tendances avec certaines des lexicographies les plus centrales, mais il est clair qu'il serait nécessaire de croiser ces informations à celles présentes dans d'autres sources pour affiner notre recherche. En ce qui concerne la France, certains grammairiens rédigèrent dès le 14^e siècle des lexiques bilingues inspirés du Catholicon, comme ce fut le cas pour l'Aalma ou pour le Dictionnaire de Firmin (1440) qu'il conviendrait de consulter et d'ajouter à notre échantillon. Ces derniers continuent en effet à donner la priorité au latin⁵¹ et minimisent en un sens les biais transposés par l'effort de traduction.

1.1.4 Dictionnaires latins contemporains

Pour corroborer ces aspects, un dernier angle d'attaque serait de comparer les dictionnaires franco-latins des décennies précédentes afin d'appréhender les perceptions contemporaines des historiens et des linguistes au sujet de cette problématique. Les dictionnaires les plus emblématiques de l'histoire récente étant probablement le *Du Cange*, créé

51 JACQUIN, G., *L'oeuvre de Robert Estienne et les dictionnaires de la Renaissance*, dans *Revue 303 arts recherches créations. Patrimoine linguistique*, n°120, 2012, p.29.

par l'Institut de France, le « Dictionnaire latin-français des auteurs chrétiens » de Albert Blaise⁵² ou le *Mediae Latinitatis Lexicon Minus*, auxquels on peut rajouter le Gaffiot, prolongateur du latin classique.

Le *Du Cange* est de loin le dictionnaire le plus exhaustif. La recherche de *servus* renvoie à une liste multipaginale de types d'utilisation, le tout référencé⁵³. On peut y trouver toutes les tournures possibles, les formules latines dans lesquelles il est associé à d'autres mots pour lui conférer un sens spécifique. Cette notice s'illustre par une rigueur de la nuance. On peut y consulter des définitions latines qui renvoient aussi bien aux écrits de Tacite qu'à la loi salique, certaines qui attestent d'un statut de prisonnier, d'autres encore qui les désignent comme étant attachés à une terre (*serui fiscales*), ou comme des fugitifs (*serui fugitivi*). Mais le plus important de tous, comme des serviteurs du Christ (*servi Christi/Dei*). C'est également le seul dictionnaire, à notre connaissance à soulever l'importance de la titulature pontificale du *servus servorum Dei*. Les différences sémantiques sont pour la première fois visibles, même si elles ne sont pas explicites. Malgré une compréhension partielle, il faut toutefois préciser que notre connaissance limitée de la langue latine n'a pas été suffisante pour tirer davantage d'informations.

De manière attendue, le Gaffiot traduit simplement *servus* par « esclave »⁵⁴, probablement à cause de son sens plus antique. Vient ensuite le Blaise Patristic, dictionnaire franco-latin qui étudie le sens des mots, en couvrant la période s'étendant de Tertullien à la fin de la période mérovingienne⁵⁵ et plus particulièrement le latin des auteurs chrétiens. L'auteur intègre ici un sens nouveau. Son approche a ceci d'intéressant que la traduction première est ambivalente et intègre la double notion d'esclave et de serviteur. Mais les deux sont liés. L'esclave ne renvoie pas à un statut ou à une condition servile, mais se dit péjorativement d'un « esclave du démon »⁵⁶. En outre, la version en ligne du Blaise Medieval disponible dans la Database of Latin Dictionaries traduit *servus* non plus par l'esclave, mais comme « le Serviteur (de Yahvé, en parlant du Messie) ». En recensant les vocables des auteurs chrétiens, Albert Blaise modifie partiellement la notion de *servus* en lui attribuant un particularisme lié au culte catholique. En cela, il rejoint la définition qu'en donne J.F. Niermeyer dans son *Mediae Latinitatis Lexicon Minus*, que ce dernier définit comme « valet, sans rapport à la condition de la personne ; [...] loc. Servus Dei. En parlant d'évêques et de prêtres ; [...] En parlant de moines »⁵⁷. Autrement dit, ces deux auteurs nuancent les définitions socio-juridiques habituelles et font part d'une autre

52 PARISSE, M., *Un lexique latin médiéval-français ?*, dans *Les Historiens et le latin médiéval*, p.315-319

53 <http://ducange.enc.sorbonne.fr/SERVUS#SERVUS-71>

54 <https://gaffiot.org/85072>

55 FAVEZ, C., *Review of Dictionnaire latin-français des auteurs chrétiens by A. Blaise*, dans *Latomus*, t.14/1, p.135-139.

56 BLAISE, A., *Dictionnaire latin-français des auteurs chrétiens*, 1954, p.757.

57 NIERMEYER, J.-F., *Medieval Latinatis Lexicon Minus*, p.967.

réalité sociale jusque-là absente des lexiques : celle du service. Les *serui* ne sont plus les esclaves, mais les serviteurs. La question est maintenant de comprendre comment les deux approches se juxtaposent et quelle vérité, nuancée, se cache derrière ce terme mystérieux.

1.2 Historiographie de l'emploi de *seruus* : les courants, les penseurs.

Comme nous l'avons soulevé, une grande partie du sens de *seruus* tel que conçu par les modernes et les contemporains fait écho au statut servile, à la question de la servitude. De nombreux historiens ont tenté de reconstruire ce concept, d'en comprendre la dynamique et les tendances depuis le 19^e siècle. Souvent dans l'historiographie, utilisaient-ils le terme pour fonder leurs hypothèses, comme c'est le cas pour le servage. Or, comme l'explique Didier Bondue, « le problème de l'évolution des conditions serviles depuis la fin du Bas-Empire jusqu'au X^e siècle a suscité, depuis la première moitié du XIX^e siècle, un foisonnement de positions contrastées, tant sur les plans de l'idéologie de la chronologie que du vocabulaire »⁵⁸. Les sens sont changeants et avec eux les réalités sociales. Les racines historiographiques de cette problématique sur le servage trouvent leur origine au 19^e siècle. Le temps passant, les premières théories furent contestées, retravaillées. D'autres historiens écrivirent à leur suite, et approchèrent la question avec un angle nouveau, avant d'être eux-mêmes critiqués, complétés.

Nous allons, dans cette partie, essayer de retracer le parcours théorique de cette notion de la servitude au travers des différentes écoles de pensée ayant jalonné l'historiographie depuis le 19^e siècle. À cette fin, nous nous sommes inspirés de la synthèse de Didier Bondue, « De servus à sclavus : la fin de l'esclavage antique (371-918) » qui dénombre six tendances majeures⁵⁹. Malgré un apparent anachronisme, cet ouvrage demeure très pertinent car il faut rappeler que c'est bien au cours de la transition entre le Bas-Empire et le Haut Moyen Âge que le sens de *seruus* est le plus susceptible d'avoir évolué. De fait, comme nous l'avons constaté au travers des dictionnaires, le mot revêt un caractère, au Moyen Âge, lié au sacré. Or, c'est à partir du 4^e siècle que l'Occident adopte favorablement le christianisme, contexte dans lequel le terme est le plus employé. Cette analyse historiographique sera principalement tournée vers la France, un peu vers l'Allemagne et vers l'Italie, plus superficiellement.

1.2.1 L'école historique de la deuxième moitié du 19^e siècle et du début du 20^e siècle

Plusieurs auteurs ont abordé la question des institutions publiques au Moyen Âge et ont, pour ce faire, compilé bon nombre de documents qui furent employés par leurs suc-

⁵⁸ BONDUE, D., *De servus à sclavus : la fin de l'esclavage antique (371-918)*, p.17.

⁵⁹ *Idem*.

cesseurs⁶⁰. Nommons parmi ces précurseurs des auteurs tels que Fustel de Coulanges et son *Histoire des institutions politiques de l'ancienne France*, dans lequel il traite des origines de la féodalité et pour qui « le servage, communément objecté, n'est pas lié en fait à la féodalité (il lui préexistait) »⁶¹. Il brise avec la tradition romantique qui le précède et estime une certaine continuité avec les schémas antiques. Là où la tradition de l'époque pointait une véritable rupture induite par les invasions barbares⁶², lui portait la nuance avec méthodologie. Il fut l'un des pionniers en matière de critique historique. Pour le citer : « Lois, chartes, formules, chronique et histoires, il faut avoir lu toutes ces catégories de documents sans en avoir omis une seule. Car aucune d'elles, prise isolément, ne donne une idée exacte de la société »⁶³. Son analyse était donc transversale et attaquait partiellement chaque typologie.

Nous pouvons également citer le travail de Benjamin Guérard dans ses *Prélégomènes*, *Polyptyque* de l'abbé Irminon, qui soutient que « l'esclavage est toujours allé en se mitigant [...] depuis la conquête des Gaules par Jules César jusqu'à l'abolition de la féodalité »⁶⁴. Avec un certain recul, nous constatons qu'une partie de ses travaux proposent une méthodologie similaire. Guérard s'est fait l'analyste de nombreux autres polyptyques et cartulaires desquels, comme la citation susmentionnée en atteste, il tire des conclusions parfois généralistes et transversales. Néanmoins, ses travaux et ceux de Fustel de Coulanges, traitant de l'étude de la condition des terre et des personnes étaient, dans la première moitié du 20^e siècle, toujours perçus comme considérables et d'actualité⁶⁵.

À l'inverse de ceux que nous venons de présenter, certains auteurs ont le mérite d'avoir consacré un ouvrage entier à cette question du servage. C'est notamment le cas de Jean Yanoski qui rédige en 1860 « De l'abolition de l'esclavage ancien au Moyen Âge et de sa transformation en servitude » à la suite du travail de Henri Alexandre Wallon et de son « Histoire de l'esclavage dans l'Antiquité »⁶⁶. Notons que le titre du livre réfère non pas au concept de servage, comme nous pourrions l'attendre, mais plutôt à la servitude, plus englobante. Il approche la question de l'esclavage au travers d'un prisme chrétien, et bien qu'il diversifie ses sources (lois barbares, Grégoire de Tours, *Vitae*, chartes,...) il considère qu'en « dehors de la loi, le christianisme, par ses doctrines et par ses prédications

60 *Ibid*, p.18.

61 HARTOG, F., *le XIX^e siècle et l'histoire : le cas Fustel de Coulanges*, 1988, p.81.

62 PFISTER, C., BLOCH, M. et DIEHL, C., *Le centenaire de Fustel de Coulanges (1830-1889)*, dans *Revue internationale de l'enseignement*, t.84, 1930, p.174.

63 DE COULANGES, F., *Histoire des institutions de l'ancienne France : la monarchie franque*, 1888, p.29

64 BONDUE, D., *De servus à sclavus : la fin de l'esclavage antique (371-918)*, p.18.

65 DÉLÉAGE, A. et BLOCH, M., *Les caractères originaux de l'histoire rurale française. Compte rendu de OSLO, H.*, dans *Annales du midi : revue archéologique, historique et philologique de la France méridionale*, t.44, n°174, p.211.

66 BONDUE, D., *De servus à sclavus : la fin de l'esclavage antique (371-918)*, p.18.

avaient agi sur les mœurs des hommes d'une manière plus efficace »⁶⁷. Autrement dit, pour lui, la doctrine chrétienne impacte davantage la vie des médiévaux que les lois. Il va plus loin encore lorsqu'il affirme que « la civilisation romaine et le christianisme avaient introduit dans l'esclavage des changements qui auraient amené infailliblement, après un temps plus ou moins long, une émancipation générale »⁶⁸.

Quelques années plus tard, Paul Allard s'inscrit dans la continuité idéologique de Yanoski lorsqu'il annonce que « la diminution [du nombre d'esclaves] eut deux causes : l'une toute extérieure à laquelle le christianisme n'eut point part (décadence de l'empire) : l'autre toute rurale, qui doit lui être incontestablement attribuée »⁶⁹. Il rédigea deux ouvrages majeurs intégralement consacrés à la définition de l'esclavage et du servage, le premier datant de 1884 intitulé « Esclaves, serfs et mainmortables », et le second plus tardif, datant de 1913 titré « Les origines du servage en France ». Pour le second, il est intéressant de noter qu'il emprunte à Guérard son sujet d'étude, puisqu'il s'est également intéressé au Polyptyque d'Irminon⁷⁰. Enfin, nous pouvons également brièvement citer les ouvrages de Henri See (« Les classes rurales et le régime domanial de la France ») ou de Emile Lesne (« Histoire de la propriété ecclésiastique ») qui s'inscrivent aussi dans cette tendance propre à la fin du 19^e siècle et du début du 20^e.

Tous ces auteurs partagent donc pour la plupart une vision fort chrétienne de l'esclavage. Peut-on leur reprocher le tableau apologétique qu'ils dressent parfois à l'égard de l'Église ? En sachant qu'une grande part des sources qu'ils ont employées provient d'auteurs chrétiens (évêques, clercs, moines), il est compréhensible qu'ils abordent les sources avec un léger biais de confirmation. Cette position succède à une tendance précédente axée sur le romantisme et a le mérite d'offrir une base historiographique relativement stable.

1.2.2 La révolution Marc Bloch et l'essai de synthèse de Charles Verlinden

Le second tournant historiographique notoire advient lors de la publication de deux travaux majeurs de Marc Bloch dans la première moitié du 20^e siècle. Au travers de ses deux articles, il reconsidère les positions de ses prédécesseurs, il rompt avec leur héritage. Dans « Liberté et servitude personnelles au Moyen Âge, particulièrement en France : contribution à une étude des classes », publié en 1933, Marc Bloch s'intéresse à la question sémantique de *servus* dont l'importance fut reconnue par Georges Duby⁷¹, notamment, et

67 YANOSKI, J., *De l'abolition de l'esclavage ancien au Moyen Âge et de sa transformation en servitude*, p.35.

68 *Ibid*, p.2.

69 BONDUE, D., *De servus à sclavus : la fin de l'esclavage antique (371-918)*, p.18.

70 DUFFO, J. et ALLARD, P., *Les origines du servage en France*, dans *Revue d'histoire de l'Église de France*, t.5, n°28, p.512.

71 DUBY, G., *Sur les voies ouvertes par Marc Bloch : esclavage et servage au Moyen Âge*, p.126.

s'adonne à la définition juridique de la condition servile. D'après lui, ce sont « trois charges surtout, par leur réunion, [qui la] caractérisaient, au regard de la jurisprudence »⁷², à savoir le cheutage, la mainmorte et le formariage. Mais ses thèses furent réfutées dès 1939 par l'historien belge Léo Verriest dans un article de la Revue du Nord où celui-ci affirme que ces trois charges pesaient autant sur les libres que sur les serfs et étaient non pas des signes de servitude, mais de dépendance⁷³. Dans l'ensemble, comme le soulève Alain Derville, Marc Bloch n'était pas toujours pionnier et empruntait à ses devanciers. À l'instar de Fustel de Coulanges, il estimait que la servitude médiévale était un découlement de l'esclavage antique. C'est en tout cas la thèse qu'il évoque dans son second article, publié de manière posthume en 1946 : « Comment et pourquoi finit l'esclavage antique ? ». Selon ses mots, « en rompant avec l'esclavage, le Moyen Age, dont les mœurs sociales n'étaient pas tendres, n'a certes pas détruit, ni prétendu détruire, les inégalités de fait ou de naissance ; mais il leur a donné, si l'on peut dire, une tonalité plus humaine »⁷⁴.

Dans sa continuité, nous pouvons également brièvement citer l'apport de Léo Verriest, historien de la seconde moitié du 20^e siècle et détracteur de Marc Bloch. Son ouvrage « Institutions médiévales : Introduction au « Corpus des Coutumes » et des Lois de chefs-lieux de l'ancien comté de Hainaut » s'intéresse à la question de la liberté avec un regard de juriste. Il pense « des problèmes anciens, placés dans un éclairage nouveau »⁷⁵. Malgré l'envergure et la qualité de l'entreprise qu'il a effectué, nous pouvons cependant lui reprocher son manque de rigueur, il « n'échappe pas non plus à cette pauvreté des mots » qu'il attribuait si aisément à Marc Bloch⁷⁶.

L'historien Charles Verlinden s'est également illustré sur cette problématique de la servitude. Ce dernier a écrit plusieurs ouvrages dans lesquels il interroge d'innombrables sources, et il se distingue notamment par la zone qu'il couvre. Son premier ouvrage majeur est publié en 1955 (« L'esclavage dans l'Europe médiévale, tome 1. Péninsule ibérique-France ») et rassemble la synthèse du dépouillement de « lois, Fueros, cartulaires, chroniques, registres de notaire, documents de la pratique commerciale »⁷⁷. Impossible de les lister toutes, car comme il le dit lui-même, « les sources [...] utilisées sont excessivement nombreuses et diverses. Elles le sont même tellement que la publication d'une bibliographie n'aurait guère de sens »⁷⁸. Un travail colossal dont l'ampleur ne sera que multiplié

72 BLOCH, M., *Liberté et servitude personnelles*, dans *Mélanges historiques*, p.290.

73 DERVILLE, A. et LOCH, M., *Rois et serfs et autres écrits sur le servage (Compte rendu)*, 1996, p.191-192.

74 BLOCH, M., *Comment et pourquoi finit l'esclavage antique ?*, dans *Mélanges historiques*, p.285.

75 DOEHAERD, R., *Institutions médiévales : un livre, une méthode*, dans *Annales. Économies, Sociétés, Civilisations.*, 4^e année, n°1, 1949, p.23.

76 *Ibid*, p.27

77 RENOUEAU, Y., *Charles Verlinden, compte-rendu*, dans *Bulletin hispanique*, t.58, n°3, 1956, p.374.

78 VERLINDEN, C., *L'esclavage dans l'Europe médiévale*, t.1, p.23.

quelques années plus tard, en 1977, lorsque paraît un second volume, cette fois centré autour de l'esclavage en Italie et au Levant⁷⁹. Mais son travail ne s'arrête pas là, car une grande majorité de celui-ci tourne autour de la question de la servitude. Il était partisan de l'idée que « slaven in de middeleeuwen een zeldzaamheid waren, die men alleen aantrof gedurende de eerste paar eeuwen na de val van Rome »⁸⁰ et rejoignait donc Marc Bloch au sujet d'une continuité entre système esclavagiste antique et servitude médiévale. Il le rejoint aussi dans son constat d'un certain déclin de l'esclavage à partir de la période carolingienne. Enfin, il est également l'auteur d'un article en sémantique historique : « L'origine de *sclavus* = esclave »⁸¹, dans lequel il retrace l'histoire étymologique et contextuelle du mot *sclauus*. Toutefois, comme le rappelle Didier Bondue, il n'a pas satisfait « les tenants d'une approche plus économique et anthropologique du sujet ».⁸²

1.2.3 L'entreprise néo-marxiste

Ce courant de pensée en vogue depuis l'époque de Georges Duby remet au centre de l'analyse historique le prisme marxiste de la lutte des classes au travers d'études juridico-économiques. Auteur d'une thèse archétypale, représentative de la « grande monographie régionale à la française »⁸³, Duby trouve sa place dans l'historiographie car il est à l'origine de la création du paradigme dit « mutationniste ». Lui, ainsi que d'autres historiens de même tendance, considèrent le 11^e siècle comme un siècle de mutation, de « révolution féodale »⁸⁴ « de l'an mil »⁸⁴, qui « tend à maintenir et renforcer la domination de classe de l'aristocratie »⁸⁵. L'hypothèse rejoint une tendance « marxisante » en cela qu'elle suggère une véritable dichotomie entre les classes. Mais Duby soutient également qu'il existe une opposition forte entre l'antiquité et le féodalisme. Cependant, il ne nie ni ne conteste la continuité du système antique jusqu'aux alentours du 11^e siècle. Dans son ouvrage « Guerriers et paysans VIII-XII^e siècles, premier essor de l'économie européenne »⁸⁶, il définit la situation servile comme suit :

« Dans l'Europe du VII^e et VIII^e siècles, les textes qui subsistent révèlent parfois la présence d'hommes et de femmes très nombreux, que le vocabulaire latin nomme *servus* et *ancilla*, ou qui sont désignés par un substantif neutre, *mancipium*, exprimant plus clairement leur situation d'objets. Ils appartiennent en effet totalement à un maître depuis leur naissance jusqu'à la

79 VERLINDEN, C., *L'esclavage dans l'Europe médiévale*, t.2, *Italie, colonies italiennes du Levant, Levant latin, empire byzantin*.

80 VERLINDEN, C., *Slavenhandel en economische ontwikkeling in midden-, oost-, en noord-Europ gedurende de hoge Middeleeuwen*, p.3.

81 VERLINDEN, C., *L'origine de *sclavus* = esclave*, 1943.

82 BONDUE, D., *De *servus* à *sclavus* : la fin de l'esclavage antique (371-918)*, p.19.

83 CARRIER, N., *De l'esclavage au servage dans le royaume de Bourgogne (VIII-XII^e siècle)*, p.2.

84 Le concept de « révolution féodale » est dû à Georges Duby dans son ouvrage *Les trois ordres ou l'imaginaire du féodalisme*, Paris, 1978, p.183-205.

85 URBANCZYK, P., *Europe around the year 1000*, 2001, p.12.

86 DUBY, G., *Guerriers et paysans VIII-XII^e siècles, premier essor de l'économie européenne*.

mort, et les enfants que la femme esclave met au monde sont appelés à vivre, à l'égard du propriétaire de leur mère dans la même dépendance que celle-ci. Ils n'ont rien à eux. Ce sont des instruments, des outils animés que leur possesseur utilise à sa guise, entretient s'il lui plaît, dont il est responsable devant les tribunaux, qu'il punit comme il veut, qu'il vend, qu'il abrite ou qu'il donne ».

Son impact dans l'historiographie est incontestable, et cela malgré ses nombreux détracteurs contemporains. Des historiens plus radicaux émergèrent à sa suite, comme c'est le cas de Pierre Bonnassie pour qui il n'existait pas de raisons de ne pas traduire le terme *servus* par esclave avant le 11^e siècle⁸⁷. Dans son article « Survie et extinction du régime esclavagiste dans l'Occident du Haut Moyen Âge »⁸⁸, Bonnassie aborde le problème servile avec une approche proprement économique-juridique et dans lequel il recense trois facteurs possibles du recul de l'esclavage, à savoir un facteur religieux, un facteur militaire et un facteur économique⁸⁹. Son raisonnement est axé sur la recherche de rentabilité. Il se veut héritier de Marc Bloch et conteste une partie des travaux réalisés avant lui, comme c'est le cas pour la synthèse de Charles Verlinden dont il regrette l'étude uniquement juridique, sans laisser la place à une analyse anthropologique. En cela, il rejoint les tendances de son époque, à savoir le mouvement de la « Nouvelle Histoire » initié par Jacques Le Goff, dont la génération portait intérêt à cette discipline.

D'autres approches du même genre ont été réalisées dans le courant de la décennie. Des auteurs tels que Claude Meillassoux⁹⁰, Guy Bois⁹¹ ou Adriaan Verhulst⁹² se sont également démarqués de la tradition historiographique du début du siècle en abordant la question sous couvert d'anthropologie. À leur suite, un ouvrage de synthèse parut en 1991, fruit de la collaboration de Pierre Poly et d'Eric Bournazel⁹³. Au lendemain de sa parution, de vives critiques à l'égard de ce paradigme dominant (ré-)émergèrent. À leur tête, nous retrouvons Dominique Barthélémy qui relativise les observations des deux auteurs et regrette que « la documentation [soit] lacunaire et aléatoire »⁹⁴. Les documents seraient inégalement répartis et occasionneraient en cela un biais d'échantillonnage initial dont les auteurs ne se seraient pas départis, bien au contraire. Mais il ne se limite pas à l'ouvrage

87 CARRIER, N., *De l'esclavage au servage dans le royaume de Bourgogne (VIII-XII^e siècle)*, p.3.

88 BONNASSIE, P., *Survie et extinction du régime esclavagiste dans l'Occident du haut moyen âge (IV^e-XI^e s.)*, dans *Cahiers de civilisation médiévale*, 28^e année, n°112, 1985, p.307-343.

89 BONDUE, D., *De servus à sclavus : la fin de l'esclavage antique (371-918)*, p.20.

90 MEILLASSOUX, C., *Anthropologie de l'esclavage*, Paris, 1986

91 BOIS, G., *La mutation de l'an mil. Lournand, village mâconnais de l'Antiquité au féodalisme*, Paris, 1989.

92 VERHULST, A., *The decline of Slavery and the Economic expansion of Early Middle Ages*, dans *Past and Present*, n°133, 1991, p.195-203.

93 POLY, P. et BOURNAZEL, E., *La mutation féodale, X^e-XII^e siècles*, Paris.

94 BARTHÉLÉMY, D., *La mutation féodale a-t-elle eu lieu ?*, dans *Annales, Économies, Sociétés, Civilisations*, 47^e année, n°3, 1992, p.770.

et reconsidère la théorie de la mutation féodale dans son ensemble depuis les thèses de Georges Duby (1953) et de Pierre Bonnassie (1975)⁹⁵.

Il ne faut cependant pas croire que le désaccord débute avec Barthélémy. En effet, la genèse du débat prend racines bien avant les années 1970 et donne lieu à bon nombre de publications et d'évolutions. Mais ce n'est vraiment qu'à partir de 1975 que se révèlent les problèmes liés au paradigme mutationniste, et les premières critiques, issues alors du monde anglo-saxon⁹⁶. Barthélémy apparaît dans une troisième phase d'opposition, plus assumée, clamant que « ce paradigme doit être remis en cause »⁹⁷. Cette thèse fut très récemment remise à l'ordre du jour par Nicolas Carrier, qui résume le débat historiographique animé en soulevant une fois de plus les problèmes posés par la théorie mutationniste. Ce qu'il pointe principalement du doigt, c'est le désintérêt des partisans de Duby pour l'analyse sémantique à proprement parler. Pour citer Dominique Barthélémy et son appel à revenir à « l'ancienne école » (Bloch, de Coulanges) :

« Ceux-là savaient voir, sous la fixité du vocabulaire, les changements insensibles, ou malgré les mutations sémantiques, les permanences réelles »⁹⁸.

En réponse aux mutationnistes, plusieurs historiens en sont venus par la suite à ré-envisager l'esclavage médiéval non plus par le statut économique-juridique (comme le firent les auteurs à tendance marxiste), mais plutôt par le statut politico-juridique des hommes sujets à asservissement⁹⁹. Nous retrouvons en figure centrale de ce mouvement Elisabeth Magnou-Nortier et son article « Servus – Servitium : une enquête à poursuivre » dans lequel elle fait état de la confusion qu'il existe entre la servitude propre et la servitude d'État. C'est lorsqu' « un prisonnier de guerre est [...] arraché à sa patrie, à sa famille, à ses biens, et [qu'] il a perdu sa liberté » et que son statut devient celui du *servuus* que l'on peut parler de servitude d'État¹⁰⁰. Elle se sert, dans sa démarche, des taxes qui pèsent sur leur tête, du poids des institutions sur les individus.

1.2.4 Les écoles allemande et italienne

L'historiographie allemande a ceci d'intéressant qu'elle ne s'inscrit pas dans la tradition française d'une forme de liberté dichotomique, mais envisage des « catégories, des degrés différents dans la liberté »¹⁰¹. Pour ce faire, les historiens ne se basent plus sur les

95 *Ibid*, p.768.

96 URBANCZYK, P., *Europe around the year 1000*, 2001, p.18.

97 BARTHÉLÉMY, D., *La mutation féodale a-t-elle eu lieu ?*, dans *Annales, Économies, Sociétés, Civilisations*, 47^e année, n°3, 1992, p.767.

98 CARRIER, N., *De l'esclavage au servage dans le royaume de Bourgogne (VIII-XII^e siècle)*, p.3.

99 BONDUE, D., *De servus à sclavus : la fin de l'esclavage antique (371-918)*, p.21.

100 MAGNOU-NORTIER, É., *Servus - servitium : une enquête à poursuivre*, dans *Media in Francia. Recueil de mélanges offert à Karl Ferdinand Werner à l'occasion de son 65^e anniversaire par ses amis et collègues français*, p.279.

101 BONDUE, D., *De servus à sclavus : la fin de l'esclavage antique (371-918)*, p.22.

vocables latins mais sur les termes allemands de « Sklave, Knecht, Unfrei », par exemple. C'est particulièrement le cas de Hermann Nehlsen, dans son ouvrage « Sklavenrecht zwischen Antike und Mittelalter », publié en 1972, dans lequel il se base sur une remarque de Tacite dans la Germanie révélant que « d'autres asservis sont utilisés d'une autre manière que dans nos mœurs dans des emplois familiaux »¹⁰². Les schémas romains et les schémas germaniques ne sont donc originellement pas les mêmes lorsque est abordée la question de la servitude. À sa suite, citons également Gabriele Van Olberg¹⁰³ qui s'évertue à décrire et à recenser tous les mots proches des trois grandes catégories d'hommes libres au travers des actes législatifs ou juridiques des peuples barbares. Enfin, Bondue fait également mention de Heike Grieser et d'un ouvrage de synthèse au sujet de l'esclavage en Gaule entre le 5^e et 7^e siècle, d'après des sources chrétiennes¹⁰⁴.

Les historiens italiens se démarquent à leur tour des théories occidentales. Francesco Panero distingue deux écoles parmi les médiévistes, une école européenne et une école italienne, dont la différence majeure résiderait dans la perception de la distinction entre libres et non-libres, ce qui a engendré pour les Européens la création de la notion de servitude mitigée appelée servage. Les Italiens sont davantage partisans de l'idée que cette distinction a persisté tout au long du Moyen Âge et que les *servi* du Haut Moyen Âge n'étaient pas sujets de droit. Cela s'oppose quelque peu à la définition française qui observe dans les chartes et dans les documents du droit un flou juridique qui rend difficile la différenciation des statuts¹⁰⁵. Auteur d'une bibliographie des plus foisonnantes, Panero a surtout consacré ses lignes à la contradiction des thèses des spécialistes européens de l'histoire des campagnes et des seigneuries, comme c'est le cas notamment avec Marc Bloch¹⁰⁶. Il reste néanmoins un pilier de l'historiographie italienne.

Les approches, les sources, les méthodes, tout cela diffère dès lors que l'on quitte la Francophonie. Ce que les Allemands et les Italiens nous apprennent, c'est qu'il existe d'autres manières d'aborder cette question de la servitude, qu'il s'agisse d'un traitement sémantique basé sur l'allemand et non le latin, ou bien de la reconsidération du concept dans son ensemble.

102 *Idem*.

103 VAN OLBERG, G., *Freie, Nachbarn Gefolgsleute*, 1983.

104 GRIESER, H., *Sklaverei in spantiken und früe mittelalterlichen Gallien*, 1997.

105 CARRIER, N., *De l'esclavage au servage : pour une étude des dynamiques de la servitude*, dans *Médiévales*, n°81, 2022, p.179-196.

106 PANERO, F., *Le nouveau servage et l'attache de la glèbe aux XII^e et XIII^e siècles : l'interprétation de Marc Bloch et la documentation italienne*, dans *Mélanges de l'école française de Rome*, t.112/2, 2000, p.551-561.

1.2.5 L'historiographie récente

D'autres auteurs se sont démarqués au cours des dernières années dans leur entreprise de réévaluer les théories et concepts des auteurs susmentionnés. Tout d'abord, le premier ouvrage qu'il nous semble important de présenter est celui de Didier Bondue, publié en 2011. Celui-ci centre sa réflexion sur la thématique de l'asservissement dans le monde occidental entre le 4^e et le 10^e siècle, et plus particulièrement de « l'entrée en servitude, [des] tenants et aboutissements de l'asservissement, son cadre juridique, [de] l'affranchissement et [de] la situation des affranchis »¹⁰⁷. Son introduction a le mérite de présenter la problématique de la polysémie des termes relatifs à la servitude, mais la qualité de son travail est limitée par un manque de transversalité. Les dynamiques générales sont laissées de côté au profit de tableaux détaillés par lieux et par périodes. Son travail pose ses bases sur un réexamen des sources dans un cadre chronologique, en débutant par le décret de Valentinien Ier, jusqu'au diplôme de Konrad Ier en 918.

Un an plus tard, sous la plume de Nicolas Carrier, paraît un autre ouvrage qui se veut complémentaire, mais tout de même différent dans la méthode. « Les usages de la servitude. Seigneurs et paysans dans le royaume de Bourgogne (VI^e – XV^e siècle) » ne s'enracine pas dans une tradition historiographique, il renouvelle même les approches d'autres historiens (Bonnassie, Moses, Finley,...), et se structure d'une telle manière que les étapes de l'évolution de la servitude sont facilement mises en exergue¹⁰⁸. L'auteur a également le mérite de rappeler en introduction les définitions des termes centraux du sujet abordé, à savoir la servitude, l'esclavage, le servage et la dépendance. Mais l'ouvrage n'est qu'un exemple. Nicolas Carrier a en effet publié plusieurs travaux sur la question, dont le plus récent traite en partie de l'historiographie la plus contemporaine. Son importance n'est pas à sous-estimer.

En 2017, l'historienne Alice Rio publie son livre intitulé « Slavery after Rome (500-1100) », une excellente synthèse sur la transformation des institutions serviles entre l'Antiquité tardive et le Moyen Âge, qui brasse un territoire très large puisqu'elle intègre la totalité de l'Europe médiévale ainsi qu'une partie du territoire byzantin¹⁰⁹. L'autrice se penche sur les mécanismes du passage entre homme libre et non-libre, qu'il s'agisse de la guerre, de la pauvreté ou encore dans la recherche de protection. L'intérêt de cet ouvrage concerne justement l'amplitude qu'elle propose, car la création d'un modèle passe par la généralisation des similitudes. Or, à ce sujet, Alice Rio se montre critique à l'égard des

107 CARRIER, N., *De l'esclavage au servage : pour une étude des dynamiques de la servitude*, dans *Médiévales*, n°81, 2022, p.179-196.

108 ANDRIEUX, P., *Les usages de la servitude. Seigneurs et paysans dans le royaume de Bourgogne (VI^e-XV^e siècle)*, compte rendu dans *Revue historique du droit français et étranger*, vol.92, n°2, p.322.

109 FELLER, L., *Alice Rio, Slavery after Rome (500-1100)*, dans *Cahiers de civilisation médiévale*, n°243, 2018, p.318-320.

médiévistes français et de leur tendance généralisante. Elle leur reproche d'établir des modèles globaux à partir de sources uniquement françaises, de tirer des conclusions à partir d'échantillons trop limités. En outre, à l'instar de Nicolas Carrier, elle ne juge pas pertinent de comparer les catégories juridiques avec les catégories sociales¹¹⁰. Par plusieurs aspects donc, elle parvient à se montrer novatrice.

Enfin, citons aussi la contribution de Sandrine Victor et de son ouvrage « Les fils de Canaan »¹¹¹. Spécialiste d'histoire économique et sociale, l'autrice s'intéresse à la question de l'esclavage dans laquelle elle inclut la notion de servage. Elle y aborde les méthodes de renouvellement de main d'oeuvre servile détaillant le processus évolutif entre ces pôles que représentent l'esclavage antique et l'esclavage moderne. La brièveté de l'ouvrage force toutefois à reconsidérer son importance, car la place laissée à la période médiévale ne permet pas de rentrer dans le détail. En outre, l'ouvrage souffre également de quelques lacunes méthodologiques, comme le souligne Vincent Corriol. En effet, qu'il s'agisse d'une bibliographie trop ancienne ou lacunaire, ou encore d'un problème fondamental de définition, l'exhaustivité, la clarté et la méthodologie ne sont pas toujours au rendez-vous¹¹².

1.3 Approche complémentaire : la sémantique historique

Une fois les méthodes des historiens du passé abordées, nous nous sommes demandé comment faire pour dépasser leur approche. Si nous devons faire mention d'une lacune de la part de l'historiographie, ce n'est pas tant son désintérêt pour la question sémantique, au contraire, que l'efficacité limitée dont elle a faite preuve jusque-là. Depuis Isidore, la question du vocabulaire interpelle, questionne et de nombreux historiens se sont reposés, à sa suite sur les mots pour définir le monde. Certains comprennent qu'ils sont des outils clés pour décrire le réel, et qu'il faut être pointilleux dans leur usage (Bloch, Duby), d'autres considèrent les sens comme acquis et n'estiment pas les biais par lesquels ils attaquent leurs sources (Bonnassie). Or, s'il existe bien une difficulté en médiévistique aujourd'hui, c'est la compréhension des sens latins. Alain Guerreau a bien compris cela. À ses yeux, un mot n'a pas de sens intrinsèque, unique et statique, mais il dépend des relations qu'il entretient avec d'autres mots¹¹³. Sitôt que l'on soutient cette hypothèse, les concepts développés à partir des mots latins, s'ils ne sont exprimés avec nuance, ne sont plus si solides. Nous faisons face à de véritables colosses aux pieds d'argile.

¹¹⁰ *Idem.*

¹¹¹ VICTOR, S., *Les fils de Canaan*, Paris, 2019.

¹¹² CORRIOL, V., Sandrine Victor, *Les fils de Canaan : l'esclavage au Moyen Âge*, dans *Cahiers de civilisation médiévale*, t.249, 2020, p.95-96.

¹¹³ GUERREAU, A., slide 6 de <http://www.glossaria.eu/minorque/cours/Guerreau-5-Semantique.pdf>

1.3.1 La sémantique historique : une piste pertinente ?

Cette tâche de définition requiert toutefois prudence et compétence. Les médiévistes qui se penchent sur la question, se heurtent rapidement à l'ambiguïté inhérente de la langue latine qui empêche de fixer un sens univoque aux mots les plus fréquents, et c'est avec pragmatisme qu'ils se doivent de faire le lien entre les mots employés et les réalités sociales qu'ils traduisent. Nous devons donc d'abord penser à exclure les explications sémantiques des dictionnaires contemporains, pris dans le piège de la traduction et de la définition illusoire¹¹⁴. Il est en effet nécessaire d'en faire abstraction, car comme nous l'avons soulevé précédemment, une grande partie des dictionnaires fige les mots. Or, rappelons-le, les sens changent et avec eux les réalités. Nous savons grâce à Alain Guerreau que « les structures qui apparaissent dans un corpus antique et dans un corpus médiéval sont, presque toujours, radicalement différentes »¹¹⁵. Les mots prennent racine dans une société et s'imprègnent de ses codes, et en même temps, ils permettent à celle-ci de se réinventer, de se forger continuellement. Les mots sont à la fois les vecteurs et les témoins du changement.

Vers quoi se tourner alors ? Qu'existe-t-il comme outil pour réduire le champ de notre ignorance ? L'une des pistes les plus prometteuses reste à ce jour la sémantique historique qui essaie de développer des méthodes descriptives d'un point de vue diachronique. Elle « retrace le changement dans le sens des mots individuels, [...] elle essaie de systématiser et de catégoriser les types de changement de sens [et] décrit la façon dont le vocabulaire d'une langue évolue »¹¹⁶. Elle prend donc en considération ce caractère sémantique changeant. Cette nouvelle méthode emploie des outils différents, dont la lexicométrie, depuis l'essor des technologies informatiques.

Nous allons nous intéresser dans un premier temps à cette méthode se trouvant à la frontière de plusieurs disciplines (histoire, linguistique, sémantique) et évaluer ses résultats, ses avantages et ses limites épistémologiques. Notre intervention sera plus spécifiquement dirigée vers l'étude des termes médiévo-latins, puisqu'il s'agit de notre sujet d'étude. Puis, nous nous pencherons pareillement sur l'outil d'analyse lexical qu'est la lexicométrie avec une approche tout aussi critique.

1.3.2 La sémantique diachronique

La sémantique historique s'est affirmée comme une discipline en tant que telle, avec ses problématiques propres, au cours de la seconde moitié du 20^e siècle, dans le sillon des travaux de Reinhart Kosseleck. Selon Brigitte Nerlich, la sémantique historique se

114 BON, B. et GUERREAU-JALABERT, A., "*Pietas*" : réflexions sur l'analyse sémantique et le traitement lexicographique d'un vocable médiéval, dans *Médiévales*, n°42, 2002, p.77.

115 GUERREAU, A., slide 22.

116 HANCIL, S., *Histoire de la langue anglaise*, p.129.

distingue entre tradition française, tradition allemande et tradition anglaise (même si à la base, toutes émergent de la *Begriffsgeschichte* allemande)¹¹⁷. Chacune a connu une évolution qui lui est propre, mais l'une des leçons fondamentales que toutes enseignent, c'est que pour qu'un mot puisse changer librement dans l'usage ordinaire du langage, il faut passer par un oubli total de l'étymologie¹¹⁸. Seuls comptent le mot et le contexte dans lequel il apparaît. Il n'est pas nécessaire de s'attarder ici sur l'historiographie de la discipline¹¹⁹, mais il reste peut-être pertinent de soulever le manque d'engouement auprès des historiens, les plus réticents étant les historiens français qui tiendraient les chiffres en horreur¹²⁰. Et c'est bien là un problème, car l'un des aspects majeurs de la discipline aujourd'hui, c'est justement l'art de compter les mots.

Très tôt les sciences humaines (linguistiques particulièrement) ont opté pour l'intégration des méthodes mathématiques afin d'explorer d'autres champs de possibles. Dès lors que le vocable fut établi en tant qu'unité lexicale, il fut permis de quantifier le vocabulaire d'une œuvre, d'établir des index et des tableaux de concordances, de schématiser des distributions de fréquence ou encore d'observer l'accroissement du vocabulaire. Derrière une lourdeur non dissimulée se cachait alors une précision extrême permettant d'apprécier les variations stylistiques d'un même auteur, ou mesurer la distance entre deux textes d'auteurs différents¹²¹.

Initialement, donc, la sémantique statistique outillait principalement les linguistes dans leurs recherches, mais depuis quelques années, certains historiens s'en sont armés à leur tour. L'histoire médiévale française s'est vue enrichie par plusieurs études de bonne facture qui ont visé très spécifiquement un terme ou un concept en particulier. En 2002, un premier article publié conjointement par Bruno Bon et Anita Guerreau-Jalabert partait déjà de l'analyse d'un vocable défini (*pietas*) pour débroussailler la réflexion sur la sémantique historique. Ils abordent le problème de la diachronie de manière conventionnelle, un peu à l'image de notre méthode première : l'évolution dans les dictionnaires et les lexiques médiévaux. Leur démarche est complétée par une confrontation de leurs résultats avec les dictionnaires de latin classique, ainsi que par les dictionnaires modernes et contemporains. Mais ils s'éloignent de toute tentative de traduction et ne recherchent du sens qu'à travers des expressions et des termes qui s'y rattachent. Selon eux, « le dévoilement des

117 KAYNAR, E., *De la philologie à la sémantique historique en études ottomanes*, dans *European Journal of Turkish Studies*, n°31, 2020, p.5.

118 NERLICH, B., *Avant-propos : La sémantique historique en Allemagne, en Angleterre et en France*, dans *Histoire Épistémologie Langage*, t.15, fasc.1, 1993, p.6

119 L'article d'Erdal Kaynar est suffisamment clair et exhaustif, et donne une vision riche et détaillée de la bibliographie.

120 GUERREAU, A., *Pourquoi (et comment) l'historien doit-il compter les mots ?*, dans *Histoire & Mesure*, vol.4, 1989, n°1-2, p.81.

121 MULLER, C., *La statistique lexicale*, dans *Langue française*, n°2, 1969, p.39.

valeurs d'un vocable ne peut éviter une approche plus large, de nature sémantique : tout vocable est inscrit dans un réseau lexical et sémantique et c'est sa place dans le réseau, les relations qu'il entretient avec d'autres vocables connexes ou antonymes qui déterminent ses usages »¹²².

Quelques années plus tard, en 2008, les deux auteurs signent une nouvelle publication tournée cette fois-ci sur la signification du terme trésor, approchée similairement. De fait, c'est une nouvelle fois au travers des dictionnaires que débute leur travail, bien que leur regard soit plus critique. Les dictionnaires sont désormais vus comme un point de départ efficace, utile pour cerner le réseau lexical gravitant autour de leur vocable et comme ils le disent si bien, « le bon usage des dictionnaires passe aussi par la prise en compte de leurs défauts »¹²³. Autrement dit, conscients des limites de leur outil de travail, ils ont prolongé leur réflexion par une enquête statistique dans les textes, résultant à l'évaluation des fréquences et la distinction de cooccurrences.

Le tournant de la sémantique historique se trouve dans l'emploi récent des technologies numériques. Alain Guerreau dans l'« Avenir d'un passé incertain », critiquant déjà les dictionnaires, jugés peu fiables quand ils donnent des traductions, annonçait aussi la centralité de l'outillage électronique. Son axiome de base est double et repose d'une part sur le besoin pour les médiévistes de s'inspirer des notions du linguiste Jost Trier : le Wortfeld et le Sinnbezirk, que l'on traduit par espace lexical et champ sémantique (ces deux concepts représentant à ses yeux « l'échelle la plus adéquate pour toute analyse sérieuse du sens des mots, parce que c'est l'échelle qui permet de reconstruire une structure »¹²⁴) et de l'autre les outils numériques. Car si on allie la démarche numérique à ces outils sémantiques, le potentiel d'objectivité devient alors considérable. C'est ce dont nous sommes témoins depuis les années 1990. La discipline a beaucoup gagné en popularité ces dernières années, grâce notamment au rôle phare joué par l'intégration des outils numériques.¹²⁵

En clair, comme c'est au travers des dictionnaires qu'ont commencé les premières études en sémantique des termes médiévo-latins, aborder le sens de *seruus* par son étymologie avait donc de pertinent la fixation d'une borne synchronique initiale. Mais l'étymologie a ses limites. C'est pourquoi, dans un second temps, il semblait judicieux de se tourner vers les dictionnaires et les lexiques médiévaux et de les consulter en nombre, car cela permet de réaliser une première comparaison diachronique de la sémantique. Toute-

122 BON, B. et GUERREAU-JALABERT, A., « *Pietas* » : réflexions sur l'analyse sémantique et le traitement lexicographique d'un vocable médiéval, dans *Médiévales*, n°42, 2002, p.86.

123 BON, B. et GUERREAU-JALABERT, A., *L'apport des dictionnaires dans une recherche sémantique, l'exemple du trésor*, dans *Listy Filologické/Folia Philologica*, vol.131, n°1-2, 2008, p.93.

124 GUERREAU, A., *L'avenir d'un passé incertain*, p.108.

125 KAYNAR, E., *De la philologie à la sémantique historique en études ottomanes*, dans *European Journal of Turkish Studies*, n°31, 2020, p.1.

fois, comme cela a été expliqué précédemment, à un certain stade, les dictionnaires sont à leur tour obsolètes. Il fallait donc employer un autre outil pour dépasser ces paradigmes conventionnels. C'est là qu'intervient l'outil principal de notre travail : la lexicométrie.

1.3.3 Définition de la méthode : qu'est ce que la lexicométrie ?

De manière très synthétique, le terme de lexicométrie désigne « la discipline qui prend en charge l'analyse informatisée du discours et du lexique »¹²⁶. Selon Leimdorfer et Salem, sous ce vocable se regroupe aussi « toute une série de méthodes qui permettent d'opérer, à partir d'une segmentation, des réorganisations formelles de la séquence textuelle et des analyses statistiques portant sur le vocabulaire »¹²⁷. Dans un sens plus large, nous pourrions même parler d'un « ensemble de méthodes de description des textes fondées sur des indicateurs statistiques »¹²⁸. Il s'agit donc d'une discipline plutôt récente qui exploite les capacités des technologies numériques pour appliquer aux textes et à la sémantique les techniques statistiques. Ses précurseurs, ses initiateurs, sont Pierre Guiraud (*Les caractères statistiques du vocabulaire*, publié en 1954) et les travaux du Père Roberto Busa (*Index Thomisticus*), véritable pionnier des Humanités Numériques. À partir des années 1965-1968, des laboratoires commencent à investir ce champ d'étude et le premier numéro de la revue *Computers and the Humanities* paraît également¹²⁹.

Inspirés par les succès de cette nouvelle linguistique, les historiens ont rapidement cherché à se l'approprier et l'ont perçue comme une méthode d'analyse propre à conférer à l'histoire sa qualité de science. Dans un premier temps, les débats sont nombreux au sein de la communauté des historiens, et les critiques, violentes. Les détracteurs de la lexicométrie vilipendent l'usage des statistiques en avançant le postulat que les chiffres n'ont aucun sens en eux-mêmes. Malgré l'opprobre, certains historiens continuent de s'y intéresser (Alain Guerreau, André Salem, Jean-Philippe Genet) et peaufinent les logiciels de lexicométrie (Alceste, Lexico3, Hyperbase)¹³⁰. L'avantage de la mathématisation de la recherche en linguistique est qu'elle permet de travailler sur des corpus bien plus étendus qu'auparavant et qu'elle assure une certaine part d'objectivité dans le traitement informationnel¹³¹.

126 SOUTI, C., *La lexicométrie ou l'analyse du discours assisté par ordinateur : ce que l'informatique et les mathématiques peuvent apporter à la littérature et la linguistique*, dans *Expressions*, n°1, 2015, p.123.

127 *Idem*.

128 POLO DE BEAULIEU, M.-A., *Panorama de la lexicométrie*, dans *Histoire & Mesure*, vol.2, 1989, p.173-197.

129 *Ibid*, p.176.

130 https://devhist.hypotheses.org/898#_ftn23

131 SOUTI, C., *La lexicométrie ou l'analyse du discours assisté par ordinateur : ce que l'informatique et les mathématiques peuvent apporter à la littérature et la linguistique*, dans *Expressions*, n°1, 2015, p.119.

Aujourd'hui, les outils ont encore gagné en qualité. En parallèle, de nombreuses bases de données ont été assemblées afin de donner aux quantitativistes des bases de travail plus solides (*CEMA*, *Corpus corporum*,...). Avec des données de plus en plus foisonnantes et des outils de plus en plus performants, la lexicométrie a de beaux jours devant elle. Des travaux d'envergure ont déjà pu être menés en sémantique historique sous le couvert de la lexicométrie, à commencer par le mémoire de Nicolas Perreaux (*L'eau, l'écrit et la société. Étude statistique sur les champs sémantiques dans les bases de données [CBMA]*) qui reprend exactement les codes précédemment décrits. Alain Guerreau se livrait une dizaine d'années auparavant à un exercice similaire en définissant cette fois le mot *vinea*¹³².

L'emploi de la voie sémantique, et particulièrement de l'outil lexicométrique, apparaît donc pertinent malgré les reproches qui ont pu leur être adressés. Tout d'abord, parce que cela permet de briser les barrières internationales de l'historiographie et des concepts, frontière à la communication internationale reconnue déjà par Marc Bloch au début du siècle précédent. « Une fois de plus, sous peine de sclérose, ce sont de vieux casiers, périmés et faussement commodes, qu'il s'agit de briser », disait-il¹³³. Mais également et surtout car cela attaque le problème à la base, à savoir le mot lui-même. En cela, il ne nie ni ne refuse les postulats des historiens qui nous ont précédés, mais permet justement de prolonger ou nuancer leurs hypothèses en les soumettant au jugement de la sémantique.

Attention toutefois à ne pas tomber dans les mêmes travers que certains philosophes que l'on rattache au *Linguistic Turn*, mouvement historiographique des années 70. Il s'agit là d'un mouvement de pensée initié aux États-Unis et inspiré par certains penseurs occidentaux, notamment français (Roland Barthes, Jacques Derrida). Celui-ci part du constat que toute réalité est médiatisée par le langage et par les textes, et que, par conséquent, toute recherche en dépend pour penser le discours. L'historien n'appréhende pas le réel directement et ne peut dépeindre qu'une image discursive de la réalité, non la réalité elle-même. Si l'on suit ces préceptes, la connaissance historique dans son ensemble ne peut être que relative, et aucune vérité propre ne peut en émerger. Contrairement aux sciences dites « exactes », il est difficile de tomber dans l'affirmation d'une vérité absolue, mais cela ne retire en rien le caractère scientifique de la discipline. Dès lors, nous nous retrouvons confrontés à une double barrière linguistique. L'une analytique, où nous dépendons des mots qui témoignent des réalités que l'on étudie, et l'autre expressive, car nous sommes otages de nos propres lexiques. Mais dans la recherche de la connaissance, le mot est la limite. C'est l'outil primordial pour penser, s'exprimer. Plus son sens sera clarifié, moins cette barrière posera problème. C'est pourquoi, aussi loin que nous poussera notre travail,

132 GUERREAU, A., *L'avenir d'un passé incertain*.

133 SCHMITT, J.-C., *L'anthropologie historique de l'Occident médiéval. Un parcours*, dans *L'Atelier du Centre de recherches historiques*, t.6, 2010.

nous garderons en tête que la réalité que le terme *seruus* exprime ne pourra jamais être absolument vraie.

2 Analyse synonymique de *seruus* au sein de son champ lexical

Deux synonymes absolus n'existent pas, il y a et aura toujours une subtile nuance entre les deux qui fait que l'un n'est pas l'autre. Nonobstant, ils peuvent être, relativement à un autre mot, plus ou moins proches sémantiquement. Pour le deviner, il existe des approches multiples qui feront l'objet de ce chapitre. Pour commencer, il convient de légitimer le focus établi sur *seruus* et de relativiser son importance vis-à-vis des autres lemmes au sens similaire et associé. Puis, viennent les techniques de comparaison. Les hypothèses et les approches étaient plurielles, mais ont été limitées à trois. Le premier axe est une axe de comparaison des cas sous la justification suivante : l'inégalité des occurrences des cas traduit des usages distincts. Le second relève d'une analyse statistique classique si ce n'est le filtre par lequel il passe. En effet, les analyses des moyennes, médianes, fréquences et écarts-types seront abordées au travers du prisme d'un indice appelé *TF-IDF*. Enfin, la dernière méthode conceptualisée pour l'occasion est celle du calcul d'indice de similarité entre les corpus de cooccurrences de chaque mot, sous le présupposé qu'une similitude de collocation pourrait exprimer la synonymie.

La volonté de dissocier *seruus* des autres termes latins auxquels on l'associe habituellement permettrait, en cas de vérification, de l'individualiser et de minimiser la quasi-synonymie qu'on lui attribue parfois dans l'historiographie. En outre, cela équivaldrait aussi à lui reconnaître une nuance qu'il s'agira de décoder par la suite. Dans cet ouvrage, seules les compétences en code peuvent s'avérer lacunaires et déformer drastiquement les résultats.

2.1 Le choix de *seruus*

Le choix spécifique du terme *seruus* a été décidé après plusieurs étapes de discrimination. Notre recherche brassait initialement tous les termes relatifs à l'esclavage médiéval dans les documents diplomatiques, mais au vu des déséquilibres observés et de l'importance constatée pour ce lemme, nous avons choisi de nous intéresser exclusivement à son cas. Ce constat fut d'abord reflété par une approche simple : une comparaison basique des occurrences des termes liés à notre question de recherche. Leur sélection possède une part d'arbitraire, car elle est issue de la lecture du livre d'Alice Rio¹³⁴ et des *Étymologies* d'Isidore de Séville¹³⁵. Initialement, le projet était de traiter des occurrences de chacun des termes ou expressions suivantes : *mancipium*, *ancilla*, *sclauus*, *aldius*, *seruus*, *ministerialus*, *seruus rusticanus*, *colonus*, *schiaui*, *villani*, *famulus*, *serua*, *serui casati* (*non casati*)

¹³⁴ RIO, A., *Slavery after Rome (500-1100)*, 2017.

¹³⁵ REYDELLET, M., *Isidore de Séville : Étymologies*, t.IX, p.178.

et *seruula*, mais la recherche dans la base de donnée des *CEMA* n'en a retourné qu'une liste réduite, comprenant *mancipium*, *ancilla*, *sclauus*, *seruus*, *colonus*, *famulus*, *serua* et *seruula*. Une subtile nuance reste à apporter cependant. Ne sont traités ici que les lemmes faisant référence à des substantifs, autrement dit, les termes possédant une racine similaire ne sont pas pris en compte s'ils ne font pas partie de cette classe grammaticale.

À y regarder de plus près, lorsqu'on l'on interroge la base de données pour les siècles concernés, la distribution des occurrences paraît équilibrée (Figure.4), à la double exception des termes *seruula* et *seruus*. En ce qui concerne la première, son importance peut être rapidement négligée car nous ne retrouvons que 20 occurrences du lemme dans l'entière du corpus, tandis que la seconde dépasse de loin les autres valeurs. Avec ses 34235 occurrences, le terme *seruus* surclasse d'un facteur sept le second terme le plus fréquent, *mancipium*. Un bref coup d'oeil au graphique permet de se rendre compte de l'importance de *seruus* mais il faut toutefois émettre une première réserve vis-à-vis de cette observation.

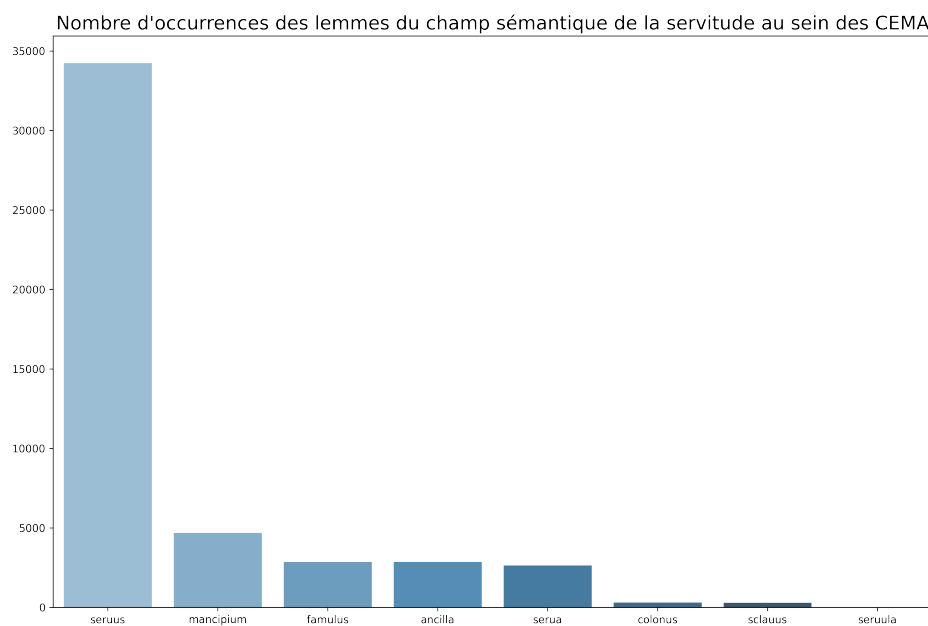


Figure 4 – Différence d'occurrences entre les lemmes.

Nous faisons face à un biais d'échantillonnage plutôt perceptible en cela qu'il influe drastiquement sur les conclusions que nous pourrions tirer. En effet, si l'on s'intéresse aux formes que *seruus* prend dans les chartes, nous remarquons qu'une très grande majorité de ses cooccurents renvoient à lui-même. Sur les 34235 occurrences, 27204 sont employées avec *seruus*. Après vérification, il semble que l'expression *seruus seruorum dei*

monopolise une grande majorité de ces cas de figure. Ce ne sont pas moins de 13498 fois que l'expression se présente dans notre corpus, soit 26996 occurrences de *seruus*. Lorsque *seruus* est employé avec lui-même, il est presque absolument certain qu'il le sera au sein de la formule d'humilité. Dès lors, la comparaison avec les autres termes se légitimise et trouve son sens uniquement lorsque la liste d'occurrences est épurée de ce bruit. Il n'existe en effet pas de formule similaire impliquant un autre des lemmes. Pour comparer ce qui est comparable, il est nécessaire d'en faire abstraction et de comparer les occurrences de *seruus* apparaissant dans le corps des chartes et qui ne sont pas suivies d'une autre forme d'eux-mêmes.

Pour y parvenir, nous avons développé un algorithme prenant en entrée un tableau .csv des occurrences de *seruus* à partir de la base de données des *CEMA*, sous forme lemmatisée. Après un passage par cet algorithme, le document initial supprime toutes les lignes où l'expression dont on désire se délester est présente. *In fine*, sur les 34235 lignes de tableau originelles, seules 7223 subsistent, soit à peine 20%. Autrement dit, le mot *seruus*, dans les chartes médiévales est principalement employé comme formule d'humilité. Cette fonction est absolument centrale dans notre recherche car elle est à la base de toutes les observations que nous allons effectuer.

```
colnames = ['indice', 'pre', 'word', 'post']
df = pd.read_csv('/home/noe/Documents/Master_2/Mémoire/final.csv', header = 4, names=colnames)
df = df[['pre', 'word', 'post']]
df['new'] = df['pre'] + ( ' ' ) + df['word'] + ( ' ' ) + df['post']
df['new'] = df['new'].str.lower()
df['drop'] = df['new'].str.find('seruus2 seruus2 deus ')
df = df[df['drop'] == -1]
df = df.reset_index()
df.pop('index')
df.pop('drop')
df.pop('new')
```

Figure 5 – Exemple de l'algorithme de suppression de *seruus seruatorum dei*.

Une fois cette précaution prise, nous nous sommes posés une question fondamentale, à savoir quelle était la part de synonymie entre *seruus* et tous les autres termes susmentionnés. Nous avons cherché à évaluer cette dernière au travers de deux axes comparatifs principaux. Tout d'abord, leur usage est-il semblable ? Sont-ils employés de la même manière, les retrouve-t-on accordés aux mêmes cas ? Les synonymes étant généralement employés dans des contextes similaires, il paraît logique que deux termes sémantiquement semblables soient employés dans des situations semblables. Ensuite, nous avons comparé les lexiques de cooccurrences de chaque lemme afin d'en calculer des indices de similarité. Les mots s'insèrent dans le texte selon certains schémas, encadrés par d'autres mots. Or le sens d'un vocable, en plus d'être fluctuant, dépend directement du sens des mots avec lesquels il est employé. Nous avons donc cherché à illustrer, à quantifier les différences ou les similarités qu'entretient chacun de ces termes avec *seruus*.

2.2 Comparaison des cas

Dans chaque charte, les termes sont accordés d'après un cas spécifique qui témoigne de sa fonction dans la phrase. Si l'analyse des cas entre nos lemmes concorde, qu'on les retrouve sous des formes aux proportions similaires, il devient possible que leur emploi soit associé. Encore fallait-il extraire ces données des .csv téléchargés à partir des *CEMA*. En effet, une première recherche en langage *CQL* sur l'interface en ligne permet d'obtenir un tableau manipulable de quatre colonnes. Tous les tableaux extraits à partir de la base de données ont la même apparence, à savoir :

- Une première colonne qui liste les numéros des chartes (pour l'identification).
- Une seconde qui comporte le texte de la charte avant l'occurrence du lemme ou du mot recherché.
- Une troisième qui se compose uniquement des lemmes ou des mots demandés.
- Et enfin une dernière qui se fait le miroir de la seconde, contenant quant à elle le texte de la charte qui suit l'occurrence.

2.2.1 Démarche méthodologique

Afin de pouvoir quantifier la proportion de chaque cas, il était nécessaire de nettoyer quelque peu le document. Toutes les colonnes ne sont pas pertinentes, et les chaînes de caractères ne sont pas uniformes. De fait, certains lemmes apparaissent selon deux typographies distinctes, l'une débutant en majuscule, l'autre non. Nous avons donc créé une fonction en langage python prenant en entrée un tableau csv et retournant en sortie un *dataframe* nettoyé et homogène des différentes formes trouvées dans la colonne `df['lemme']` avec leurs occurrences respectives.

```
def clean_csv(path):  
    df = pd.read_csv(path, sep=",", header = 4, encoding = "UTF-8")  
    df.columns = ["indice", "pre", "lemme", "post"]  
  
    df_final = df['lemme'].str.lower().value_counts()  
  
    return (df_final)
```

Figure 6 – Fonction python pour écarter l'expression *seruus seruorum dei*.

Cette fonction a été appliquée à chacun de nos fichiers à l'exception de celui du lemme *seruus*. Pour ce dernier, le nettoyage du bruit précédemment expliqué (*seruus seruorum dei*) a dû être effectué avec des manipulations préalables. Il a donc fallu récupérer ce .csv et le formater pour qu'il s'accorde aux autres lemmes. Pour ce faire, nous l'avons d'abord fait passer par le filtre du premier algorithme détaillé précédemment (Figure.5), puis lui

avons appliqué le même traitement que pour les autres lemmes.

Dans un second temps, nous avons attribué manuellement les valeurs d'index pour chacun des *dataframes* obtenus en fonction des cas des lemmes ¹³⁶. Puis, afin de travailler avec des valeurs relatives comprises entre 0 et 1, nous avons divisé les occurrences de chacune des formes par le nombre d'occurrences totales. Une fois cela fait, il ne restait qu'à fusionner les *dataframes* et appliquer une fonction d'illustration. Pour cela, nous avons choisi d'employer la librairie *seaborn* et sa méthode *heatmap*. Le résultat final nous renvoie une matrice dont chaque case indique doublement sa valeur. D'une part, la couleur permet de mesurer le poids relatif de chacune des cases, et de l'autre, les valeurs indiquées quantifient la dispersion au sein d'une variable. Toute cette analyse a été menée sur deux *datasets*, le premier exempt de nettoyage initial (Figure. 7) et le second filtré au moyen du code précédemment détaillé (Figure. 8). Le but de cette démarche était de voir comment évolue la proportion des cas lorsque l'on ôte du corpus 80% de ses occurrences.

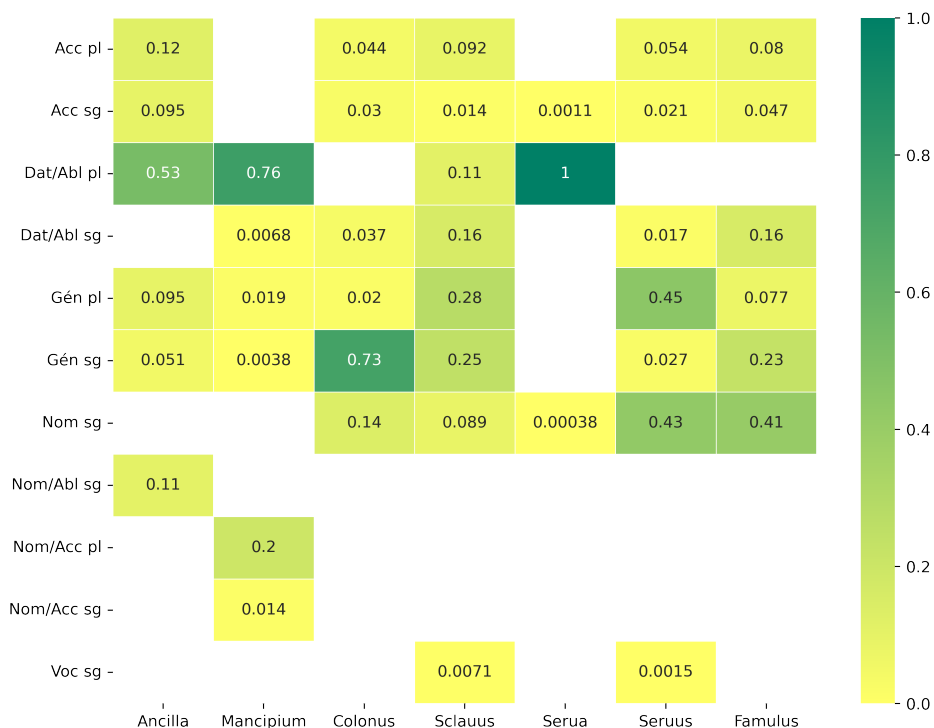


Figure 7 – Heatmap des cas avec la formule *seruus seruorum dei*.

¹³⁶ `dfSeruus.index = ['Gén pl', 'Acc pl', 'Nom sg', 'Gén sg', 'Acc sg', 'Dat/Abl sg', 'Voc sg']`

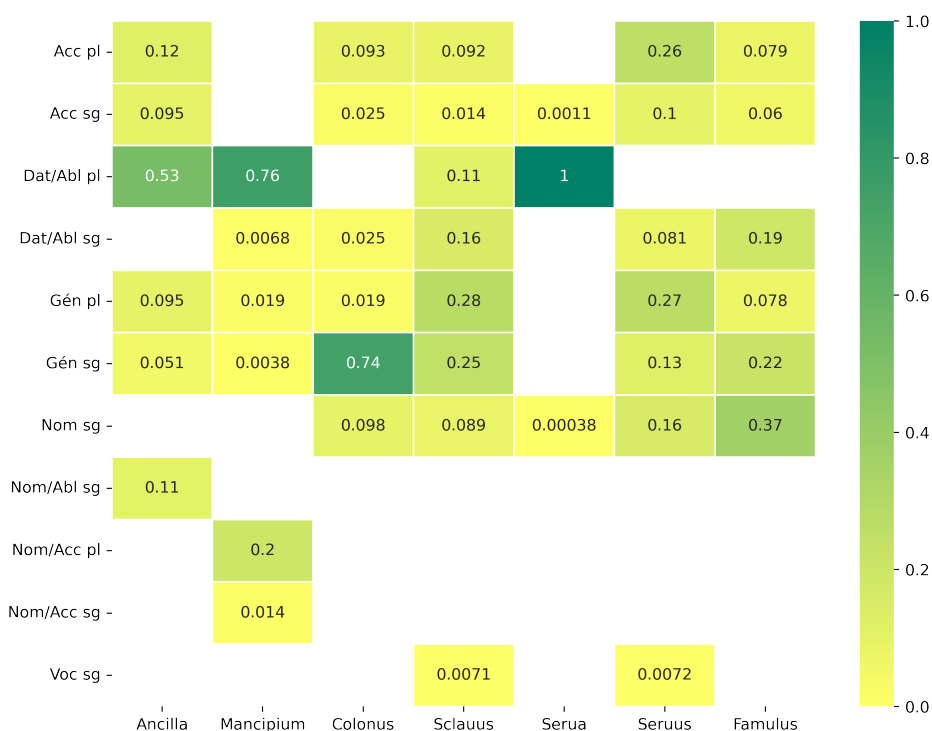


Figure 8 – Heatmap des cas sans la formule *seruus seruorum dei*.

2.2.2 Interprétation des résultats

À première vue, la première *heatmap* paraît disproportionnée. Près de 90% des cas totaux sont soit du nominatif singulier (*seruus*) soit du génitif pluriel (*seruorum*). La prépondérance de ces cas s'explique simplement par l'expression *seruus seruorum dei* et ses 13498 apparitions. Une fois que l'on s'en sépare, la répartition s'aplanit, s'équilibre. Bien que le génitif pluriel soit toujours l'expression majoritaire (27%), elle n'est désormais plus talonnée par le nominatif singulier, mais par l'accusatif singulier. Autrement dit, le *seruus* est moins sujet dans les chartes qu'il n'est objet. En outre, si l'on néglige le vocatif, sensiblement minoritaire, voire anodin, chaque cas a de l'importance et apparaît en moyenne au moins près de 10% du temps.

Qu'en est-il des autres lemmes? Pour les plus importants d'entre eux, la dispersion est bien moins harmonieuse. Par exemple, *mancipium* concentre à hauteur de 76% ses occurrences dans sa forme au datif ou à l'ablatif pluriel (*mancipiis*), et laisse encore un peu de place à sa forme au nominatif singulier (*mancipium*). Tous les autres cas se disputent les quelques pourcentages restants. Autant dire que leur impact frôle l'anecdotique. Ce problème, on le retrouve aussi, moins prononcé cependant, auprès du lemme *ancilla*. Si

l'on fait abstraction du génitif pluriel, du datif et de l'ablatif pluriel, la répartition s'accorde déjà plus avec celle de *seruus*. Il reste cependant difficile d'oublier l'importance de *seruorum*, d'autant plus si l'on ignore la formule d'humilité.

Une anomalie mérite également que l'on s'y arrête. La colonne de *serua* concentre la quasi-totalité de ses occurrences dans sa forme au datif ou ablatif pluriel. Sur ses 2628 apparitions, 2623 se retrouvent sous cette forme (*seruis*), et parmi elles, 834 proviennent de l'expression latine *seruis et ancillis*. On peut donc légitimement se poser deux questions. Tout d'abord, *seruis* est également la forme du datif et ablatif pluriel du terme *seruus*. Existe-t-il une faible possibilité pour que les attributions des lemmes effectuées dans les CEMA aient été mal réalisées ? Que toutes ces occurrences affectées à *serua*, l'aient été alors qu'une partie provenait de *seruus* ? Cette hypothèse expliquerait la désertion de cette forme dans sa colonne, et l'omniprésence dans celle de *serua*. En outre, dans l'hypothèse où cela s'avérerait vrai, cela permettrait de distinguer un premier écart sémantique entre *seruus* et *ancilla*. La présence de la particule 'et' sous-entend en effet que les deux termes sont différents. Maintenant, si les affectations ne sont pas erronées, cela implique que *serua* et *ancilla* fonctionnent de pair, ce qui paraît étonnant. La coordination de *seruus* et *ancilla* se justifie par la complétude des genres. *Seruus* renvoie au genre masculin tandis qu'*ancilla* est propre au genre féminin. Étant donné que *serua* appartient au genre féminin également, la formule devient moins logique. D'autant plus si l'on sait que pour tous les autres cas, c'est *seruus* qui est employé (*seruos et ancillas* ; *seruus et ancilla* ; *seruum et ancillam*,...).

Famulus, quant à lui, s'équilibre autour des cas au singulier, principalement le nominatif. C'est en cela qu'il diffère de l'usage de *seruus*. Celui-ci, en effet, fonctionne majoritairement au moyen de terminaisons au pluriel. Mais avant tout, c'est la dualité entre sujet et objet qui une nouvelle fois interpelle. De manière cumulée, *famulus* comptabilise la même proportion de nominatifs que *seruus* propose d'accusatifs. Il devient difficile de croire que ces deux vocables aient pu partager un usage intrinsèquement commun.

Enfin, *colonus* est disproportionnellement centré autour de sa forme au génitif singulier (74%), ce qui jure avec le partage pondéré constaté pour *seruus*. Contre toute attente, c'est *sclauus* qui se rapproche le plus de cet état de mesure, d'équilibre, même si ses occurrences les plus nombreuses ne se superposent pas tout à fait à son homologue.

Dans l'ensemble, aucun lemme ne trouve un équilibre sensiblement similaire à *seruus*. Peut-être pourrait-on relever la similarité des génitifs qu'il entretient avec *sclauus*, mais si l'on regarde le reste du tableau, on constate rapidement que la similarité s'arrête là. En ce qui concerne *ancilla* et *mancipium*, eux qui sont supposément les plus proches de se-

ruus, ils brillent par leur différence. Tous deux sont principalement accordés au datif ou à l'ablatif pluriel, cas pour lesquels *seruus* ne comptabilise aucune occurrence. Même si on lui ajoutait les occurrences de *serua* potentiellement mal attribuées, cela ne représenterait qu'un grand 25%, ce qui n'est toujours pas comparable. De la même manière, le désert d'occurrences à l'accusatif pluriel pour les deux lemmes est problématique car c'est là l'un des emplois majeurs de *seruus* dans les chartes.

Après toutes ces observations, il est nécessaire de prendre un peu de recul et d'analyser notre jeu de données différemment. La première réserve que nous nous devons d'émettre concerne la différence de taille des échantillons. Les comparaisons effectuées entre *seruus* et d'autres termes se justifient seulement si ces derniers détiennent un nombre d'occurrences significatif. En ce qui concerne ceux qui détiennent une fréquence basse, n'atteignant pas le millier d'apparitions, il est nécessaire de nuancer les résultats, qu'il s'agisse de *colonus* et de ses 298 occurrences ou de *sclauus* et ses 282 apparitions.

Bien entendu, cette approche n'a d'intérêt que jointe avec d'autres indicateurs. Utilisée seule, elle permet seulement de certifier que dans les documents diplomatiques, des documents formatés, aux formules abondantes, les emplois de tous ces lemmes ne sont pas identiques. Les causes peuvent être multiples comme la présence d'expressions fréquentes qui biaisent les résultats. Tout cela pour dire que, bien que pertinente, la démarche n'est pas suffisante.

2.3 Les schémas de pesage traditionnels pour les termes : *Term Frequency and Inverse Term Frequency*

Étudier la synonymie des mots, c'est également estimer le rôle qu'ils peuvent jouer au sein d'un corpus donné. En cela, tous les mots ne sont pas égaux. Certains apparaissent plus dans certains documents, d'autres moins, mais sont mieux répartis dans le corpus. Il est possible de mesurer le poids des mots. Pour cela il existe différents modèles. Le plus simple, et ce faisant, le moins significatif, consiste à attribuer une valeur à chaque mot (binaire, booléenne) selon sa présence ou son absence dans le document. Ce modèle s'es-souffle rapidement car il ne permet que de comparer de manière synthétique des textes et alloue des importances égales à des termes plus ou moins discrets ¹³⁷.

Une autre manière d'aborder les rapports de sens est de passer par le calcul de la fréquence. Celle-ci se définit comme étant la somme des apparitions d'un mot dans un document divisé par la totalité des mots du texte. Il s'agit là d'une approche plutôt basique qui essaye de mesurer l'impact qu'un mot détient sur le document. Dans l'absolu, si un mot

¹³⁷ NEGRE, E., *Comparaison de textes : quelques approches...*, 2013, p.6.

revient plus fréquemment qu'un autre, cela signifie qu'il dirige plus le sens global, qu'il est plus central pour la compréhension.

Au sein de l'étude de chartes médiévales, la comparaison de fréquence permet d'indiquer la gravité de chacun des mots. Mais cette démarche a une limite. Elle ne prend en compte que le poids d'un mot sur un seul et unique document, pas sur le corpus. Certes, la moyenne des fréquences sur l'entièreté du corpus est possible, et indiquera une tendance, mais elle n'exprime *in fine* que l'influence du mot, en moyenne, dans un document, pas dans l'intégralité du corpus. Il s'agit donc d'une première étape sensée et pertinente, mais incomplète.

Pour pallier ce manque, il existe une autre formule qui « étudie l'importance d'un terme contenu dans un document, relativement à une collection »¹³⁸ : le TF-IDF (*Term frequency – Inverse document frequency*). Cette notion mêle la fréquence des termes avec l'IDF (*Inverse Document Frequency*), qui assigne des poids plus faibles aux mots plus fréquents, et inversement¹³⁹. Ce faisant, les écarts entre les mots sont diminués, les « mots-valises » sont compensés. Le TF-IDF compile donc ces deux approches. En ce qui concerne les termes dont la fréquence est plus élevée, plus la fréquence des textes dans lesquels ils apparaissent est faible, plus ils sont représentatifs de leur document. Car s'il apparaît de nombreuses fois dans un document, mais que peu l'emploient, cela signifie qu'il est plus constituant du sens. Deux mots dont le score TF-IDF est proche jouent donc un rôle également important dans la charte et à l'inverse, s'il diffère, cela implique que l'un des deux est bien plus déterminant, et qu'en cela, son sens s'autonomise. La différence entre les scores TF-IDF peut donc indiquer une forme de dissimilarité dans l'emploi des mots.

Dans cette partie, l'objectif est de passer notre corpus au crible de ces méthodes. La comparaison porte toujours sur les mêmes termes : *mancipium*, *ancilla*, *famulus*, *serua*, *colonus* et *sclauus*, les deux derniers étant d'autant plus intéressants qu'ils comportent moins d'occurrences que leurs homologues. Nous allons donc présenter dans un premier temps les données de base et la méthode employée pour passer des tableaux bruts à la schématisation de nos résultats. Puis, nous aborderons l'analyse et l'interprétation de ces résultats. Que disent les fréquences booléennes, fréquences des termes et le TF-IDF de la place de chacun de ces termes dans les chartes médiévales ?

¹³⁸ *Ibid*, p.7.

¹³⁹ SHAHZAD, Q. et RAMSHA, A., *Text mining : Use of TF-IDF to Examine the Relevance of Words to Documents*, dans *International Journal of Computer Applications*(0975– 8887), vol.181, 2018, p.25.

2.3.1 Méthodologie

Les données initiales ne permettent pas une lecture spontanée des informations que nous recherchons. Il a fallu les manipuler quelque peu pour qu'elles répondent à nos attentes. Comme pour la réalisation de la *heatmap*, nous avons extrait ces données des *CEMA* sous forme de tableaux .csv comprenant quatre colonnes distinctes : le numéro de la charte (sa brève fiche d'identification), les quelques mots qui précèdent le terme de la recherche, le mot en question ainsi que les mots qui lui succèdent. Néanmoins, dans ce cadre spécifique, il a été nécessaire d'ajouter le texte intégral dans le tableau. En effet, l'une des variables les plus importantes dans le calcul des fréquences, c'est le nombre de mots que comporte chaque texte. Cette information était exportable et suivait l'indice de la charte dans la première colonne.

In fine, ce que nous désirons obtenir, c'est un tableau synthétique composé de quatre colonnes également, mais pas les mêmes. La première reprendrait la liste de toutes les chartes dans lesquelles un certain lemme apparaît. La seconde indiquerait le nombre d'occurrences de ce lemme au sein du document. La troisième listerait quant à elle la fréquence de ce mot dans la charte, tandis que la dernière en renverrait le score TF-IDF. Dans le cas de notre étude, la fréquence booléenne n'est pas utile car les indices qu'elle renverrait dans un tel tableau seraient tous égaux à 1.

Une nuance est également à apporter. Comme nous l'avons spécifié plus tôt, le terme *seruus* comporte un nombre relativement plus élevé d'occurrences dans le corpus. Mais ces occurrences sont, pour une très grande majorité d'entre elles, issues de l'expression latine *seruus seruorum dei* et représentent donc un bruit non négligeable dans l'analyse du sens du mot dans le texte. Se débarrasser de cette répétition a donc été le point de départ de notre travail. Toutefois, il importe de comparer les deux états du tableau, avec l'expression et sans elle. En ce qui concerne les autres termes, une fonction commune a pu être programmée pour tous.

Le nettoyage et l'organisation du tableau de *seruus* ont dû s'effectuer en trois temps et ont pris en entrée deux tableaux différents. Ces tâches ont pu être réalisées au moyen des lignes de code suivantes :

1. Normalisation du tableau initial en vue de se séparer de l'expression latine et transformation en un premier tableau :

```
def tab_1(path):

    colonnes = ['indice', 'pre', 'lemma', 'post']
    df = pd.read_csv(path, sep=',', header=4, names=colonnes, encoding='UTF-8')
    df['new'] = df['pre'] + (' ') + df['lemma'] + (' ') + df['post']
    df['new'] = df['new'].str.lower()
    df['drop'] = df['new'].str.find('seruus2 seruus2 deus')
    df = df[df['drop'] == -1]
    df = df.drop("drop", axis=1)
    df = df['indice'].value_counts()
    df = df.reset_index()
    df.columns = ['indice', 'occurrence']

    return(df)
```

Figure 9 – *Fonction d'évincement de seruus serurorum dei.*

2. Obtention du nombre de mots par texte et mise sur pied d'un tableau intermédiaire pour chaque lemme :

```
def csv_scatter_TF_IDF(path, path2):

    #le deuxième path sert à alléger la commande et ne pas itérer dans le gros dataframe

    colonnes = ['indice', 'pre', 'lemma', 'post']
    df = pd.read_csv(path, sep=',', header=4, names=colonnes, encoding='UTF-8')

    df2 = df['indice'].str.replace(',', ' ')
    df2 = df2.str.split()
    texte = df2.drop_duplicates()

    texte_final = []
    for i in texte:
        texte_final.append(i[1:])
    taille = []
    for i in texte_final:
        taille.append(len(i))

    dataframe = pd.read_csv(path2, sep=',', header=4, names=colonnes, encoding='UTF-8')
    final = dataframe['indice'].value_counts()
    final = final.tolist()
    indice = dataframe['indice'].drop_duplicates()
    indice = indice.tolist()

    df_final = pd.DataFrame(list(zip(indice, final, taille)))
    df_final.columns = ['indice', 'occurrence', 'nbr_tot_mots']

    df_final['TF-IDF'] = (df_final['occurrence']/df_final['nbr_tot_mots']) * (math.log10(244085/len(indice)))
    df_final['Fréquence'] = (df_final['occurrence']/df_final['nbr_tot_mots'])

    return df_final.describe()
```

Figure 10 – *Fonction de préparation au TF-IDF.*

Dans l'ensemble, la fonction prend en entrée deux .csv différents alors qu'un seul est nécessaire. Ce choix est motivé par le poids du document et la difficulté pour l'interface de Jupyter Notebook d'y faire tourner un code. Afin donc d'y minimiser les itérations, il a été décidé de prendre un second document dans lequel tous les textes n'apparaissent pas. Il restait cependant on ne peut plus nécessaire de le manipuler afin d'obtenir, pour chaque occurrence de chaque lemme, le nombre total de mots dans le texte, variable primordiale dans le calcul du TF-IDF.

3. Superposition des deux tableaux :

```
final = pd.merge(df_final,df, on = 'indice')
final = final.drop('occurrence_x', axis=1)
final.columns = ['indice', 'nbr_tot_mots', 'TF-IDF', 'Occurrences']
```

Figure 11 – *Tableau TF-IDF de seruus.*

Ceci étant fait, nous obtenons en sortie un premier *dataframe* à l'image de ce que nous décrivions en introduction (Figure.12). Nous y retrouvons donc une colonne listant toutes les chartes dans lesquelles se trouve le lemme *seruus*. Cette liste correspond à la fréquence booléenne d'apparition du lemme dans le corpus des *CEMA*. Les deux suivantes donnent les variables essentielles dans les calculs de fréquence et de TF-IDF dont les indices remplissent les deux dernières colonnes.

	indice	nbr_tot_mots	Occurrences	TF-IDF	Fréquence
0	0800_passau_224441	178	1	0.006344	0.005618
1	0800_pflugk-harttung_236150	469	1	0.002408	0.002132
2	0800_saintgall_224984	209	2	0.010806	0.009569
3	0800_stclaudes_53057	123	1	0.009181	0.008130
4	0801_mgh-dd-kar-i_216522	545	1	0.002072	0.001835
...	...	...	...	...	...
5279	1299_wub_59832	225	1	0.005019	0.004444
5280	9999_hersfeld_229038	1117	1	0.001011	0.000895
5281	9999_marmoutiermanceau_255656	482	1	0.002343	0.002075
5282	9999_marmoutiervendomois_256306	229	1	0.004931	0.004367
5283	9999_marmoutiervendomois_256373	215	1	0.005252	0.004651

5284 rows × 5 columns

Figure 12 – *Fusion des tableaux*

4. Affichage des données :

Pour mieux se rendre compte de la distribution de nos données, et ainsi faciliter leur appréhension, nous avons choisi de les représenter sous forme de *scatter plot* comparé, avec en abscisse, les textes comprenant le lemme, et en ordonnée, leur indice de fréquence ou leur score TF-IDF. La lecture de gauche à droite permet de retracer l'évolution diachronique de ces indices car les chartes sont triées en fonction des dates (Figure. 13). Fréquence et TF-IDF sont proportionnels, puisque le second emploie le premier comme valeur multiplicatrice. Cela n'en réduit pas la pertinence, car les proportions varient. La comparaison

des schémas pour chacun des lemmes analysés permettra d'interpréter leur rôle respectif dans le langage des actes juridiques médiévaux.

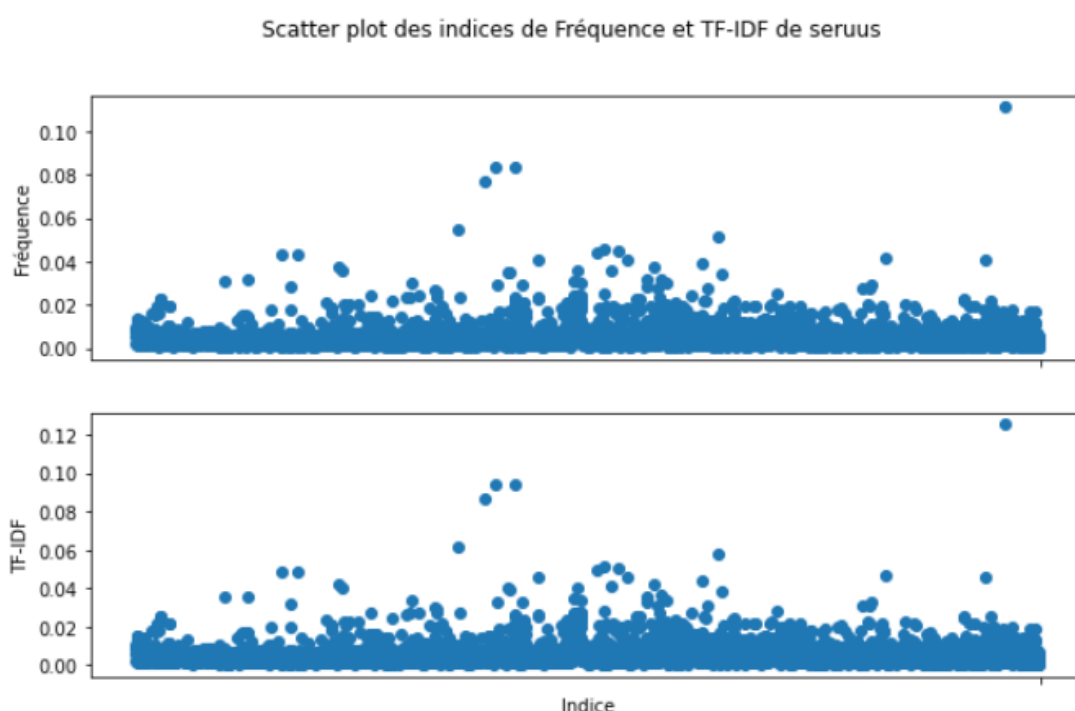


Figure 13 – Scatterplot des indices TF-IDF et de la Fréquence de *seruus*.

Pris seuls, ces tableaux n'ont pas beaucoup de sens. Ils indiquent juste dans l'absolu la dispersion à travers le temps de la fréquence et de l'indice TF-IDF de chaque mot. L'analyse devient pertinente dès lors que l'on compare les résultats entre eux. Pour faire parler les graphes, nous avons choisi de nous intéresser à une comparaison statistique de base, en mesurant les différences de moyenne et d'écart-type entre *seruus* et les autres mots.

2.3.2 Présentation des données et interprétation

— Les occurrences :

Lorsque l'on s'intéresse aux statistiques comparées du nombre d'occurrences dans les chartes pour chacun des termes latins liés à la question de la servitude, plusieurs indicateurs permettent de soulever des différences. D'abord et avant tout, il faut noter que le nombre de mots pour chaque corpus est différent, bien qu'un grand nombre (à l'exclusion de *sclauus* et *colonus*) appartient à une classe plutôt similaire. Tous les minima affichent 1, ce qui est logique lorsque l'on parle d'occurrences, tandis que les maxima ont une légère tendance à la fluctuation. Les mots les plus rares sont les moins employés dans les textes. À l'inverse, les lemmes *ancilla*, *famulus* et *serua* détiennent le plus d'apparitions par texte,

davantage que *seruus* lorsqu'il est allégé de sa formule invocatrice. La différence la plus notoire concerne la forme brute de *seruus* dont le maximum est nettement plus élevé que sa forme nettoyée. Cela signifie qu'en plus de sa présence dans l'invocation de la charte, l'expression se retrouve encore maintes fois dans son corps.

En moyenne, *seruus* est employé plus de fois par texte que ses homologues, à l'exception de *mancipium* et d'*ancilla*. Dans l'absolu, cette information ne signifie pas grand-chose, tant qu'elle n'est pas introduite relativement au nombre de mots par texte. Difficile donc de mesurer l'importance de cet indicateur sans aborder la fréquence des lemmes dans chaque texte, ce qui sera l'objet d'un prochain point de comparaison. Nous pouvons néanmoins constater que la moyenne des occurrences de *seruus* lorsqu'elle est calculée avec l'expression *seruus seruorum dei* tend à se rapprocher de 2, ce qui est tout à fait cohérent. En effet, l'expression à elle seule, composée d'une double utilisation, influence la courbe de la moyenne.

	<i>seruus</i>	<i>mancipium</i>	<i>ancilla</i>	<i>colonus</i>	<i>famulus</i>	<i>sciausus</i>	<i>serua</i>	<i>seruus_brut</i>
nombre de textes	5284.000000	3342.000000	2023.000000	248.000000	2176.000000	214.000000	2034.000000	18127.000000
moyenne	1.366957	1.402454	1.409293	1.201613	1.314798	1.317757	1.291544	1.888564
écart type	0.891998	1.017903	1.285833	0.561661	1.187940	1.075534	1.195193	0.954640
minimum	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
25%	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	2.000000
médiane	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	2.000000
75%	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	2.000000
maximum	20.000000	17.000000	28.000000	5.000000	28.000000	11.000000	25.000000	52.000000

Figure 14 – Dataframe *comparatif des occurrences par charte de chaque lemme*.

Enfin, le pointage de l'écart-type nous apprend l'étalement des valeurs autour de la moyenne, autrement dit, elle met en exergue les comportements atypiques et affiche les valeurs les plus extrêmes. Plus l'écart-type est faible, moins les valeurs s'éloignent de la moyenne. Dans notre situation, l'écart-type de *seruus* est de loin le plus faible. Cela signifie que son nombre d'occurrences par texte est plus proche de sa moyenne que pour les autres, qu'il y a moins de documents avec des valeurs extrêmes, comparativement aux autres termes. Cela pourrait signifier une nouvelle forme de récurrence due à une formalisation du mot. Si un mot est employé de manière régulière et similaire, peut-être qu'il exprime une formalité propre à la rédaction d'un acte. Que l'écart-type soit plus faible signifie que le terme est moins répété, donc potentiellement davantage structurant – éventuellement autour d'autres expressions latines – ou encore qu'il est plus institutionnalisé. *Seruus* a donc plus tendance à apparaître dans les chartes un nombre fixe de fois, à l'inverse des autres mots pour lesquels d'une charte à l'autre leur nombre d'occurrences est moins prévisible. Cela indiquerait donc un emploi différent et nuancerait l'approche sy-

nonymique.

— La taille :

La seconde composante de la formule de fréquence est le nombre de mots présents dans le document. Le nombre d'occurrences en soi n'est pas un mauvais indicateur, mais les mots spécifiques se diluent dans une masse d'autres mots. Plus il y en a, moins on peut les considérer comme véritablement impactants. Au premier coup d'oeil, les textes nettoyés dans lesquels on retrouve *seruus* sont en général plus longs que ceux débutant par la formule d'humilité. Étant donné que cette appellation est jointe au nom du pape, émissaire de la bulle, si elle est la seule forme de *seruus* du texte, elle peut faire allusion à n'importe quelle décision papale et ne référer en cela à aucun domaine spécifique. La taille de ces chartes est alors indéterminée, mais significativement moindre que lorsqu'on se déleste de l'expression. Pour le reste, à la double exception de *mancipium* et de *famulus*, tous les autres termes apparaissent dans des actes globalement plus longs.

	<i>seruus</i>	<i>mancipium</i>	<i>ancilla</i>	<i>colonus</i>	<i>famulus</i>	<i>sclaus</i>	<i>serua</i>	<i>seruus_brut</i>
nombre de textes	5284.000000	3342.000000	2023.000000	248.000000	2176.000000	214.000000	2034.000000	18127.000000
moyenne	483.261734	384.800120	531.556599	545.971774	446.493107	544.546729	537.861849	381.530479
écart type	414.543880	299.163969	474.326740	487.132595	438.684820	446.282206	457.506177	362.591337
minimum	9.000000	23.000000	25.000000	99.000000	27.000000	22.000000	23.000000	9.000000
25%	237.000000	232.000000	246.000000	267.500000	178.000000	281.500000	273.000000	152.000000
médiane	380.000000	301.000000	391.000000	411.000000	310.500000	416.000000	409.000000	265.000000
75%	572.000000	427.000000	642.500000	627.250000	538.000000	683.750000	622.500000	485.000000
maximum	3009.000000	3001.000000	3002.000000	3001.000000	3060.000000	3001.000000	3031.000000	3009.000000

Figure 15 – Dataframe *comparatif du nombre de mots par charte pour chaque lemme*.

Tout l'intérêt de l'étude de la taille des chartes réside dans l'appréhension de son écart-type. Car rappelons-le encore, plus il est faible, plus les valeurs sont proches de la moyenne, ce qui laisse à croire à l'existence d'une certaine homogénéité de la taille des chartes aux thématiques proches. Au sein du graphe de dispersion, les valeurs comprises dans l'espace d'une distance d'écart-type correspondent en moyenne à 68% des valeurs totales, et si l'on prolonge cet espace à deux écarts-types, cela intègre jusqu'à 95% des valeurs totales¹⁴⁰. La taille des textes est donc, au vu des données de notre tableau, plutôt inconstante, et cela, davantage pour certains lemmes.

¹⁴⁰ <https://support.minitab.com/fr-fr/minitab/18/help-and-how-to/statistics/basic-statistics/how-to/store-descriptive-statistics/interpret-the-statistics/interpret-the-statistics/>

— La fréquence :

De ce que l'on sait jusque-là, le terme *seruus* est en moyenne plus présent dans les textes, et il apparaît dans des textes généralement plus courts, ce qui devrait suggérer une fréquence plus élevée. C'est bien ce que l'on observe dans ce tableau (FIG), bien que les distinctions ne soient pas nettes. *Mancipium*, dont le nombre d'occurrences lui était supérieur et dont la taille moyenne inférieure, est sans surprise plus fréquent. Cet angle analytique permet de dire que *seruus* reste relativement moins fréquent qu'une partie des autres termes, qu'il est donc moins structurant du texte en lui-même.

	seruus	mancipium	ancilla	colonus	famulus	sclauus	serua	seruus_brut
nombre de textes	5284.000000	3342.000000	2023.000000	248.000000	2176.000000	214.000000	2034.000000	18127.000000
moyenne	0.004619	0.005092	0.004844	0.003422	0.005152	0.004255	0.004022	0.008646
écart type	0.005226	0.006094	0.007340	0.002642	0.005145	0.005673	0.006835	0.008278
minimum	0.000333	0.000333	0.000333	0.000333	0.000327	0.000333	0.000330	0.000333
25%	0.001988	0.002667	0.001706	0.001745	0.002188	0.001552	0.001688	0.003893
médiane	0.003165	0.003805	0.002959	0.002747	0.003831	0.002699	0.002743	0.006431
75%	0.005263	0.005405	0.005398	0.004126	0.006711	0.004651	0.004425	0.011364
maximum	0.111111	0.137931	0.108527	0.020408	0.081395	0.045455	0.159292	0.300813

Figure 16 – Dataframe *comparatif de la fréquence de chaque lemme par charte*.

Par ailleurs, il est également plus stable. Avec un écart-type plus faible que les autres, on peut considérer sa fréquence comme plus régulière. Il n'y a donc pas d'anomalies significatives, des chartes dans lesquelles il apparaîtrait un nombre disproportionné de fois. En tout cas, pas dans sa forme épurée. Car lorsque l'on se penche sur la version brute, qu'il s'agisse de sa moyenne ou de son écart-type, les valeurs augmentent drastiquement. Si les occurrences augmentent, cela peut s'expliquer par des répétitions des l'expression au sein même de la charte, dans d'autres parties que dans la titulature initiale. Cela s'exprime bien au travers de son écart-type, bien plus élevé que pour les autres termes. L'envol de cette variable peut être dû au fait qu'une grande majorité des chartes présentes ne contient *in fine* que la formule *seruus seruorum dei* en guise de présentation, par opposition à une autre masse de documents manifeste dont la concentration est plus élevée.

Le calcul de fréquence permet de mettre en avant des différences mineures en apparence, mais fondamentales sur le fond. D'abord, la suppression du bruit que représente le *seruus seruorum dei* diminue presque d'un facteur 2 la fréquence dans les textes, ce qui souligne une nouvelle fois son importance dans les chartes médiévales. Cela étant dit, une fréquence plus élevée permet de mettre en exergue un emploi plus régulier qui peut s'expliquer de deux manières. Ou bien le terme est employé de la même manière et la fréquence nous permet de faire émerger des expressions figées propres à certaines pratiques administratives, ou bien son contexte d'apparition varie, ce qui lui donne dès lors un poids

bien plus élevé.

— L'indice TF-IDF :

Une fois toutes ces variables passées en revue, nous pouvons désormais nous tourner vers l'analyse comparative des indices TF-IDF. Ce que l'on observe directement, c'est un déséquilibre flagrant entre *seruus* et tous les autres termes. En moyenne, on le retrouve près de deux fois moins que les autres, y compris lui-même dans sa version brute. Comme nous l'avons mentionné précédemment, plus l'indice d'un mot est élevé, plus son importance croît. Cette différence majeure laisse donc penser qu'il est moins structurant de la charte que les autres.

	seruus	mancipium	ancilla	colonus	famulus	sclauus	serua	seruus_brut
nombre de textes	5284.000000	3342.000000	2023.000000	248.000000	2176.000000	214.000000	2034.000000	18127.000000
moyenne	0.005216	0.009490	0.010083	0.010241	0.010561	0.013007	0.008363	0.009763
écart type	0.005901	0.011356	0.015278	0.007909	0.010547	0.017344	0.014212	0.009348
minimum	0.000376	0.000621	0.000693	0.000997	0.000670	0.001019	0.000686	0.000376
25%	0.002244	0.004969	0.003552	0.005224	0.004486	0.004745	0.003510	0.004396
médiane	0.003573	0.007090	0.006158	0.008223	0.007854	0.008251	0.005704	0.007262
75%	0.005943	0.010073	0.011236	0.012350	0.013758	0.014219	0.009200	0.012832
maximum	0.125468	0.257039	0.225904	0.061083	0.166851	0.138960	0.331198	0.339683

Figure 17 – Dataframe *comparatif de l'indice TF-IDF pour chaque lemme*.

2.4 Les indices de similarités

En dehors du calcul du TF-IDF, il existe d'autres indicateurs qui permettent d'établir les connexions entre les mots. Il est possible de mesurer leur degré de proximité au travers d'indices dits de « similarité », autrement dit, d'évaluer la distance qui les sépare sémantiquement. Il faut cependant admettre que l'analyse de cette similarité sémantique est quelque peu artificielle, en cela qu'elle emploie des outils destinés à une approche initialement syntaxique¹⁴¹. Néanmoins, tous les indices employés ont pour objectif initial de comparer des textes, ce qui n'est pas le cas de cette étude. Pour comparer les mots, nous avons choisi de sélectionner les listes des cooccurents pour chacun des lemmes latins précédemment cités. Le calcul de similarité se base donc sur des tableaux d'occurrences, tandis que le rapprochement que l'on peut établir entre les mots dépend des ressemblances entre leurs lexiques de cooccurrences.

Pour faire simple, et pour reprendre la définition qu'en donne Damon Mayaffre, une cooccurrence est « la co-présence ou présence simultanée de deux unités linguistiques

¹⁴¹ NEGRE, E., *Comparaison de textes : quelques approches...*, 2013.

(deux mots par exemple ou deux codes grammaticaux) au sein d'un même contexte linguistique (le paragraphe ou la phrase par exemple), ou encore une fenêtre arbitraire»¹⁴²

L'intérêt des cooccurrences réside dans le constat que les combinaisons de mots sont rarement aléatoires, que leur agencement reflète une certaine volonté sémantique. Stefan Evert le mentionnait déjà en 2004 :

« You shall know a word by the company it keeps! With this slogan, Firth (1957) drew attention to a fact that language scholars had intuitively known for a long time : In natural language, words are not combined randomly into phrases and sentences, constrained only by the rules of syntax. »¹⁴³.

Pour traiter du latin, et plus particulièrement du sens du mot *seruus*, étant donné le bilan nuancé des techniques traditionnelles pour le définir, il fallait une piste neuve, une nouvelle approche. Nous nous sommes alors tournés vers ses cooccurents et plus particulièrement vers des méthodes d'analyse distributionnelle. Ces dernières s'appuient sur l'hypothèse que des mots issus de milieux similaires tendent à avoir des sens similaires. Les méthodes peuvent varier en fonction du contexte considéré pour représenter le sens des mots. Les modèles à base de mots, ceux qui prennent comme point de départ l'occurrence d'un mot dans le voisinage d'un autre (cooccurrent), parviennent à mieux saisir les relations paradigmatiques et à mettre en lumière plus efficacement les liens synonymiques entre deux mots¹⁴⁴.

Grâce aux différents indices résumés dans l'excellent article d'Elsa Negre "Comparaison de textes : quelques approches", nous avons pu comparer les différents lexiques de cooccurrences extraits de notre base de données. Notre approche sera multiple. Plusieurs des indices mentionnés dans l'article seront appliqués à la situation ici étudiée, et une formule de comparaison lexicale personnelle leur sera même ajoutée. Bien que certaines librairies Python permettent d'employer ces outils naturellement (Scikit – learn, par exemple), leur application première demeurerait la comparaison de textes. Or, l'objet de base de notre étude ici est la comparaison de tableau d'occurrences à deux colonnes. À défaut de savoir comment s'approprier ces codes, nous avons préféré adapter les mathématiques de ces indices à notre situation.

La comparaison aura lieu en deux temps. Tout d'abord, nous allons nous focaliser sur les rapports de similarité que chacun des mots entretient avec *seruus*, au travers de leurs

142 MAYAFFRE, D., *De l'occurrence à l'isotopie : les co-occurrences en lexicométrie*, dans *Syntaxe et Sémantique*, n°9, 2008/1, p.55.

143 EVERT, S., *The Statistics of Word Cooccurrences : Word pairs and Collocations*, p.15.

144 HETLEN, K. et BERTELS, A., *Sémantique distributionnelle en linguistique de corpus*, dans *Langages*, n°201, 2016, p.54

listes de cooccurents, avant de nous tourner vers une comparaison plus intuitive, à savoir celle des mots à proprement parler. Pour ce faire, nous nous attaquerons à trois types grammaticaux spécifiques : les noms communs (SUB), les verbes (VBE), et les adjectifs (QLF). Afin d’avoir un recul dans l’analyse de nos données, nous avons choisi d’intégrer un mot ‘test’, un échantillon aléatoire de comparaison qui n’appartient pas au champ lexical de *seruus*. Ainsi, afin de donner du sens aux indices que nous allons produire, nous comparerons également *seruus* avec *molina*, dont le nombre d’occurrences dans le corpus équivaut globalement à celui des autres lemmes (2404).

Il faut aussi noter que nous avons été contraints de laisser de côté deux lemmes (*colonus* et *sclauus*) dont le nombre d’occurrences était trop faible, et *de facto* le poids de leurs cooccurrences bien trop déséquilibré pour rendre la comparaison pertinente. À cet égard, toutes les conclusions que nous pourrions tirer ici devront faire abstraction de leur cas.

Dans l’ordre, nous comparerons les lemmes selon les indices de Dice, l’indice de Jaccard, un calcul de similarité Cosinus, le coefficient de corrélation de Pearson et pour finir, un indice personnel. Nous laissons volontairement de côté la distance d’édition de Levenshtein, jugée inadaptée à la comparaison de nos tableaux. Pour chacun d’entre eux, nous allons procéder comme suit : d’abord, nous allons présenter l’indice et expliciter en quoi il est pertinent pour notre recherche, puis nous illustrerons les résultats sous forme de trois graphiques diachroniques comparatifs. Nous essayerons ensuite, dans la mesure du possible de commenter nos observations.

2.4.1 Indice de Dice

L’un des intérêts majeurs de la démarche provient de la sensibilité que cet indice offre vis-à-vis de données hétérogènes. Les valeurs aberrantes ne se voient pas attribuer un poids conséquent¹⁴⁵. Cela a œuvré à l’essor de sa popularité dans le champ de la recherche en lexicographie informatique, afin de « mesurer le score d’association lexicale de deux mots donnés »¹⁴⁶. Son utilité est donc on ne peut plus évidente.

¹⁴⁵ https://stringfixer.com/fr/Sorenson_index

¹⁴⁶ RYCHLY, P., *Un score d’association favorable aux lexicographes. Actes du deuxième atelier sur les progrès récents dans le traitement du langage naturel slave*, dans RASLAN, 2008, p.6-9.

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|}$$

Figure 18 – *Formule de Dice*

$$sim_{dice}(d_1, d_2) = \frac{2N_c}{N_1 + N_2}$$

Figure 19 – *Indice de similarité de Dice (E. NEGRE)*

Brièvement, voici comment le calculer. $|X|$ et $|Y|$ sont les cardinalités des deux ensembles étudiés, autrement dit le nombre total d'élément de chacun des ensembles. L'indice se décrit donc comme le produit de l'intersection entre deux ensemble par deux, divisé par la somme des cardinalités. Dans le cas de la lexicographie, on multiplie donc par deux le nombre de fois où un mot se retrouve dans les deux textes, que l'on divise ensuite par la somme du nombre de mots de chacun des textes.

```
def Dice_indice(path1,path2):

    df = pd.read_csv(path1)
    df = df.drop('Unnamed: 0', axis=1)
    df2 = pd.read_csv(path2)
    df2 = df2.drop('Unnamed: 0', axis=1)

    liste = []
    for i in range(len(df)):
        for j in range(len(df2)):
            if df['lemma'][i] == df2['lemma'][j]:
                liste.append(df['lemma'][i])

    num = 2 * len(liste)
    denum = len(df) + len(df2)

    result = num/denum

    return(result)
```

Figure 20 – *Fonction de similarité de Dice.*

Dans notre cas, le sujet d'étude concerne deux tableaux de cooccurrences où chaque lemme est accompagné par une fréquence, son poids en quelque sorte. Cette formule plutôt simple fait abstraction des fréquences et ne juge que de la proximité formelle des deux tableaux. L'indice de Dice nous éclaire donc sur le contenu à proprement parler, sur les mots

que les deux lexiques partagent. Il s'agit d'une approche semblable, quoique plus complexe, à une similarité booléenne de base, jugeant de la présence ou non d'un terme dans l'un et l'autre tableau. Pour les besoins de notre recherche, nous avons défini une fonction qui calculait automatiquement l'indice en le comparant systématiquement au terme *seruus*.

Nous avons ensuite implémenté notre fonction dans un code plus vaste qui itérait les fichiers de notre ordinateur et qui donne en *output* un diagramme comparatif des indices pour chaque catégorie grammaticale (QLF, SUB, VBE)¹⁴⁷. En voici les analyses.

1. SUB :

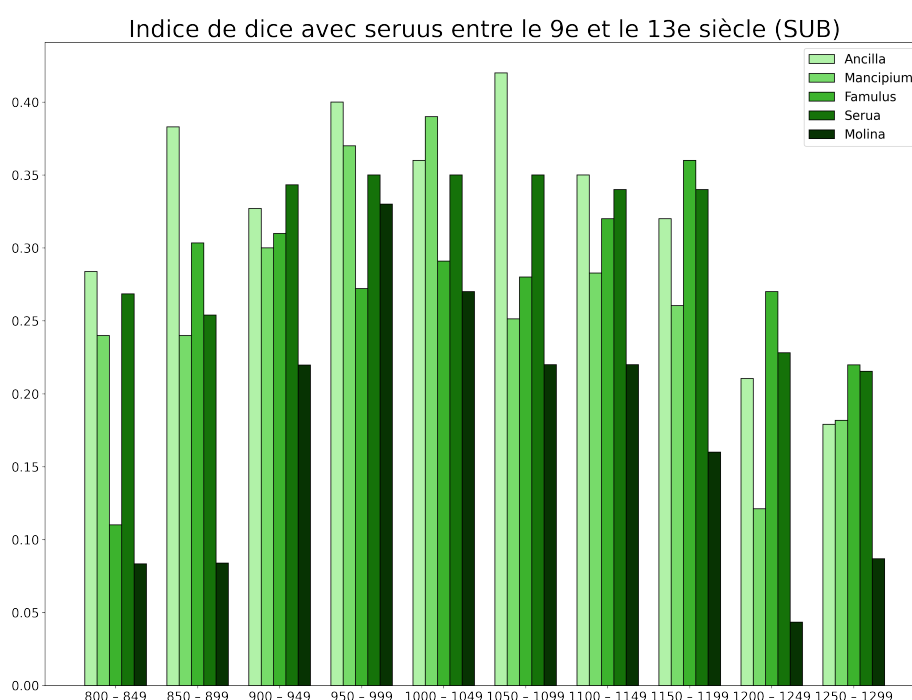


Figure 21 – *Distribution des indices de Dice par lemme entre 800 et 1300 (SUB).*

Une lecture simple du diagramme n'est pas possible dans l'absolu, car les pourcentages ne reflètent aucune vérité propre. Ce n'est que la comparaison de multiples indices qui permettra de définir la proximité entre *seruus* et les autres termes. Ce que l'on peut toutefois constater, c'est que le mot *molina*, l'échantillon de comparaison aléatoire, fait émerger une tendance en ce qui concerne les substantifs utilisés conjointement à *seruus*. Pour presque tous les siècles, sa valeur est inférieure à celle des autres termes. Cela peut donc impliquer que ces derniers sont sémantiquement plus proches de *seruus* que lui. Toutefois, il faut voir cette tendance avec recul, car elle permet certes de dissocier *molina* du

¹⁴⁷ Cfr: le github pour le détail.

champ lexical de la servitude mais sans pour autant établir le degré de ressemblance que les autres entretiennent avec *seruus*. Car *in fine*, à quoi correspond une proximité de 40% ?

1. VBE :

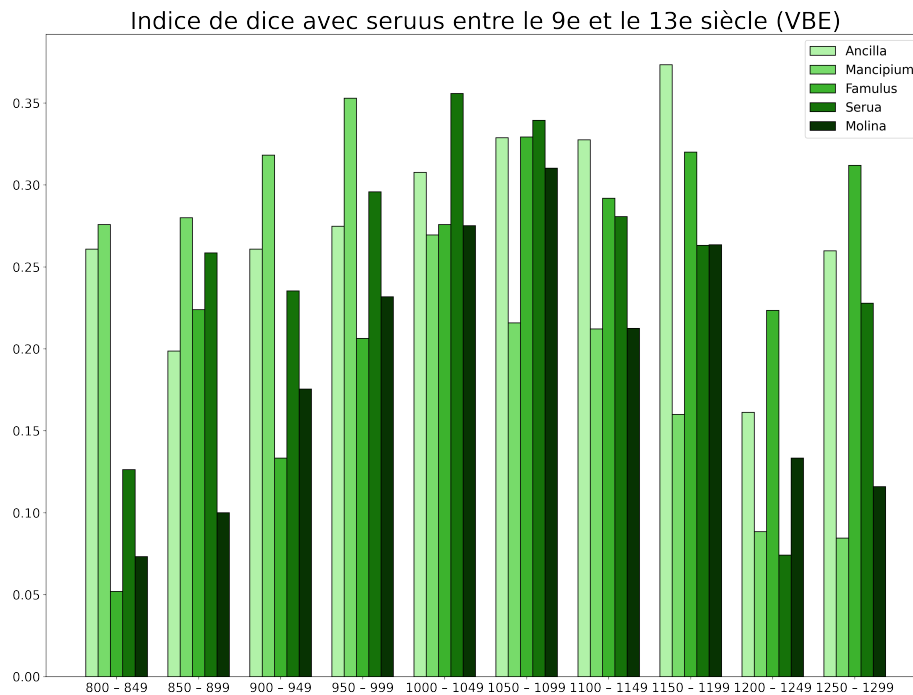


Figure 22 – Distribution des indices de Dice par lemme entre 800 et 1300 (VBE).

La tendance observée pour les noms communs ne se reflète pas d’après les comparaisons de verbes. En effet, les rapprochements entre les lexiques de cooccurrence de verbes de *seruus* et de *molina* tendent à rejoindre ceux des autres lemmes. Les différences sémantiques ne s’illustreraient, selon l’indice de Dice, pas par les actions entreprises ou subies par les *serui*. En outre, notons aussi que l’échelle de l’indice a diminué entre les deux catégories grammaticales. Les verbes sont donc plus variés dans l’usage que les noms, lorsqu’ils encadrent le mot *seruus*.

1. QLF :

En ce qui concerne les adjectifs qualificatifs, le diagramme suggère une ressemblance avec l’emploi des noms communs.

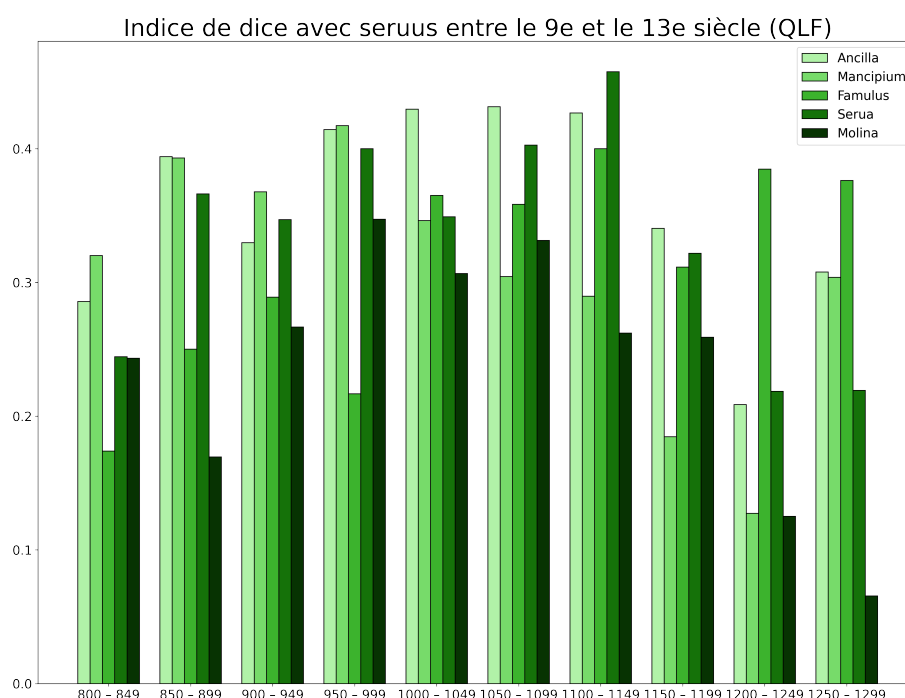


Figure 23 – Distribution des indices de Dice par lemme entre 800 et 1300 (QLF).

2.4.2 Indice de Jaccard

Il existe un second coefficient qui ressemble fortement à l'indice de Dice : l'indice Jaccard. Les deux peuvent être équivalents dans le sens où il est possible de calculer la valeur de l'indice de Jaccard à partir de l'indice de Dice, et inversement d'après la formule suivante, où S correspond à l'indice de Dice, et J à l'indice Jaccard.

$$J = S / (2 - S) \text{ et } S = 2J / (1 + J).$$

Figure 24 – Formule de dépendance entre l'indice de Jaccard et l'indice de Dice

Pour le définir plus simplement, l'indice ou coefficient Jaccard « est le rapport entre la cardinalité (la taille) de l'intersection des ensembles considérés et la cardinalité de l'union des ensembles. Il permet d'évaluer la similarité des ensembles »¹⁴⁸. La formule est également liée au calcul de l'indice cosinus. Entre leur deux formules, il n'existe qu'une unique différence : la soustraction du produit scalaire des deux ensembles au produit des normes de ces ensembles. *In fine*, il s'avère que les trois formules sont intimement intriquées.

148 NEGRE, E., *Comparaison de textes : quelques approches...*, 2013, p.9.

$$sim_{cosinus}(d_1, d_2) = \frac{\vec{d_1} \cdot \vec{d_2}}{\|\vec{d_1}\| \|\vec{d_2}\|}$$

Figure 26 – Indice de similarité cosinus (E. NEGRE)

$$sim_{jaccard}(d_1, d_2) = \frac{\vec{d_1} \cdot \vec{d_2}}{\|\vec{d_1}\| \|\vec{d_2}\| - \vec{d_1} \cdot \vec{d_2}}$$

Figure 25 – Indice de similarité de Jaccard (E. NEGRE)

Comme pour l'indice de Dice, nous avons adapté la formule pour qu'elle s'applique à nos besoins et l'avons transformée en une fonction python qui prend en entrée deux tableaux, et qui donne en sortie l'indice compris entre 0 et 1.

```
def jaccard(path1, path2):
    df = pd.read_csv(path1)
    df = df.drop('Unnamed: 0', axis=1)
    df2 = pd.read_csv(path2)
    df2 = df2.drop('Unnamed: 0', axis=1)

    for i in range(len(df2)):
        df2['Occurrences'][i] = df2['Occurrences'][i]/df2['Occurrences'].sum()
    for i in range(len(df)):
        df['Occurrences'][i] = df['Occurrences'][i]/df['Occurrences'].sum()

    num = 0
    for i in range(len(df)):
        for j in range(len(df2)):
            if df['lemma'][i] == df2['lemma'][j]:
                num = num + (df['Occurrences'][i]*df2['Occurrences'][j])

    df2['carré'] = df2['Occurrences']**2
    df['carré'] = df['Occurrences'] **2
    denom = (math.sqrt(df2['carré'].sum())) * (math.sqrt(df['carré'].sum()))
    final = num/(denom - num)

    return(final)
```

Figure 27 – Fonction indice de Jaccard

Lorsque l'on applique le même processus que pour Dice, voici les graphes qui sont générés :

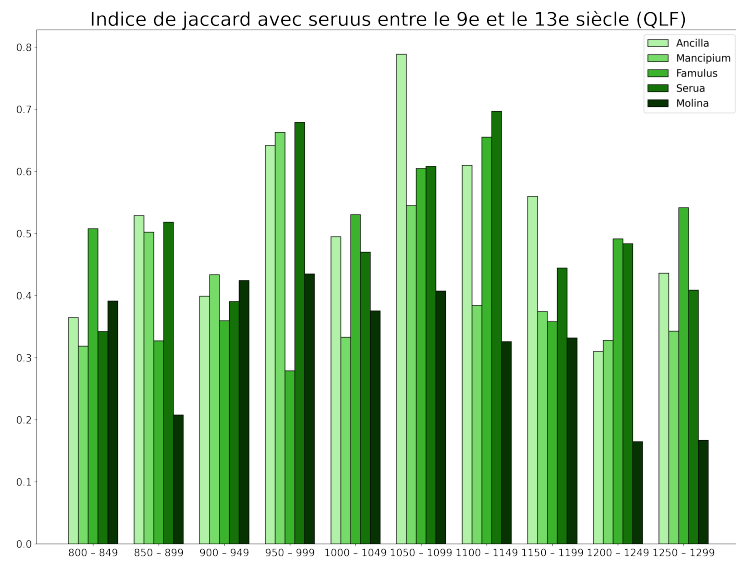


Figure 28 – *Jaccard (QLF)*

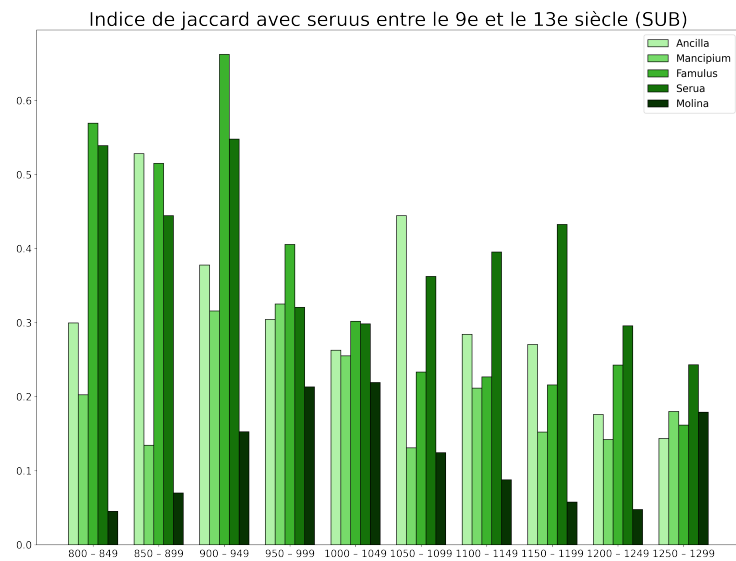


Figure 29 – *Jaccard (SUB)*

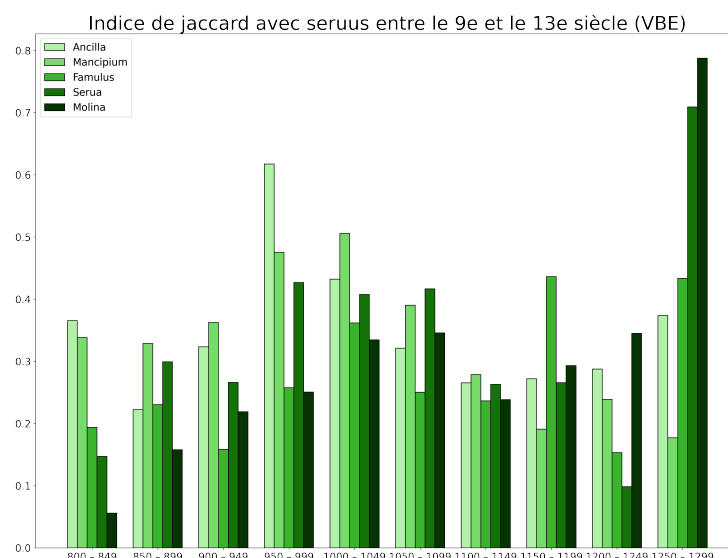


Figure 30 – *Jaccard (VBE)*

Ce que nous disent ces graphes, c’est qu’en moyenne, les mots tendent à plus de 50% de similarité avec *seruus*. Pour les QLF et les VBE, les valeurs en axe Y tendent même pour certaines vers les 80%. Comment expliquer ce phénomène ? En ce qui concerne les adjectifs, nous constatons que ces valeurs extrêmes sont marginales, et ne touchent que les deux demi-siècles qui composent le 13^e siècle. En outre, les indices les plus élevés sont ceux de *serua* et de *molina*, ce qui soulève l’interrogation. Difficile d’expliquer cette brisure soudaine avec le *pattern* constaté par l’indice de Dice, d’autant plus au vu du comportement irrégulier de *molina*. Dans l’ensemble, en ce qui concerne les noms et les adjectifs, ce dernier tient bien son rôle de modérateur et se distance très nettement des autres lemmes, si ce n’est pour ce 13^e siècle. L’explication pourrait provenir de la raréfaction des occurrences de *molina* et sa concordance avec des verbes plus simples, plus standards.

Même avec des indices parfois si élevés, il reste difficile d’interpréter les résultats. D’une part, il faut se rappeler que la quantité d’occurrences dépend de la quantité de documents, qui fluctue selon les siècles. En ce qui concerne les siècles les plus abondants en documentation, on peut noter une légère baisse sur tous les graphes. Plus il y a de mots, plus les nôtres se noient dans la masse. Dans tous les cas, les comportements de chacun des graphes exprime une idée précise. Si les ressemblances sont plus fortes avec les adjectifs, cela signifie que *mancipium*, *ancilla*, *famulus* et *serua* sont définis avec des épithètes semblables à *seruus* et sont donc décrits plus similairement. À l’inverse, la différenciation se marque avec les noms qui les jouxtent, ce qui pourrait laisser sous-entendre deux

choses. Ou bien les mots qui se trouvent en leur périphérie n'interagissent pas avec eux pour donner du sens à la phrase (peut-être par la présence d'ablatifs circonstanciels), ou bien ils sont liés mais appartiennent à un champ sémantique bien différent. Pour s'en assurer, il faudrait plonger plus en profondeur dans le texte, et à la moindre occurrence de l'un des termes, analyser son cas et ceux de ses voisins ce qui permettrait d'évaluer leur potentiel synergique.

2.4.3 Similarité Cosinus

Les métriques à base cosinus sont généralement celles qui sont le plus fréquemment utilisées dans le cadre de modèles vectoriels basiques ¹⁴⁹. Il s'agit de mesures issues de l'algèbre linéaire qui s'étendent également aux vecteurs pondérés et qui sont devenues l'une des mesures standards dans le domaine de la recherche d'informations ¹⁵⁰. Cette métrique estime la similarité entre deux vecteurs en déterminant le cosinus de leur angle, ou comme l'explique Amit Singhal :

« Typically, the angle between two vectors is used as a measure of divergence between the vectors, and cosine of the angle is used as the numeric similarity (since cosine has the nice property that it is 1.0 for identical vectors and 0.0 for orthogonal vectors) ¹⁵¹. »

La formule de la similarité cosinus est donc, comme précédemment mentionnée, très similaire à l'indice de Jaccard (Figure.26), nonobstant la soustraction au dénominateur du produit scalaire des deux ensembles. Le résultat est compris entre 0 et 1 ¹⁵².

Tout comme pour les métriques précédentes, nous avons traduit cette formule en langage Python et nous l'avons adaptée à nos tableaux d'input.

149 RAJMAN, M. et LEBART, L., *Similarités pour données textuelles*, p.5.

150 BANNOUR, S., AUDIBERT, L. et NAZARENKO, A., *Mesures de similarité distributionnelle entre termes*, 2011, p.528.

151 SINGHAL, A., *Modern Information : A brief Overview*, p.2.

152 NEGRE, E., *Comparaison de textes : quelques approches...*, 2013, p.7-8.

```
def cosine_sim(path1,path2):

    df = pd.read_csv(path1)
    df = df.drop('Unnamed: 0', axis=1)
    df2 = pd.read_csv(path2)
    df2 = df2.drop('Unnamed: 0', axis=1)

    for i in range(len(df2)):
        df2['Occurrences'][i] = df2['Occurrences'][i]/df2['Occurrences'].sum()
    for i in range(len(df)):
        df['Occurrences'][i] = df['Occurrences'][i]/df['Occurrences'].sum()

    num = 0
    for i in range(len(df)):
        for j in range(len(df2)):
            if df['lemma'][i] == df2['lemma'][j]:
                num = num + (df['Occurrences'][i]*df2['Occurrences'][j])

    df2['carré'] = df2['Occurrences']**2
    df['carré'] = df['Occurrences'] **2
    denom = (math.sqrt(df2['carré'].sum())) * (math.sqrt(df['carré'].sum()))
    final = num/denom

    return(final)
```

Figure 31 – *Fonction de similarité cosinus.*

De prime abord, ce que l'on constate, c'est que globalement, les échelles sont réduites. Pour chaque catégorie grammaticale, l'axe des Y ne dépasse pas les 0,4 points de similarité. Cela signifie donc que chacun des termes tend plus vers l'orthogonalité, autrement dit : l'indépendance sémantique. Nous retrouvons les mêmes courbes que pour l'indice précédent, accompagnées des mêmes tendances extraordinaires et difficile à décrire¹⁵³. Parmi tous les termes, *molina* reste globalement celui qui se corrèle le moins avec *seruus*.

2.4.4 Corrélation de Pearson

La corrélation de Pearson donne une indication, non pas directement de la similarité, mais davantage une indication de la linéarité ou non de la corrélation entre deux variables. La seule limite à cette observation est que la corrélation de Pearson traite des relations linéaires, ce qui exclut donc les données différemment corrélées (exponentielle, logarithmique, non-linéaire)¹⁵⁴.

Une nouvelle fois, nous constatons que la formule est liée aux autres. En effet, pour calculer le coefficient de corrélation de Pearson, il suffit de reprendre la similarité cosinus dans laquelle il faut substituer à la valeur initiale la soustraction de la moyenne à cette valeur.

¹⁵³ Cfr: github pour les graphes.

¹⁵⁴ [https://link.springer.com/chapter/10.1007/978-3-642-00296-0_5?](https://link.springer.com/chapter/10.1007/978-3-642-00296-0_5?noAccess=true)
noAccess=true

$$sim_{pearson}(d_1, d_2) = sim_{cosinus}(d_1 - \overline{d_1}, d_2 - \overline{d_2})$$

Figure 32 – *Corrélation de Pearson (E. NEGRE)*

Si l'indice permet d'estimer le degré de corrélation linéaire entre deux mots, il pourrait également être un indicateur de synonymie et d'antonymie. S'il se rapproche de 1, c'est que leur emploi est hautement similaire, et *de facto* qu'ils sont potentiellement synonymes, et s'il tend vers -1, c'est qu'ils sont inversement corrélés, autrement dit, qu'ils sont potentiellement antonymes. Autour de la valeur de 0, les vecteurs se dissocient et leur corrélation se dissipe.

L'application du coefficient de corrélation de Pearson à nos données retourne des graphiques très divers¹⁵⁵. Jusqu'alors, c'est le seul qui attribue des valeurs négatives. Elles sont multiples dans chacun des trois graphes, mais ne descendent pas au-delà de la barre des -0,3. Nous pouvons donc estimer que les mots conservent une certaine indépendance dans l'absolu, vis-à-vis de *seruus*. L'axe des Y atteint également cette valeur de 0,3, pour les QLF et le SUB. Quant au graphe des VBE, il rapporte la même anomalie que les autres indices pour la seconde moitié du 13^e siècle, à savoir une valeur disproportionnée, démesurée pour le mot *molina*.

Le graphique relatif aux verbes est le terrain de plusieurs interprétations. Parmi tous les termes, il y en a deux qui restent quasi intégralement dans le positif, à savoir *ancilla* et *mancipium*. Il n'y a que lors de la seconde partie du 13^e siècle que cela s'inverse, et plutôt drastiquement. Peut-être est-ce lié au coefficient de *molina*, tout aussi irrégulier. Est-ce là un problème d'échantillonnage ? De codage ? Cette exception mise à part, les deux termes ne détiennent *in fine* qu'un indice faible qui signale davantage une dissimilarité syntaxique. Néanmoins, il reste possible que la relation qu'entretient *seruus* avec eux ne soit pas linéaire, comme précédemment mentionné.

Les deux autres graphes ne sont pas plus éloquents, en cela qu'ils comprennent tous les deux la majorité des valeurs dans l'intervalle [-0,1 ; 0,3]. Il ne s'agit pas non plus d'indices extraordinaires.

2.4.5 Fonction de similarité personnelle

Pour terminer, lors de l'examen de Python du premier quadrimestre à l'École Nationale des Chartes, il nous avait été demandé de réaliser une fonction qui prenait en entrée deux

¹⁵⁵ Cfr: github pour les graphiques.

textes et qui renvoyait un indice de similarité entre les deux. Quelques adaptations ont été nécessaires, mais elle est en état de produire des résultats. En voici l'idée sous-jacente.

```
def indice_similarite(path1, path2):

    df = pd.read_csv(path1)
    df = df.drop('Unnamed: 0', axis=1)
    df = df[['lemma', 'Occurrences']]
    df2 = pd.read_csv(path2)
    df2 = df2.drop('Unnamed: 0', axis=1)
    df2 = df2[['lemma', 'Occurrences']]

    dic1 = dict(df.values)
    dic1 = {k: v for k, v in dic1.items()}
    dic2 = dict(df2.values)
    dic2 = {k: v for k, v in dic2.items()}

    for mot in dic1:
        if mot not in dic2:
            dic2[mot] = 0
    for mot in dic2:
        if mot not in dic1:
            dic1[mot] = 0
    differences_occurrences = 0
    for mot in dic1:
        differences_occurrences += abs(dic1[mot] - dic2[mot])
    min_somme = 0
    max_somme = sum(list(dic1.values())) + sum(list(dic2.values()))
    differences_occurrences /= max_somme

    return 1 - differences_occurrences
```

Figure 33 – Comparaison de lexiques de cooccurrences

La fonction prend en entrée deux *dataframes* et les transforme en deux dictionnaires. La clé d'accès du dictionnaire correspond au lemme, et la valeur qui lui est attribuée au nombre d'occurrences. L'idée derrière la fonction est de comparer les deux dictionnaires, et pour chaque valeur présente dans l'un mais pas dans l'autre, d'ajouter cette valeur manquante avec 0 pour valeur. Puis, de calculer par lemme la différence d'occurrences entre les deux en valeur absolue, pour ne garder que des valeurs positives. Dans un second temps, elle calcule ensuite la somme de toutes ces valeurs et établit un dénominateur du nombre total d'occurrences relatives.

Cette fonction ne poursuit pas un but de comparaison sémantique, elle ne prend même pas en compte les lemmes à proprement parler, juste leur pondération. En revanche, elle permet d'estimer d'une certaine manière, le degré de superposition de deux dictionnaires en fonction des fréquences d'apparition de lemmes. En cela, il était pertinent de la proposer.

Ce qu'on observe grâce aux graphiques, c'est une nouvelle fois une évolution des superpositions des lexiques de cooccurents dès lors que la documentation est à la hausse, et dans le même ordre d'idées, une baisse de similarité lorsqu'elle diminue. Les adjectifs et les noms écartent plus promptement *molina* que le graphique des verbes, une nouvelle fois. Cela pourrait signifier que les verbes qui permettent de décrire *molina* et *seruus* sont relativement similaires.

2.4.6 Conclusion

Tous ces indices ne servent donc *in fine* que d'indicateurs. Ils ne permettent pas d'établir dans l'absolu le degré de similarité entre deux mots, juste à pointer lequel, parmi notre liste initiale, détient le lexique de cooccurents qui se rapproche le plus de celui de *seruus*. Une fois les graphiques générés, il reste à « compter les points », et estimer quel lemme lui ressemble plus. Pour ce faire, notre choix s'est tourné vers une méthode d'évaluation basique : l'addition de chacun des indices par demi-siècle et par catégorie grammaticale pour un mot donné (Figure. 34). Étant donné que la valeur 1 de l'axe Y signifiait à chaque fois la similarité, il n'est pas incohérent de postuler que plus le résultat de la somme est important, plus il se rapproche de *seruus*.

Sans surprise, on aperçoit que *molina* se distingue avec un score nettement inférieur aux autres lemmes, et ce, malgré les résultats extraordinaires qu'il avait obtenu pour certains indices pour la seconde moitié du 13^e siècle. Comme le mystère de cette aberration n'a pas été élucidé, il serait logique que le score final doive être amoindri, creusant ainsi d'autant plus le fossé entre lui et les autres lemmes. Cela nous indique une certaine tendance sémantique. *Ancilla*, *mancipium*, *serua* et *famulus* appartiennent donc bien à un champ sémantique plus similaire à *seruus* puisqu'une partie de leurs cooccurents gravitent autour des siens.

En outre, on remarque également une nette différence entre *mancipium* et les autres lemmes, un écart flagrant. Il n'arrive qu'en quatrième position, derrière *famulus*, surprenamment, lui-même en retrait par rapport à *serua* et *ancilla*. Le rapprochement morphologique de *serua* avec *seruus* lui prodigue un léger avantage, de quelques centièmes seulement. De cela, nous pouvons supposer que le score de *serua* provient d'une potentielle erreur d'encodage dans les CEMA, qui lui aurait attribué toutes les formes du datif et ablatif pluriel de *seruus*. Cela l'aurait ainsi davantage apparenté à *seruus*, toute proportion gardée. L'expression *seruus et ancilla*, présente 242 fois entre le 9^e et le 13^e siècle biaise aussi les données, en cela que le rapprochement entre les deux lemmes est accentué par celle-ci.

Si l'on se penche sur les graphiques analytiques de la somme des indices pour chaque catégorie grammaticale, on se rend compte que l'ordre ne varie sensiblement pas, sauf pour les verbes où *mancipium* prend la seconde valeur la plus élevée. C'est toutefois aussi pour les verbes que les indices sont les plus faibles, il faut donc rester prudent.

Tout cela étant dit, que peut nous apprendre une comparaison sémantique par des tableaux de cooccurrences nous apprend à propos du rapprochement entre deux mots ? La théorie suit l'intuition initiale, les cooccurents peuvent être de bons outils de comparaison sémantique. Un mot placé dans un environnement similaire à un autre, aux cooccurents semblables à un autre se rapproche de lui sémantiquement. La multiplicité des indices de similarité permet de s'attaquer à la question selon différents angles et donner différentes perspectives analytiques. Dans notre cas, les résultats obtenus restent toutefois à nuancer.

De manière synthétique, les constats majeurs portent sur la valeur des maxima en axe Y et de la place qu'occupe *molina* dans les graphes. Pour le premier, les intervalles peuvent varier fortement selon les indices, mais en dehors de celui de Jaccard qui donnait des valeurs démesurées atteignant jusqu'aux 0,8 points de similarité, la plupart semblent plus fébriles, et peinent à dépasser la barre des 0,3. Ces chiffres ne signifient rien en eux-mêmes, ils permettent juste d'établir une comparaison des champs sémantiques.

Aucune conclusion définitive n'est donc permise. Néanmoins, quelques hypothèses peuvent toujours être soulevées. Par exemple, au vu des résultats, il est légitime de considérer *ancilla*, *mancipium*, *serua* et *famulus* comme davantage proches de *seruus* que *molina*, tout en gardant à l'esprit que les indices n'étaient pas non plus en faveur d'une ressemblance significative. Autrement dit, il reste tout à fait plausible d'imaginer qu'il existe toujours un écart sémantique entre eux.

Si l'on compare les mots, à aucun moment *seruus* n'est le plus compatible de la liste tout en restant malgré tout l'un des plus proches. *Ancilla* est *in fine* plus similaire de *mancipium* ou de *serua* qu'il ne l'est de *seruus*. L'échelle de comparaison varie entre 0 et 150. Affirmer une ressemblance sémantique lorsque le score le plus élevé n'atteint même pas les 50 ne serait pas judicieux.



Figure 34 – *Comparaison des lemmes en fonction des indices.*

In fine, bien que prometteuse, la démarche reste trop en surface pour permettre une analyse pertinente et percutante. D’abord, les échantillons ne sont pas bien équilibrés, à cause de la disparité des occurrences pour chaque terme. Tous les siècles ne doivent peut-être pas être pris en considération pour travailler sur la question, car tous ne possèdent pas un lexique de cooccurrence d’une envergure suffisante et comparable. C’est d’ailleurs entre autres pour cette raison que les termes *sclauus* et *colonus* n’ont pas été incorporés dans les graphes de comparaison. La fluctuation irrégulière des données abîme la qualité des résultats. Les informations que l’on peut extraire de ce travail ne peuvent faire office que de tendances et doivent toujours être nuancées.

Par ailleurs, travailler sur un corpus qui brasse uniquement les siècles, c’est laisser délibérément de côté les pluralités sémantiques caractéristiques à des régions bien définies. Notre démarche nécessiterait un prolongement qui prendrait en considération les différences typologiques, géographiques et temporelles comme les *CEMA* le proposent, une tâche bien trop laborieuse pour un travail de Master.

3 Définition du terme *seruus*

La comparaison avec d'autres mots issus du même champ lexical aide donc à évaluer la distance entre eux et *seruus*. Mais, bien que cette étape soit fondamentale pour l'isoler, elle ne permet pas d'effectuer une analyse qualitative, et de cerner, à défaut de la définition, au moins le sens de l'environnement dans lequel il apparaît. C'est à cela qu'est dédié le chapitre qui suit.

Avant toute chose, il est important d'estimer la fluctuation sémantique de *seruus* entre le 9^e et le 13^e siècle. Pour ce faire, il est possible d'utiliser les indices précédemment développés grâce auxquels des tendances évolutives entre les lexiques de cooccurrences s'observent. Il n'y a pas de stabilité absolue entre les demi-siècles, même si avec du recul on constate que les écarts-types ne sont pas drastiquement élevés. Qu'il s'agisse des adjectifs, des noms ou des verbes, les indices fluctuent toujours entre 0.3 et 0.6, à l'exception de l'indice de corrélation de Pearson qui stagne majoritairement autour des 0.2. Pour rappel, plus il tend vers 0, plus la similarité est haute.

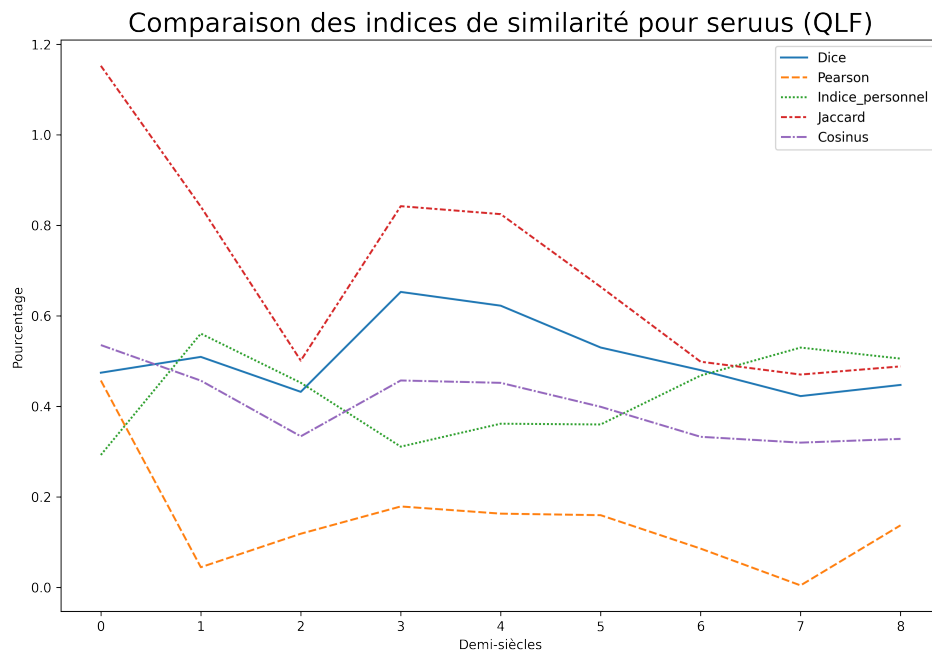


Figure 35 – *Comparaison des indices de similarités pour seruus (QLF).*

Grâce à ces graphes (Figure.35, Figure.36 et Figure.37), il est possible de déduire que *seruus* connaît une évolution plutôt régulière. Son sens varie entre les siècles, c'est certain. Aucun indice n'atteint jamais la barre des 1, ce qui révèle que le lexique se renouvelle. Mais il le fait de façon régulière, sans véritable pic de croissance ou de décroissance. Un commentaire que l'on peut émettre relève de l'indice de Jaccard qui, pour une raison que l'on

ignore, dépasse la barre des 1 pour les adjectifs du 9^e siècle, ce qui est mathématiquement impossible. Nous sommes dans l'incapacité d'expliquer ce phénomène.

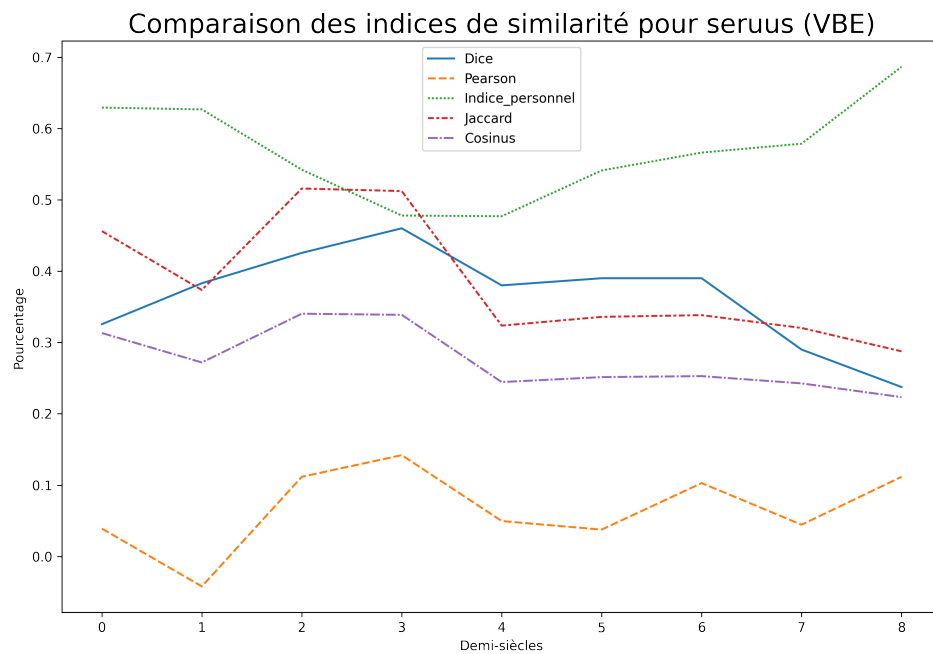


Figure 36 – *Comparaison des indices de similarités pour seruus (VBE).*

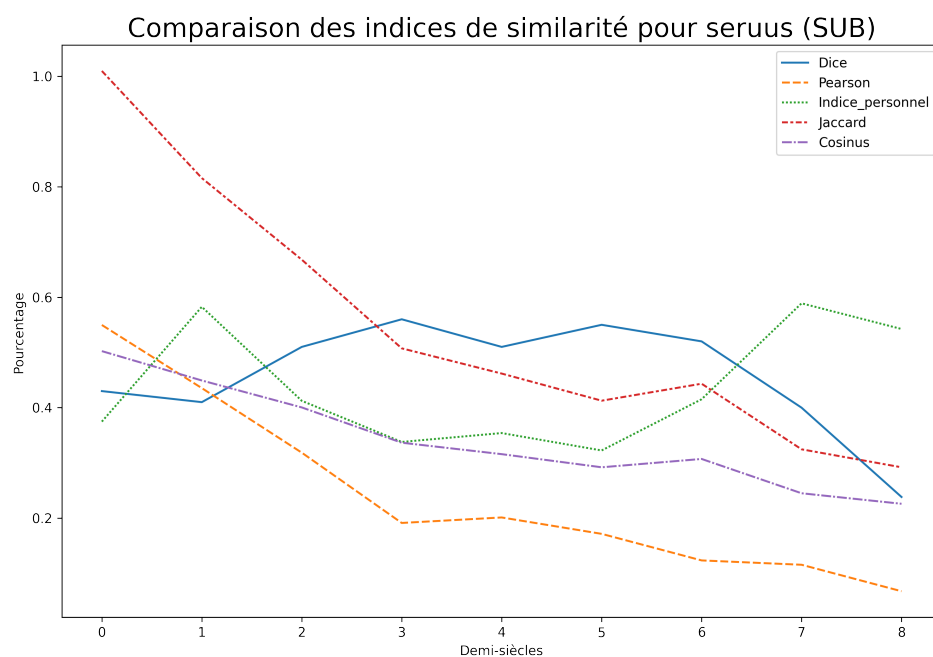


Figure 37 – *Comparaison des indices de similarités pour seruus (SUB).*

Le fait que les indices fluctuent indique que les mots qui encadrent *seruus* changent aussi. Et si le contexte diffère, peut-être en est-il de même pour son sens. Pour s'en assurer, il est nécessaire d'aborder le problème par un prisme qualitatif et de s'attarder sur ces mots qui l'accompagnent. À cette fin, il faut prendre en compte plusieurs éléments. Tout d'abord, il convient de différencier les comportements selon les catégories grammaticales. En effet, les interactions que chacune entretient avec notre lemme sont spécifiques. Un verbe ne le définit pas de la même manière qu'un autre nom ou encore qu'un adjectif, c'est pourquoi cette partie sera scindée en trois axes succincts, un pour les adjectifs (QLF), un second pour les verbes (VBE) et un dernier pour les noms communs (SUB).

Une deuxième question méthodologique s'est posée ensuite. L'exportation des données depuis les *CEMA* compliquait la réalisation d'une étude particulariste mais favorisait les analyses plus globalistes. Il était difficile de régionaliser les informations de manière équivalente car, parmi les délimitations géographiques envisageables, seule la distinction nationale était abordable, en cela qu'elle ne comptait que 18 pays, nombre dérisoire face aux 393 diocèses ou aux 125 régions. Néanmoins, cette répartition nationale est fortement limitée aussi, par son anachronisme d'abord car ce sont bien les frontières actuelles qui sont prises en compte, mais aussi et surtout car les documents sont très inégalement répartis entre les pays. En outre, il faut aussi signaler qu'une partie non-négligeable des documents médiévaux ne sont pas géo-référencés, ce qui en réduit d'autant plus la masse documentaire exploitable. Malgré tout, afin d'éprouver les méthodes d'analyses, il a été décidé de mener l'attaque sur les deux fronts. Une première partie sera consacrée à l'intégralité des *CEMA* afin de créer un modèle avec lequel il sera possible, dans un second temps, de comparer les résultats de certains pays.

Puis, qu'il s'agisse des *CEMA* dans leur ensemble ou des comparaisons entre États, il sera dans un premier temps question de définir *seruus* par ses cooccurents proches, dans un intervalle arbitraire et défini, puis de tronquer ces listes de cooccurrences en fonction du cas auquel il est accordé.

3.1 QLF

Syntaxiquement, le nom est défini par l'article et l'adjectif. Ces derniers entretiennent une relation de détermination vis-à-vis du nom qu'ils accompagnent, ce qui est encore plus le cas pour les adjectifs qualificatifs qui s'apparentent aux noms pour une partie de leur rôle syntaxique¹⁵⁶. Il paraît donc compréhensible, voire même obligatoire, d'étudier cette catégorie grammaticale.

¹⁵⁶ DELESALLE, S., *De l'adjectif : lecture d'une grammaire*, dans *Documentation et recherche en linguistique allemande contemporain - Vincennes*, n°41, p.180.

3.1.1 Étude générale dans les *CEMA*

— Méthodologie

Une première démarche, plus grossière et plus générale, se penche sur les tableaux de cooccurrences de *seruus* dans un intervalle de 10 lemmes, 5 qui le précèdent et autant qui lui succèdent. Un tel intervalle se justifie par la distance que les adjectifs peuvent prendre en latin avec le nom auquel ils sont accordés. Parmi ce tableau, il a été décidé d'en sélectionner certains afin d'alléger la suite de la recherche. Ainsi, seuls les lemmes dont la fréquence atteint les dix pourcents de la fréquence du lemme le plus fréquent feront partie de l'échantillon de comparaison. Néanmoins, dans le cas où les différences de fréquence seraient trop importantes, il a été décidé de prolonger la liste jusqu'à un maximum de 10 lemmes. Le choix du pourcentage s'inspire de la loi de Zipf, propre aux statistiques textuelles et qui lie l'importance d'un mot à sa fréquence. L'apparition des mot suit une distribution linéaire, à savoir que le plus fréquent apparaît généralement dix fois plus que le dixième, cent fois plus que le centième¹⁵⁷. Par ce fait, la plus grande partie des mots du textes sera analysée, même si cela sous-entend de laisser tomber une partie de sa diversité.

Cette discrimination écarte tant des mots aux occurrences multiples que de simples hapax. Certains linguistes sont partisans de l'idée qu'il faut faire abstraction de ces derniers, d'autres, à l'instar d'Alain Guerreau les assimilent même à de la « cacographies »¹⁵⁸. Pour Saussure, « chaque occurrence est un hapax dont la valeur diffère avec le contexte, l'intonation, etc »¹⁵⁹. Chaque mot, chaque occurrence d'un même mot, revêt donc à ses yeux une égale importance. Dans l'absolu, seuls les lemmes les plus fréquents devaient être considérés, mais par souci d'objectivité, 5 hapax seront à chaque fois examinés conjointement à eux.

Une fois les listes dressées, chacun des mots retenus sera défini doublement. D'abord en suivant la définition du dictionnaire moderne Blaise Patristic, mais aussi au moyen de son homologue médiéval (Blaise Medieval). Une fois la tâche effectuée, les adjectifs seront regroupés selon certaines catégories, à savoir : les adjectifs qui connotent une certaine appartenance ou vertu religieuse, les adjectifs renvoyant à une condition sociale ou un état de liberté, les déterminants (possessifs, indéfinis,...), les adjectifs qui renvoient à des formules textuelles (susdit, susmentionné,...) et les adjectifs qualificatifs qui ne rentrent pas dans ces catégories.

157 BULLY, P., *Zipf, créateur de la linguistique statistique*, dans *Communication et langages*, n°2, 1969, p.24.

158 GUERREAU, A., *L'avenir d'un passé incertain*, p.104.

159 MAYAFFRE, D., *De l'occurrence à l'isotopie : les co-occurrences en lexicométrie*, dans *Syntaxe et Sémantique*, n°9, 2008/1, p.55.

Le deuxième type d'analyse repose encore davantage sur les outils informatiques. L'idée est d'exporter depuis les *CEMA* les tableaux avec les formes non-lemmatisées et d'évaluer les cas de *seruus* dans leur contexte d'apparition. Toujours dans le même intervalle arbitraire, les adjectifs seront examinés selon leur terminaison et si elle semble associable, accordable, avec *seruus*, ils seront enregistrés, ce qui permet d'élaguer la première liste de cooccurrence. Cette initiative n'est qu'expérimentale, car autant l'entreprise d'analyse de *seruus* est simple, autant les distinctions entre les terminaisons d'adjectifs de première et de deuxième classe ainsi que la pluralité des cas pour une terminaison donnée rend la tâche trop ardue. Néanmoins, elle liste les possibilités d'accords entre *seruus* et les autres adjectifs, ce qui élimine déjà tous ceux qui ne peuvent véritablement le qualifier.

— Analyse quantitative

L'étude sémantique par cooccurents apporte une connaissance de l'environnement, et donc indirectement permet de comprendre à quel moment *seruus* était employé. Après avoir extrait les listes d'adjectifs cooccurents demi-siècles par demi-siècles, et après les avoir distribués dans plusieurs catégories distinctes, la comparaison devient possible.

La représentation de ces données est capitale pour percevoir tant l'évolution diachronique des catégories susmentionnées que pour leur dispersion intra-classe. Même si la création des catégories paraît purement arbitraire, elle repose partiellement sur la problématique générale de ce mémoire et sur l'intuition que le mot *seruus* ne peut intégralement se dissocier du contexte ecclésiastique dans lequel il a évolué. En ce sens, la distinction entre les adjectifs relatifs à des valeurs chrétiennes (*iustus, uenerabilis, augustus, sanctus, diuinus, ecclesiasticus,...*) et d'autres plus tournés vers le monde laïc ou juridique (*ingenuus, liber, temerarius, proprius, regalis,...*), prend du sens. Au travers de l'observation des lexiques de cooccurrences, on constate également que les déterminants possessifs monopolisent une grande partie des occurrences, il convenait donc d'ajouter une catégorie à part. Une quatrième distinction a été opérée en ce qui concerne les adjectifs au style plus « administratif »¹⁶⁰ (*iamdictus, prenomatus, predictus, prefatus,...*). Enfin, la dernière catégorie rassemble les adjectifs soit qui peuvent s'accoler aussi bien aux deux premières (*innocens, exterus, masculinus,...*) soit qui ne peuvent appartenir à aucune autre (*alter, omnis, alius, quisque,...*).

¹⁶⁰ Par là, nous entendons plutôt un vocabulaire diplomatique courant qui relève de l'expression courante, ordinaire, stylistique, une manière de dire les choses de manière simple et concise.

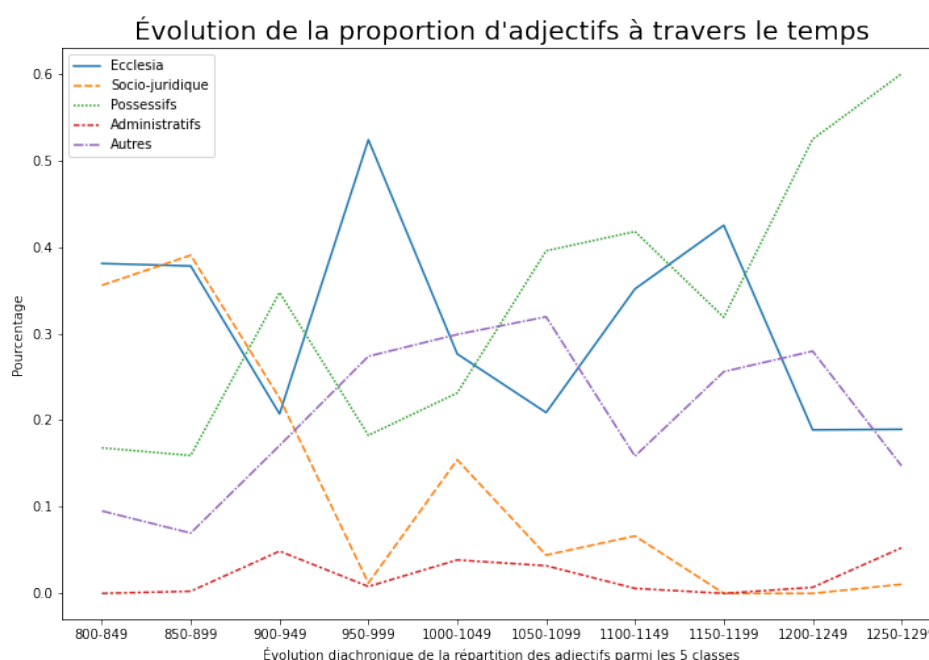


Figure 38 – *Diagramme comparatif des types d'adjectifs entre le 9^e et le 13^e siècle.*

Avant toute chose, on observe une certaine constance en ce qui concerne les adjectifs administratifs. Ils ne prennent jamais un rôle prépondérant, et s'effacent même comparativement aux autres adjectifs. Il n'en est pas de même pour les autres catégories dont les ratios oscillent et varient plus frénétiquement. Prenons l'exemple des adjectifs de la première catégorie, représentée en bleu. Au fur et à mesure du temps, la courbe évolue selon un *pattern* très irrégulier, tantôt gagnant en importance, tantôt l'inverse. *In fine*, même si ces adjectifs sont globalement en déclin, ils conservent une importance relative non-négligeable dans le quota de la fin du 13^e siècle. Ce déclin est d'autant plus marqué en ce qui concerne les adjectifs de la seconde classe. Presque majoritaires au cours de la seconde moitié du 9^e siècle, la chute est très prononcée au cours du siècle suivant, et malgré un regain mineur d'importance au cours de la première moitié du 11^e siècle, le constat est flagrant : sa concentration est quasi-nulle à la fin du 13^e siècle, plus minime encore que pour les adjectifs administratifs. Les adjectifs inclassables, quant à eux, varient peu. Bien sûr, ils gagnent de l'importance dans les premiers temps, mais finissent toutefois par décliner. Dès le début du 10^e siècle, les déterminants possessifs commencent à être de plus en plus employés. La courbe de progression est constante, si ce n'est quelques légères régressions. Mais cela ne perturbe pas l'ascension de la courbe qui, à la fin du 13^e siècle, atteint la proportion record des 60% d'occurrences.

L'idéal aurait été de s'attarder sur chacune de ces catégories et de voir comment elles évoluent au sein des différents types d'actes diplomatiques (chartes, bulles, lettres, no-

tices,...), ce qui n'a malheureusement pas été réalisé par manque de temps.

Dans les faits, il n'est possible de relier aucun de ces adjectifs directement à l'emploi de *seruus*. Néanmoins, leur proximité soulève plusieurs interrogations. Que des possessifs soient employés conjointement à *seruus* peut suggérer une certaine relation d'appartenance. En fonction du déterminant, la tournure de l'emploi peut prendre une autre teinte. Dans la mesure où *seruus* peut exprimer le service, s'il est accompagné d'un déterminant possessif, cela pourrait également relever de la formule d'humilité.

- Dans le cas de *tuus*, les exemples renvoient à l'humilité.¹⁶¹ :
- À la première personne du singulier (*meus*), le déterminant est empreint d'un sens d'appartenance et est souvent lié à des verbes comme *donare*¹⁶²
- Ces schémas se répètent pour le pluriel. Dans le cas de *noster*, on retrouve la même idée d'appartenance et de possible transmission¹⁶³, tandis qu'à la seconde émane la même intention d'humilité¹⁶⁴.
- En ce qui concerne la forme à la troisième personne du singulier, le *suus*, il s'aligne sur le sens de la première personne et, bien que plus ambigu quand même, donne la même impression de privation de liberté, de rattachement¹⁶⁵.

On observe donc une légère altération sémantique en fonction du possessif employé. Le ton de la phrase n'est pas le même dès lors qu'il s'agit de la première ou de la deuxième personne (singulier ou pluriel). À ce stade, seule la différenciation est explicitée, pas le poids de chacun d'entre eux. Si l'on se penche sur leur proportions, le constat est on ne peut plus limpide. Les formules de politesse sont très minoritaires, voire quasiment absentes des 10 lemmes les plus fréquents. Dans l'environnement de *seruus* se retrouvent donc de manière prépondérante des déterminants possessifs illustrant l'appartenance ou en tout cas une forme de liaison sociale¹⁶⁶ Et parmi tous les possessifs, les plus fréquents sont dès le début du 9^e siècle *noster*, *meus* et *tuus*. Ce n'est véritablement qu'à partir du 11^e siècle que les formes à la deuxième personne apparaissent progressivement, bien que toujours minoritairement¹⁶⁷.

161 doc.id = 1051_plpape_239271 : « Beate Petre apostole ego Leo episcopus seruus tuus et omnium seruorum dei de tui ».

162 doc.id = 1212_bari-snicolas_257600 : « ut Andreas seruus meus quem eidem sire Guillelmo in dotem dedi ».

163 doc.id = 0949_cluny_2160 : « Et dono uobis seruos nostros is nominibus Dalmacius et uxor sua nomine Dominica et infantibus is nominibus Ingelbergi et Uuandalbergi et Ioanni ».

164 doc.id = 1074_sahagun_254320 : « Ego igitur exiguus seruus uester Iustus presbiter adhesit in anima mea ».

165 doc.id = 1166_cluny_5636 : « Preterea homines suos liberos ac seruos suos nemo sine ipsius permissione distringat »

166 Impossible de faire correspondre les teintes des graphes, le seul indicateur qui permet de connecter le bâtonnet au graphique en fromage, c'est la légende au-dessus du bâtonnet.

167 Cfr: github pour le détail des graphes.

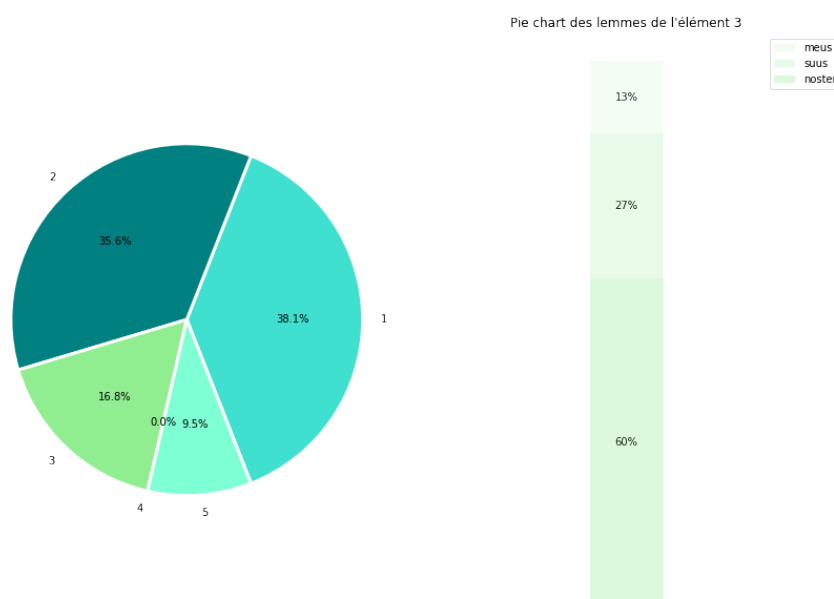


Figure 39 – *Pie charts en bâtonnets des différentes catégories d'adjectifs.*

L'environnement adjectival de *seruus* prête donc à penser que les épithètes propres à l'*ecclesia* sont initialement les plus abondants. Ils sont aussi ceux qui donnent le plus d'indications sur l'appartenance à un champ lexical donné, ce qui se confirme par la disparition des lemmes qui sous-entendraient une condition socio-juridique quelconque. Mais cela est contrebalancé à partir du moment où l'on se tourne vers les déterminants possessifs. Il apparaît alors que l'objectivation de l'individu désigné par *seruus* reste une réalité sur laquelle il est difficile de s'étendre par manque de qualificatifs précis. Car parmi tous ces adjectifs, il reste très probable qu'une partie d'entre eux n'ait *in fine* qu'un lien négligeable, qu'il ne le qualifie pas directement.

Nous ne nous étendrons pas sur le sujet et ne présenterons pas les adjectifs de chacune des catégories, justement car il n'est pas possible de les relier explicitement à *seruus*. Malgré tout, la démarche restait pertinente car elle a permis de mettre à jour le lien privilégié et validé entre les déterminants possessifs et lui.

— Analyse qualitative

Pour palier à cette imprécision, nous avons mis au point une fonction qui permet de sélectionner dans l'environnement de *seruus* les adjectifs qui s'accordent en cas et en nombre avec lui, de manière imparfaite et limitée, mais avec des résultats probants et davantage pertinents. Cela étant, elle permet de quantifier les cas des adjectifs les plus employés avec

lui mais également de pointer les formes des lemmes les plus fréquentes.

La fonction prend en entrée des tableaux extraits des *CEMA* qui affichent les mots du texte ainsi que leurs tags respectifs. La mise en parallèle des cas s'effectue par une comparaison des terminaisons. La difficulté de la réalisation, la limite même de cette fonction, correspond à la potentielle pluralité des choix dans l'analyse. Un adjectif de la première classe et un de la seconde peuvent partager une terminaison sans partager leur cas, et étant donné qu'il est impossible de discerner par sa forme à quelle catégorie il appartient, il est possible que la fonction joigne à *seruus* un adjectif qui n'est pas supposé l'être. Toutefois, cette méthode permet d'élaguer considérablement la liste des épithètes cooccurents et d'en affiner la définition. Puisqu'ils sont au même cas, puisqu'il sont dans un espace proche, la probabilité qu'ils interagissent augmente significativement ¹⁶⁸.

Grâce à cette manipulation, il est possible d'extraire deux types d'informations. D'une part, cela permet de mettre en exergue les cas dont l'apparition est la plus prolifique aux côtés de *seruus* siècle par siècle, et de l'autre d'améliorer la précision et la pertinence des graphes issus de l'analyse des cooccurents. La représentation en donuts permet de percevoir la répartition synchronique, en laissant de côté la lecture diachronique. L'importance des cas fluctue entre les siècles, mais des tendances peuvent s'illustrer parmi les déséquilibres. Le 9^e siècle s'illustre par la prépondérance de l'accusatif pluriel qui enregistre à lui seul un taux supérieur à 60%. À sa suite, le génitif pluriel tend à gagner en importance, toute proportion gardée.

Cas des adjectifs accordés à *seruus* entre 800 et 899

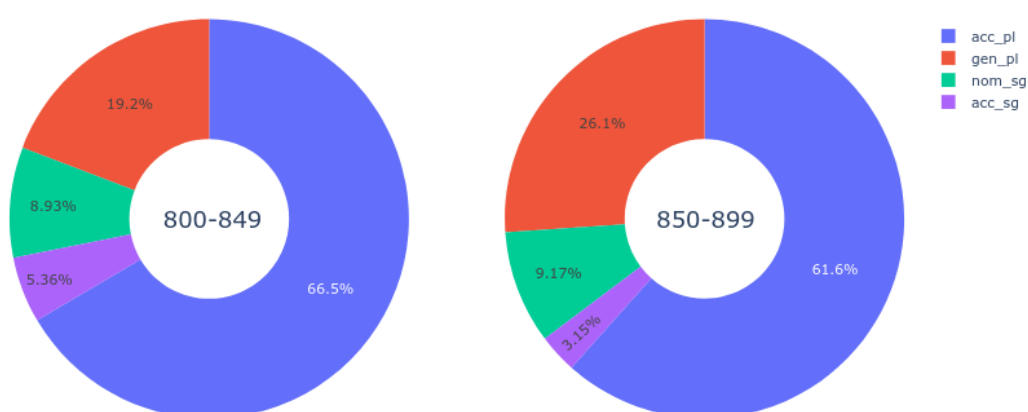


Figure 40 – Pie charts de la distribution des cas des adjectifs accordés à *seruus* au 9^e siècle.

¹⁶⁸ Au vu du nombre de lignes de code de la fonction, il a été jugé préférable de la partager uniquement sur le github.

Cas des adjectifs accordés à *seruus* entre 1200 et 1299

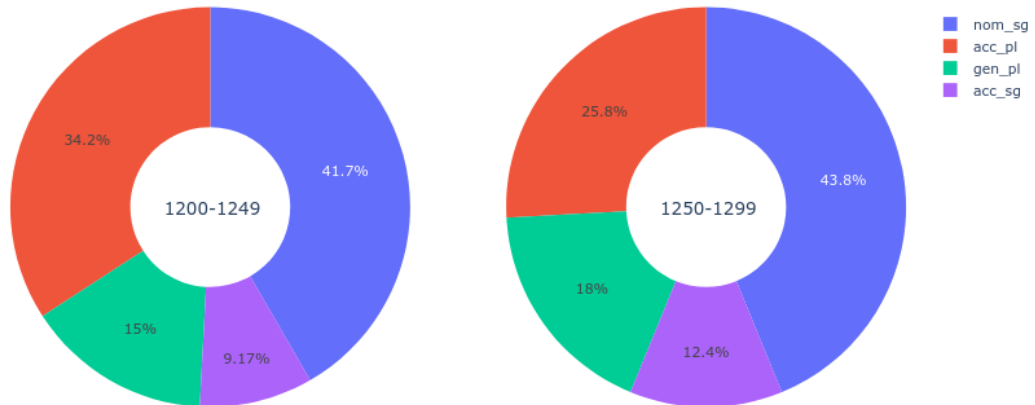


Figure 41 – *Pie charts de la distribution des cas des adjectifs accordés à *seruus* au 13^e siècle.*

Les autres siècles sont témoins d'un déclin progressif de l'emploi d'épithètes à l'accusatif pluriel et de l'essor du nominatif singulier. Dans les faits, cette information ne guide pas quant à la compréhension du mot si ce n'est que plus les siècles passent, plus les adjectifs semblent employés pour définir un *seruus* sujet plutôt qu'objet.

Ici s'arrêtent les facilités de l'automatisation. En effet, l'algorithme n'est pas capable de reconnaître la différence entre un nominatif d'un adjectif de la seconde classe terminant par *-is* et un adjectif de la première classe au datif ou à l'ablatif pluriel. Il n'est alors possible de régler le problème que manuellement. Le travail est laborieux, mais pas impossible. En différenciant les classes, on constate logiquement une transmission d'occurrences du nominatif singulier à l'ablatif pluriel, ce qui réduit le poids du nominatif. L'impact semble mineur sur la tendance générale. Plus le temps passe, plus le mot *seruus* est accompagné d'un adjectif lorsqu'il est sujet ¹⁶⁹.

¹⁶⁹ Pour plus de détail et pour visualiser les graphes pour chaque siècle, *Cfr.* [github](#).

permettre une bonne comparaison. Seule une certaine différence s’illustre entre les épithètes espagnols et les autres. Pour ceux-ci, seul le possessif *tuus* est présent, sans que cela ne signifie grand-chose *in fine*.

	forme_CEMA	occ_CEMA	forme_ang	occ_ang	forme_esp	occ_esp	forme_ita	occ_ita
0	su	37	su	5.0	dilect	1.0	me	2.0
1	me	18	inutil	2.0	omn	1.0	s	1.0
2	tu	12	libe	2.0	present	1.0	NaN	NaN
3	fidel	11	me	2.0	tu	1.0	NaN	NaN
4	inutil	9	omn	2.0	NaN	NaN	NaN	NaN
5	nost	9	quiet	2.0	NaN	NaN	NaN	NaN
6	dilect	7	religios	2.0	NaN	NaN	NaN	NaN
7	iurat	6	ali	1.0	NaN	NaN	NaN	NaN
8	omn	6	dilect	1.0	NaN	NaN	NaN	NaN
9	humil	6	infern	1.0	NaN	NaN	NaN	NaN

Figure 43 – Dataframe des occurrences par pays au 13^e siècle (QLF)

Il n’y a qu’une seule solution pour remédier à ce problème : agrandir le corpus initial en insérant toutes les occurrences depuis le 9^e siècle. Avec cela se perd la pseudo-homogénéité sémantique séculaire initialement recherchée. Si l’on prend pour point de repère les déterminants possessifs, seule l’Espagne se distingue. C’est en effet le seul corpus à plus mettre en avant des déterminants qui reflètent l’humilité.

	forme_CEMA	occ_CEMA	forme_ang	occ_ang	forme_esp	occ_esp	forme_ita	occ_ita
0	ingenu	446	deuot	13.0	uest	27.0	me	14.0
1	su	198	omn	11.0	omni	22.0	libe	13.0
2	nost	156	su	8.0	exigu	18.0	cartulat	11.0
3	omn	95	uest	6.0	indign	13.0	omn	10.0
4	me	82	uniuers	4.0	humillim	8.0	uest	7.0
5	colon	67	dict	4.0	pusill	6.0	colon	6.0
6	uest	60	cunct	3.0	ali	6.0	predict	5.0
7	paupe	54	libe	2.0	nost	4.0	nost	5.0
8	omni	50	sempitern	2.0	sanct	4.0	su	5.0
9	humil	48	religios	2.0	humil	4.0	ecclesiastic	4.0

Figure 44 – Dataframe des occurrences par pays du 9^e au 13^e siècle (QLF)

3.2 VBE

Outre les adjectifs qualificatifs, d'autres catégories grammaticales peuvent intuitivement aiguiller le raisonnement. C'est le cas des verbes, qui en fonction de leur voix (active ou passive) mais surtout en fonction du cas de *seruus*, peuvent s'avérer plutôt pertinents sur les actions que les *serui* tantôt effectuent, tantôt subissent. Ces cas de figure s'expriment par l'emploi soit du nominatif (singulier ou pluriel), soit de l'accusatif (singulier ou pluriel). Les occurrences de *seruus* accordées à l'un de ces cas ont été retenues ainsi que les verbes à proximité qui grammaticalement étaient compatibles. Il aurait été possible d'étendre la recherche au-delà de la relation sujet – verbe – objet, en intégrant par exemple les ablatifs absolus, les compléments d'agent pour les verbes passifs ou encore les compléments indirects du verbe au datif, mais cela nécessite un bagage en code dont nous ne sommes pas suffisamment muni. En outre, il eût été possible d'analyser les verbes en fonction du temps auxquels ils sont conjugués, mais le peu d'occurrences rend les chiffres trop peu parlants que pour en tirer une généralité. La piste reste néanmoins ouverte pour une prolongation de recherche.

La sélection s'arrête donc juste aux nominatifs et aux accusatifs, et à tous les verbes présents, tous modes confondus, ainsi que sous forme de participe (présent ou parfait). Les participes passent par le même traitement algorithmique que les adjectifs, tandis que l'obtention des verbes actifs et passifs nécessite la conception d'une nouvelle fonction. La description des verbes sera effectuée par demi-siècle, et à chacun seront affectés quatre axes : les verbes dont *seruus* est potentiellement sujet, ceux dont il est objet, les participes qui s'accordent avec lui, et enfin les verbes à la voie passive. Toutes ces données seront représentées sous forme de tableau avant d'être détaillées par la suite. Trois jalons seront posés pour évaluer l'évolution des verbes entre le 9^e et le 13^e siècle. Un pour le début du 9^e siècle, un second pour la fin du 13^e siècle et un dernier entre les deux, au 11^e siècle.

La sporadicité de l'analyse qualitative est due à des limites de codes. Pour bien faire, il aurait été nécessaire de créer une fonction qui puisse relier toutes les conjugaisons d'un verbe au lemme initial, ce qui paraissait trop compliqué et trop chronophage. À la place, une unique moitié des demi-siècles formeront les balises de l'étude. En outre, il n'a pas été possible de relier les verbes à un même lemme, toujours en raison de la pluralité des terminaisons difficiles à considérer dans le traitement des données, c'est pourquoi seules les formes de verbes les plus fréquentes seront prises en compte.

Entre le nominatif singulier sujet et le pluriel, il apparaît une différence de lexique flagrante au début du 9^e siècle. Outre les verbes passe-partout que sont être et avoir, le singulier témoigne d'une signification qui peut se rapprocher d'actions témoignant d'une

certaine, sinon autorité, du moins responsabilité. On retrouve des verbes comme *suggessit* (déléguer, confier, attribuer) ou *commiserit* (commettre, confier, administrer un sacrement) qui propulsent le *seruus* dans les échelons d'une hiérarchie dont il ne serait pas le dernier. Néanmoins, le principal verbe avec lequel il fonctionne est *debet*, ce qui implique donc certaines obligations. Le *seruus* doit.

À l'inverse, les occurrences plurielles manifestent une position plus secondaire. Le *seruus* paie, s'acquitte (*soluant*), il revient à quelqu'un, appartient à quelqu'un, est la propriété de quelqu'un (*pertinent*). Les listes de verbes ne se recoupent pas ou peu, si l'on prend en compte *être* et *avoir*. Cette différence peut s'expliquer par l'hypothèse suivante. Dans l'expression *seruus seruorum dei*, le nominatif a valeur d'apposition, il donne des informations supplémentaires sur le sujet et s'accorde au nominatif. Par exemple, dans le document référencé "0800_pflugk-harttung_236150", qui débute comme suit « Leo episcopus, seruus seruorum Dei », *Leo* est le sujet auquel sont rattachés deux appositions : *episcopus* et *seruus*. La fonction de tri a évincé les apparitions de la formule *seruus seruorum dei* dans les corpus, mais cette apposition peut se retrouver éventuellement plus loin dans le texte. Toutefois, il est également possible que cela ne s'y réfère pas et renvoie à une autre possible hypothèse : celle du serviteur servant et non plus du serviteur de Dieu. À cet égard, soulignons un exemple dans le document "0818-stbenolltsur_87515" :

« Ibique ueniens Fredulus [...] requirebat ei quod seruus erat domno Ludouici
Imperatoris de Patricaco uilla »

Que l'on pourrait traduire prudemment comme suit :

« Et Fredulus venant ici leur demandait que le seigneur de l'empereur Louis
ait le serviteur de la maison Patricaco ».

On revient donc à une sorte d'objectivisation marquée par le datif possessif. Cela étant, aucun verbe au nominatif singulier ne renforce cette seconde hypothèse qui pourrait être tout aussi bien marginale qu'incontournable.

L'analyse des accusatifs singuliers et pluriels (sans discernement) donne davantage d'idées sur la manière dont sont traités les *seruos*. Congédiés, aliénés, vendus, pardonnés, trahis, pris en possession, liés par des obligations morales, hérités. Dès lors qu'il se trouve à l'accusatif dans une charte, le *seruus* prend une connotation tout à fait passive où il est ballotté au gré des décisions qu'il ne prend pas. En tant que complément d'objet direct, ce cas évoque les actions qu'il subit. En ce qui concerne l'aube du 9^e siècle, les verbes qui se trouvent dans un périmètre proche de *seruus* lorsque ce dernier est accordé à l'accusatif (42) sont nettement plus nombreux que pour le nominatif (23 occurrences cumulées, 30 si l'on considère les modes passifs).

La traduction des participes ne fait que renforcer cette idée. Les verbes qui accompagnent *seruus* sont pour beaucoup des verbes que l'on retrouve également dans la listes des accusatifs (*habito, pertineo, commoror, famulo,...*). Les participes les plus fréquemment utilisés avec *seruus* proviennent des verbes *ordinare* (ordonner, installer, arranger) et *famulare* (accomplir son devoir de vassal, asservir, dompter).

En raison d'un manque de données, le 10^e siècle ne sera pas abordé ici. Qu'il s'agisse des nominatifs ou des accusatifs, les occurrences sont bien trop rares pour se faire une idée de l'évolution des lexiques. Pour cause, le 10^e siècle est un siècle de pauvreté documentaire à l'origine discutée, soit liée aux troubles de l'anarchie féodale, soit liée à des troubles généralisés de leur conservation¹⁷⁰. Moins de documents signifie moins de verbes, moins d'occurrences.

Cela étant, la courbe d'occurrences repart à la hausse au début du 11^e siècle. La tendance a quelque peu évolué. Le lexique change et le nominatif singulier a perdu sa valeur précédente. Les différences entre les verbes qui s'accordaient au singulier et au pluriel se sont amenuisées. Les occurrences au pluriel restent plus nombreuses, mais deviennent moins diverses. Pour l'essentiel, les *serui* sont reliés à l'occupation de la terre, comme en atteste l'occurrence majeure de *teneo* (occuper un fief, un domaine).

Par ailleurs, les verbes les plus fréquents desquels les *seruos* sont objets suggèrent pour la plupart une idée de transmission. Ils sont cédés, donnés (*dono*), concédés (*concedo*) ou livrés (*trado*). Les *serui* sont liés à leur terre, et donc indirectement à leur seigneur. Ce genre d'échanges de *serui* n'était pas rare. Lorsqu'une terre changeait de main, les exploitants de la terre en faisait tout autant. Les déplacements collectifs de population s'inscrivent dans le cadre de l'économie domaniale. Des couples sont ainsi transférés d'une tenure à une autre pour différentes raisons, comme opérer à des défrichements¹⁷¹. Certaines chartes attestent justement de cela¹⁷².

Par la suite, les cohabitations entre *seruus* et les verbes s'estompent brutalement. Tant les participes que les verbes actifs, tant les nominatifs que les accusatifs, tous les indicateurs partent à la baisse. Il n'y pas de renouvellement significatif dans le vocabulaire employé. Si l'on peut considérer le 10^e siècle comme une période creuse d'un point de vue de la documentation, la situation aux 12^e et 13^e siècles est parfaitement inverse. L'accroissement de la production documentaire est un symptôme bien connu des médiévistes

170 BERTRAND, P, *À propos de la révolution de l'écrit (X^e-XIII^e siècle). Considérations inactuelles*, dans *Médiévales*, n°56, p.72.

171 FELLER, L., *Paysans et seigneurs au Moyen Âge*, 2017, p.38

172 0949_cluny_2160

qui parlent de véritable « révolution de l'écrit »¹⁷³. Pourquoi assistons-nous alors à une hausse des liaisons entre *seruus* et des verbes au début du 11^e siècle et à un affaiblissement de ces dernières au cours des siècles les plus riches en documentation ?



Figure 45 – Évolution des cooccurrences de verbes avec *seruus*

Au vu des décisions prises dans la confection des fonctions, il est possible que la modification de certains paramètres soit en cause ici. Si les intervalles avaient été étendus à la phrase entière par exemple, d'autres verbes auraient pu être mis en évidence. Mais il est également possible d'interpréter cela autrement. Parallèlement à cette baisse des verbes dans l'entourage de *seruus*, on remarque logiquement une baisse similaire du nombre d'occurrences de *seruus*. Rappelons que la titulature pontificale (*seruus seruorum dei*) n'a pas été considérée dans le cadre de la présente étude, l'objectif étant d'évaluer la signification du mot au sein du texte. De ce fait, si les occurrences totales de *seruus* croissent drastiquement aux 12^e et 13^e siècles, mais que par la même occasion elles s'amoindrissent dès lors que la titulature n'est plus considérée, cela indique que le mot *seruus* tend à se perdre dans la masse des mots employés dans les actes médiévaux de la fin du Moyen Âge central. Autrement dit, plus de documents sont produits, plus d'occurrences apparaissent, mais ces occurrences trompent le jugement en cela qu'elles sont majoritairement issues de cette formule pontificale. Pour le reste, et malgré la multiplication de la documentation, le mot *seruus* tend à se raréfier dans le corps des actes.

173 BERTRAND, P, *À propos de la révolution de l'écrit (X^e-XIII^e siècle). Considérations inactuelles*, dans *Médiévales*, n°56, p.79.

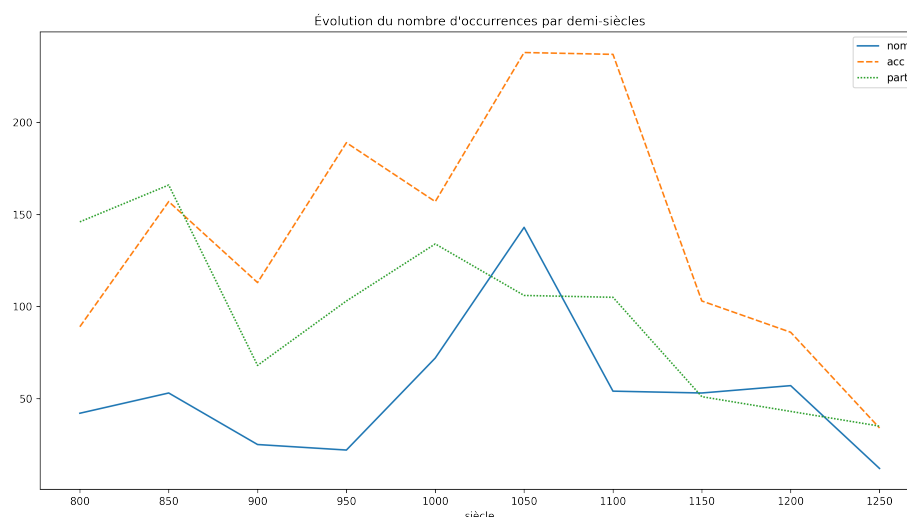


Figure 46 – Évolution du nombre d'occurrence de *seruus* par demi-siècle sans *seruus seruorum dei*.

3.3 SUB

Dans un dernier registre, l'étude des substantifs qui gravitent autour de *seruus* peut également donner des indications sur son emploi et sur le contexte lexical dans lequel il est plus prompt à apparaître. Cette dernière manipulation de données s'est limitée au même intervalle que précédemment, à savoir un intervalle [-5;5]. La collocation récurrente de certains mots témoigne d'une association privilégiée entre les mots « reliés par une structure syntaxique et dont les affinités syntagmatiques se concrétisent par [cette] récurrence en discours »¹⁷⁴. Cette répétition de syntagmes donne un aperçu de l'organisation structurale du langage au sein du document médiéval. Mais qu'en est-il des différentes valeurs paradigmatiques que peuvent prendre ces syntagmes répétés ? Sont-elles dépendantes des mots avec lesquels *seruus* interagit, sous forme d'expressions presque cristallisées ?

Cette sous-partie va chercher à éclairer les relations que certains substantifs peuvent entretenir avec *seruus* en comparant les lemmes les plus fréquents de son entourage pour les cinquante premières années de chaque siècle entre le 9^e et le 13^e siècle. Pour l'entière du corpus, le mot le plus fréquemment employé autour de *seruus* est *deus*. Il est constamment présent au moins deux fois plus que le second substantif le plus fréquent, et ce nonobstant l'absence de la titulature pontificale. De plus, dans la lignée de ce qui a été dit dans la sous-partie consacrée aux verbes, le nombre de collocation des substantifs diminue sévèrement au tournant du 12^e siècle, étant donné que les apparitions de *seruus*,

¹⁷⁴ MAYAFFRE, D., *De l'occurrence à l'isotopie : les co-occurrences en lexicométrie*, dans *Syntaxe et Sémantique*, n°9, 2008/1, p.56.

dans l'ensemble, s'amenuisent.

Au 9^e siècle donc, sur les 713 occurrences de *seruus*, *deus* lui est joint 435 fois pour faire mention des serviteurs de Dieu. L'analyse de cette expression sera le sujet d'un sous-chapitre distinct. À la place, il est intéressant de regarder la liste des autres lemmes avec lesquels il est utilisé, à commencer par *terra*. Comme en témoignent les *CEMA*, le mot *terra* est souvent relié à *seruus* par une expression quelque peu figée : « homines tam inguenos quam et seruos qui super terram residere »¹⁷⁵. Cette formule désigne les hommes, tant libres que supposément non, qui sont établis sur une terre précise. Souvent aussi, elle est complétée par une mention de détention, de provenance. Sont alors rajoutées des précisions comme « homines ipsius monasterii »¹⁷⁶ ou « ipsius ecclesie »¹⁷⁷, ou encore « homines eiusdem monasterii »¹⁷⁸. Autrement dit, beaucoup de ces *seruos* qui sont opposés aux *ingenuos* sont dépendants, et au même titre que ces derniers, d'une institution en particulier (église ou monastère). Cette expression se retrouve tant dans les territoires aujourd'hui francophones que germaniques, essentiellement à partir de 814, année de la mort de Charlemagne et de l'accès au trône de Louis le Pieux. L'expression apparaît rapidement et sporadiquement dans des institutions ecclésiastiques de tout l'Empire, aussi bien parisiennes (Saint-Denis), bourguignonnes (Mâcon) que germaniques (Ratisbonne). La première mention date de 508¹⁷⁹, mais est parfaitement isolée, et malgré quelques apparitions discrètes au 7^e siècle, il faut attendre la seconde moitié du 8^e siècle pour la voir véritablement éclore. Le *seruus* est donc rattaché à la terre. On doit cette multiplication de la formule à Louis le Pieux qui a expédié de nouveaux diplômes dès le début de son règne, diplômes « en harmonie avec les acquisitions intellectuelles de la Renaissance carolingienne, [à savoir] une écriture plus claire, avec des mots bien séparés, et obéissant à des canons précis ; une langue clarifiée et une meilleure formulation juridique des clauses »¹⁸⁰. La multiplication de la formule découle donc de la restructuration diplomatique initiée par Louis le Pieux, ainsi que de la multiplication des documents résultant de la gestion administrative d'un Empire. Cette formalisation du langage diplomatique ruisselle vers les autres institutions impériales grâce notamment à des recueils de formules (*Formulae imperiales*).

175 0814_aniane_133561

176 0816_louis-le-pieux_211720

177 0826_sens-eveques_53747

178 0843_autun-eglise_666

179 0508_mgh-dd-mer_216648

180 http://elec.enc.sorbonne.fr/cid/cid1992/art_02

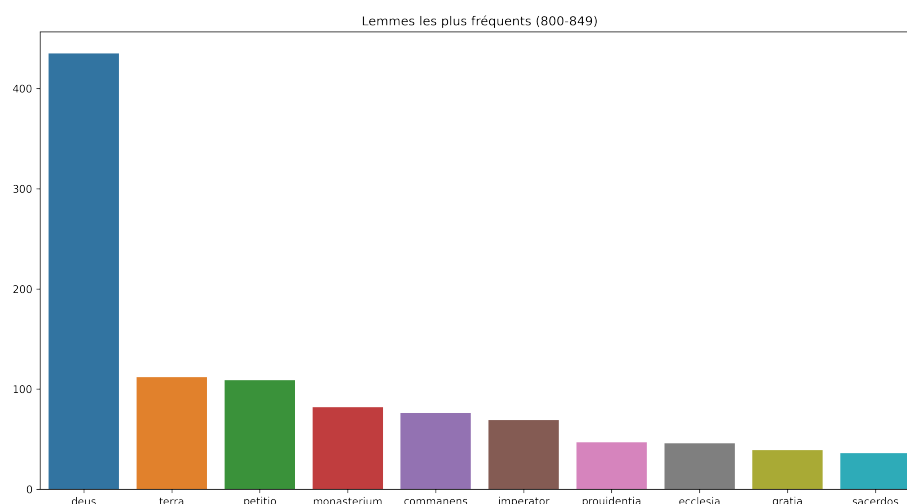


Figure 47 – *Substantifs les plus fréquemment employés avec seruus dans la première moitié du 9^e siècle.*

Par cette expression, on peut expliquer la présence d'autres lemmes, comme *monasterium* ou *ecclesia*. En ce qui concerne *petito*, il apparaît quasi-exclusivement à l'ablatif pluriel accompagné soit de la préposition *si* soit de la préposition *cum*. Cela donne lieu à une autre formule récurrente dans les chartes : « Cum/Si petitionibus seruorum dei iustis et rationabilibus diuini cultus amore fauimus », que l'on retrouve principalement dans les chartes émises par le pouvoir impérial (Louis le Pieux¹⁸¹, Pépin I d'Aquitaine¹⁸² ou encore Charles le Chauve¹⁸³). Dans sa forme plus commune, elle s'arrête à *iustis*. Il s'agit là d'une formule répandue dans les actes carolingiens. Pour bien faire, il serait nécessaire de consulter des formulaires carolingiens, sortes de « modèles de documents, le plus souvent rassemblés dans des collections, ou formulaires, et destinés à aider à la rédaction d'actes futurs »¹⁸⁴. La rédaction des actes, tous types confondus, bénéficie d'un cadre formel déterminé, autrement dit, l'écriture juridique prend une tournure figée, tournée vers l'efficacité et le désir de compréhension¹⁸⁵.

La présence du lemme *imperator*, et conséquemment de *prouidentia*, s'explique par une nouvelle formule protocolaire que l'on retrouve en fin de suscription¹⁸⁶. Lorsque l'empereur se présente, il attache à son nom une titulature impériale : « Ludouicus diuina ordi-

181 0815_louis-le-pieux_211704

182 0829_pepin-i-pepin-ii_277046

183 0849_charles-le-chauve_275706

184 RIO, A., *Les formulaires et la pratique de l'écrit dans les actes de la vie quotidienne*, dans *Médiévales*, p.4.

185 http://elec.enc.sorbonne.fr/cid/cid1992/art_02

186 GUYOTJEANNIN, O., *Diplomatique médiévale*, p.72

nante *prouidentia Imperator augustus* », qui n'apparaît dans cette forme standardisée qu'à partir de Louis le Pieux dans un diplôme de 814¹⁸⁷ et qui le présente comme l'Auguste Empereur ordonné par la providence divine, une forme d'expression alternative et antérieure à « *Ludouicus dei gratia* ». Au 9^e siècle, néanmoins, existe l'expression inversée de « *Karolus gratia dei* », une formule de dévotion qui tente de légitimer le pouvoir établi en lui donnant un fondement divin¹⁸⁸. Se référer à Dieu est indiscutablement le moyen le plus sûr de se garantir l'adhésion du plus grand nombre dans l'exercice du pouvoir.

Dans l'ensemble, la répétition la plus fréquente de substantifs dans un périmètre proche de *seruus* provient intégralement de formules protocolaires, sans renvoi au corps du document. Au 9^e siècle donc, l'emploi le plus stable de *seruus* est figé dans des expressions latines.

Les collocations tendent à évoluer tandis que le temps suit son cours inexorable. À peine un siècle plus tard, l'essentiel des mots est remplacé, et réfère de la sorte à un changement dans l'emploi de *seruus* dans l'acte. Par exemple, le second mot le plus fréquemment utilisé autour de *seruus* est désormais *nomen*. Tout comme pour *uxor*, l'occurrence rapprochée de *nomen* se déroule au sein du dispositif de l'acte médiéval¹⁸⁹, au cœur même de l'explication de la transaction généralement entamée par le verbe *dono*¹⁹⁰. Pour ainsi dire, ces deux mots sont employés pour présenter et nommer d'une part le *seruus* impliqué dans la transaction, et de l'autre pour préciser le nom de son épouse et de sa progéniture. Ce phénomène a déjà été remarqué par Eliana Magnani dans une recherche similaire au sein des *Chartae Burgundiae Medii Aevi*. Il y est question de la cooccurrence fréquente entre *uxor* et *seruus* qui interagissent eux-mêmes avec *nomen* « lors de transferts dont des hommes dépendants, et leurs familles – épouse et enfants –, sont eux-mêmes l'objet »¹⁹¹. Comme elle le souligne également, l'apposition systématique du lemme *nomen* joue comme un marqueur social flagrant, une désignation de dépendance d'un *seruus* à son *dominus*¹⁹².

L'apparition d'*ancilla* dans le graphe ne fait pas écho à un gain notable d'importance ou à l'apparition soudaine d'une nouvelle association, mais est probablement due au manque de documents datant du 10^e siècle qui met en exergue d'autres formules. En effet, ce couple de substantifs est déjà observable dans des documents datant de la fin du 6^e siècle et proliférant au cours des siècles suivants. On retrouve donc avec moins de 25 occurrences

187 0814_louis-le-pieux_211684

188 GUYOTJEANNIN, O., *Diplomatique médiévale*, p.73

189 *Ibid*, p.80.

190 0918_cluny_1562

191 MAGNANI, E., *Uxor et femina. Enquête sur la désignation des femmes dans les documents diplomatiques bourguignons (IX^e-XI^e siècle)*, p.4.

192 *Ibid*, p.8.

la formule « *seruus et ancilla* », qui attribue à *ancilla* une valeur similaire à celle de son homonyme.

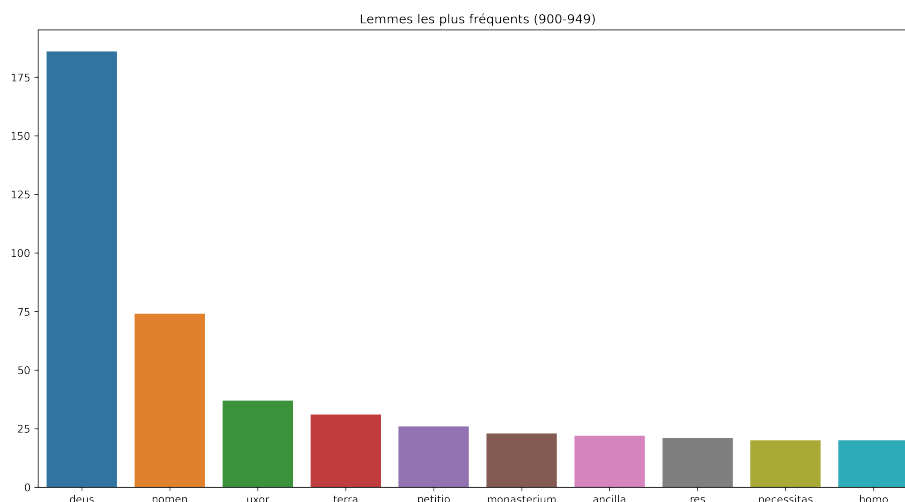


Figure 48 – *Substantifs les plus fréquemment employés avec *seruus* dans la première moitié du 10^e siècle.*

Pour le reste, les expressions impliquant les lemmes *terra* et *petitio* restent présentes en moindre mesure, et perdent progressivement en poids. Le lemme *homo* prend même l’assise sur d’autres. On assiste à un coulisement lexical lent du protocole au texte lui-même.

Le début du 11^e siècle apporte avec lui son lot de changements. L’importance de *deus* autour de *seruus* ne diminue en rien. La désignation du *seruus* par son nom prend quant à elle de l’ampleur, tandis que *uxor* sort de son giron direct. Communément, et dans un corpus plus fourni, *ancilla* s’extrait de la masse. Ce lemme est majoritairement joint à *seruus* à l’accusatif pluriel. Émerge également au 11^e siècle le lemme *seruus*. Un regard plus analytique sur la situation balaye les interrogations de sa présence. En effet, à y regarder de plus près, presque toutes les apparitions d’un second *seruus* après le premier font référence à la formule d’humilité en souscription « *seruus seruorum dei* », mais inversée (« *seruorum dei seruus* »)¹⁹³. De ce fait, elle est passée au travers des mailles de la fonction destinée à l’évincer.

En ce qui concerne les occurrences de *dominus*, elles peuvent signifier plusieurs choses. Soit elles remplacent le lemme *deus* et renvoient explicitement au Seigneur (« *seruus se-*

¹⁹³ 1051_plpape_239194

ruorum domini »¹⁹⁴), soit elles font référence à un homme au sommet d’une hiérarchie ecclésiastique (évêque, pape, abbé¹⁹⁵). Enfin, il reste également à mentionner le lemme *liber* qui, comme le fait *ancilla*, accompagne *seruus* et transmet une information relative à un statut socio-juridique, mais davantage par antonymie. Il intervient à la suite de *seruus* accompagné soit de la conjonction *et*, soit de la conjonction *uel*, soit de la conjonction *siue*.

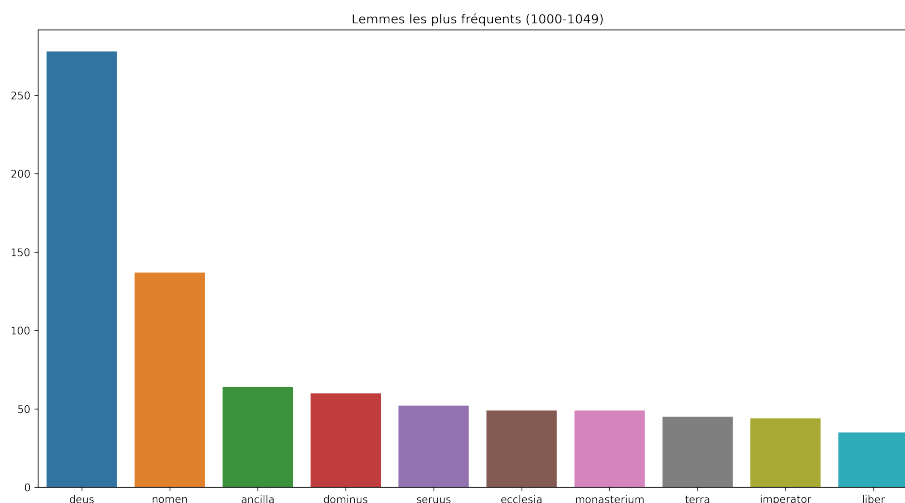


Figure 49 – Substantifs les plus fréquemment employés avec *seruus* dans la première moitié du 12^e siècle.

Avec l’arrivée du 12^e siècle, on n’assiste pas à un renouvellement des tendances, on conserve une certaine stabilité. Le lemme *filius* n’est en fin de compte qu’une cooccurrence anecdotique qui brille par les multiples contextes déstructurés dans lesquels il apparaît. *Episcopo*, quant à lui, n’est employé que dans le préambule pour préciser l’identité de l’auteur (intitulation¹⁹⁶). Dès le 12^e siècle, on remarque un essoufflement dans les collocations originaires de la baisse d’occurrences de *seruus* dans les documents, qui se traduit encore plus clairement au cours du 13^e siècle.

Tout compte fait, la première information qui transparait est l’omniprésence dans l’environnement de *seruus* du lemme *deus*, qui soit lui est directement relié¹⁹⁷, soit le talonne

194 on_250286

195 1093_lerins_83490 : on retrouve la mention suivante « Priuilegium dom[i]ni Urbani papa »

196 GUYOTJEANNIN, O., *Diplomatique médiévale*, p.72

197 L’expression de “serviteurs de Dieu” est incontournable dans les chartes médiévales, et cela même sans l’expression de *seruus seruorum dei*. Elle concerne tant l’ensemble des fidèles appartenant à la communauté chrétienne que les clercs ou les moines. Ce faisant, dès qu’un acte concerne un membre de ces groupes, elle le présentera comme étant un *seruus Deus*. Les exemples sont légions, tous cas confondus, mais plus particulièrement marqué encore au nominatif singulier et à l’accusatif pluriel.

de quelques mots à peine ¹⁹⁸. Dans tous les cas, un tel foisonnement d'occurrences de *deus* sous-entend l'importance que revêt la figure divine dans l'appréciation du sens de *seruus*.

En outre, La naissance des formules diplomatiques au 9^e siècle influe drastiquement sur l'étude quantitative telle qu'elle est menée ici. Comme nous l'avons observé, les collocations tendent à se renouveler, même si certaines occurrences demeurent comme figées dans le marbre ¹⁹⁹. Dans l'ensemble, ce que l'on peut constater et ce que l'on peut conclure surtout, c'est que les formules diplomatiques pèsent énormément. Avec de telles répétitions, il est légitime de s'interroger sur le degré de fiabilité d'une étude cooccurrence. Car même si ces formules donnent un sens à *seruus* et le mettent en contexte, leur adoption peut se révéler rapidement anachronique et uniquement destinée à l'usage, tandis que son sens réel, lui, évolue et s'en distancie.

On observe aussi l'amorçage d'une transition du protocole vers le corps du texte au cours du 9^e siècle qui aboutit dès le début siècle suivant. Cela signifie que les emplois relatifs aux titres vont à la baisse, et que les informations propres au message du document diplomatique catalysent les occurrences de *seruus*.

Nonobstant, les observations qui ont été tirées dans le cadre de l'étude des substantifs indique une nouvelle fois, et clairement, cette bipolarité sémantique qui oscille entre un sens de liaison de l'homme à la terre et de serviteur spirituel. Toutefois, il est important de rappeler que ces constats ne peuvent être soulevés que pour les 9^e, 10^e et 11^e siècles, car la chute des occurrences qui s'opère aux 12^e et 13^e siècles limite, voire empêche de les prolonger.

3.4 Le rôle central de la Bible dans le langage médiéval

Sitôt que l'on se réfère aux mots, il devient nécessaire de se tourner vers les textes. Car ce sont bien eux qui les propagent, qui les emploient. La pensée médiévale est structurée par son langage, les mots n'existent pas comme des entités indépendantes, mais s'insèrent dans un tissu large de représentations sociales dont les textes furent les plus efficaces propagateurs. Or, il est un texte dans l'Occident médiéval qui peut être considéré comme « le Livre par excellence » ²⁰⁰ : la Bible. Tous ceux qui possédaient l'art de lire étaient enjoint à s'y plonger, clercs comme laïcs ²⁰¹. L'influence linguistique d'une telle entreprise

198 0839_solignac_53787 : "Sed quicquid inde fiscus noster exigere potest, totu[m] nos pro eterne retributionis premio in lu]minaribus ejusdem ecclesiae et receptione pauperum atque stipendiis servorum Christi ibidem Deo famulantium concedimus."

199 Les exemples de *terra* et d'*ancilla* vont dans ce sens.

200 BENOIT, J.-L., *Clef du texte, clef du royaume. La lecture de la Bible au Moyen Âge comme paradigme de la littérature*, dans *Les Clefs des textes médiévaux*, 2006, p.303.

201 LOBRICHON, G., *La Bible au Moyen Âge*, p.28-54.

littéraire n'est pas à sous-estimer, loin de là et pour comprendre le système des représentations médiévales, il importe d'envisager le rôle structurant que cet ouvrage a cultivé. Parallèlement à la Bible, les exégèses, commentaires du texte sacré, ont également pulvéalisé au cours du Moyen Âge. Nombreux furent les théologiens à s'inspirer des textes des Pères de l'Église (Origène, Saint-Augustin, Saint-Jérôme,...). La production intellectuelle de l'époque tournait énormément autour de ces textes fondateurs.

Le lien fondamental entre la réalité sociale que cachent les mots et la centralité du texte biblique est absolument primordial, car les lectures de la Bible sont plurielles. L'une des plus grandes difficultés dans l'analyse des textes réside justement dans le passage du sens littéral au sens spirituel, qui peut être induit par l'emploi de métaphores et d'analogies²⁰². Dès lors qu'il est possible d'interpréter les sens des textes, l'emploi des mots se teinte de multiples facettes. Tenter de comprendre les mots pour eux-mêmes n'a pas de sens si l'on ne le compare pas aux sens qu'ils détiennent dans la Bible.

Comme le dit Frans Van Liers, « One cannot understand the medieval world without appreciating the scope of medieval people's engagement with biblical stories, characters and images »²⁰³. La Bible était lue partout et par tous ceux qui le pouvaient. La transition entre le monde antique régi par le droit romain et la genèse de l'État aux 12^e et 13^e siècles laisse un vide de plusieurs siècles au cours desquels la Bible fut l'outil normatif de référence, l'une des principales sources d'autorité. Il s'agissait là d'un « auxiliaire indispensable pour comprendre, ordonner et réformer la société »²⁰⁴. Les mots de la Bible latine ont pour beaucoup contribué à former les représentations sociales les plus centrales.

Cependant le texte ne reste pas figé, stable et intangible tout au long de la période médiévale. Des premières traductions latines issues des textes grecs, et jusqu'à la stabilisation de la « Vulgate » votée au concile de Trente en 1542, le texte connut des ajouts, des rognages, bref, des ajustements. Sous Charlemagne, le texte tend toutefois à se stabiliser et la version de la « Vulgate » prend le pas sur les modèles antérieurs de *vetus latina*²⁰⁵. Cependant, les évolutions formelles de la Bible ne sont pas significativement structurantes sur le fond, c'est-à-dire qu'il y n'y a pas de reconsidération fondamentale du système de pensée. Chaque croyant, lettré ou non, continuait de voir en la Bible une réalité indiscutable, un socle relativement stable dans sa perception du monde.

Il faut donc penser le Moyen Âge comme une structure globalement homogène, où la pensée chrétienne orchestre la vie sociale. À partir de là, on peut se demander comment

202 DAHAN, G., *Lire la Bible au Moyen Âge : essais d'herméneutique*, p.20.

203 VAN LIER, F., *An introduction to the Medieval Bible*, p.11.

204 LAUWERS, M., *Usages de la Bible et institution du sens dans l'Occident médiéval*, p.8.

205 BOGAERT, P.-M., *La Bible latine des origines au Moyen Âge*, p.140.

une croyance si répandue et omniprésente dans les textes, la culture, les rites et les traditions pourrait se différencier du langage qui a servi à l'instaurer. Est-ce réaliste de croire que les auteurs médiévaux aient pu penser les mots dans un contexte particulier sans les relier à leur sens biblique ? Est-ce que le lemme *seruus* présent dans les chartes peut être interpolé autrement que comme dans les Écritures ?

L'historiographie plus ancienne avait tôt fait d'organiser la recherche par subdivisions thématiques, à distinguer l'« histoire ecclésiastique » de l'« histoire politique ». En France, avant les années 1950, les deux disciplines étaient traitées séparément, probablement en conséquence de la scission entre Église et État au début du siècle. Avant cela, l'histoire de l'Église n'était pas au centre des préoccupations²⁰⁶. Or, comme nous venons de le préciser, étudier un aspect de la société médiévale sans tenir compte de son rapprochement avec l'institution chrétienne est un absolu nécessaire.

Dans les actes médiévaux, le sens de *seruus* est double et s'inscrit dans une sorte de dichotomie flagrante entre les membres d'une communauté ecclésiastique perçus comme des « serviteurs de Dieu », en opposition à des hommes fixés à des terres, échangeables, interchangeable, et qui répondent à une réalité sociale déspiritualisée. Si cette opposition sémantique apparaît dans les chartes, qu'en est-il de la Bible ? Si le discours de la Bible penche pour une interprétation unique du mot *seruus*, cela signifie que les scribes médiévaux devaient d'une manière ou d'une autre, dans leur système de représentation, faire le lien entre ces deux réalités. À l'inverse, la présence d'une certaine polysémie au sein même de la Bible pourrait amener à une conception plus nuancée de cette hypothèse.

C'est pourquoi, à l'aide du programme de textométrie TXM²⁰⁷ et du texte de la Vulgate lemmatisée²⁰⁸, seront extraits les cooccurents de *seruus* présents dans le texte sacré. Pour bien faire, l'analyse sera séparée en deux parties avec comme point de pivot le passage de l'Ancien Testament au Nouveau Testament. Pour cause, les deux ont été transposés par écrit en des temps bien espacés et peuvent donc témoigner d'une plus grande variance sémantique. L'Ancien Testament voit ses premiers textes, comme la Genèse, couchés par écrit dès le 8^e siècle avant Jésus-Christ²⁰⁹, tandis que le Nouveau Testament voit émerger

206 LAUWERS, M., *L'Église dans l'Occident médiéval : histoire religieuse ou histoire de la société ? Quelques jalons pour un panorama de la recherche en France et en Italie au XX^e siècle*, p.15.

207 Heiden Serge, Magué Jean-Philippe, & Pincemin Bénédicte. (2010). TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement. In JADT 2010 : 10th International Conference on the Statistical Analysis of Textual Data (pp. 12 p.). Rome, Italie.

208 Nous n'avons pas d'informations quant à la date d'édition du texte. Il se peut donc que l'édition soit relativement tardive et quelque peu anachronique. Nonobstant, elle reste une source pertinente de comparaison lexicale.

209 UEHLINGER, C., *Genèse 1-11*, dans RÖMER, T., MACCHI, J.-D. et NIHAN, C., *Introduction à l'Ancien Testament*, 2009, p.204-207.

ses textes dans le courant du premier siècle après Jésus-Christ²¹⁰.

3.4.1 Ancien Testament

Une recherche sommaire du lemme *seruus* dans le corpus, composé de 44 livres différents, retourne 631 occurrences réparties inégalement entre les textes. Le Pentateuque en référence 108 à lui tout seul, soit 17% des occurrences totales, et dont 38 proviennent de la Genèse. Pour le reste, les livres des Rois cumulent 85 occurrences, et les Livres de Samuel 101. 11 des livres n'en possèdent aucune, tandis que plus de la moitié ne voit *seruus* apparaître en deçà de dix fois. La teneur de chaque livre en est la cause car chacun aborde des événements qui lui sont propres. On peut distinguer quatre sous-catégories de textes au sein de l'Ancien Testament qui justifient une telle hétérogénéité : le Pentateuque, les livres historiques, les « Hagiographes » et les Prophètes²¹¹. À bien y regarder, les livres historiques et le Pentateuque concentrent à eux seuls une très grande majorité des occurrences, aux dépens des deux autres, bien que les psaumes comptabilisent un nombre élevé d'apparitions (43).

Parmi toutes les occurrences, 189 sont soit au nominatif pluriel soit au génitif singulier, 141 sont présentes au nominatif singulier, 96 à l'accusatif singulier, 80 à l'accusatif pluriel, 62 au génitif singulier et 60 au datif ou à l'ablatif singulier. Il n'y aucune trace de datif ou d'ablatif pluriel. En ce qui concerne la forme au nominatif singulier, *seruus*, elle apparaît de paire avec *tuus* 57 fois, et si l'on se rapporte à la combinaison lemmatisée, elle se répète 226 au total, soit près d'un tiers des occurrences. La présence de ce déterminant possessif fait écho à une interjection, à une interpolation, à une forme de dialogue.

Si l'on se penche sur les lemmes qui lui sont cooccurrents, et plus particulièrement sur ceux détenant l'indice d'association le plus élevé (fonction de TXM), on retrouve parmi eux des déterminants possessifs (*meus* (21) et *tuus* (91)), des noms communs (*ancilla* (27), *seruus* (17), *rex* (12), *domina* (10), *oratio* (6)) et même un verbe (*loquor* (8)). Certains font également partie des lemmes à la co-fréquence la plus élevée, à l'instar de *tuus* (392), *meus* (195), *rex* (121), *loquor* (63), *seruus* (55) et *domina* (44), tandis que d'autres font leur apparition (*dominus* (180), *suus* (164), *ego* (158) et *dico* (127)). Il est difficile de tirer de là une définition du mot, mais l'on peut toutefois observer que des mots tels que *dominus*, *ancilla* ou même *seruus*, que l'on retrouve au sein des actes diplomatiques, sont aussi prédominants dans cet autre contexte.

210 E. BROWN, R., *Que sait-on du Nouveau Testament ?*, 2011, p.169

211 UEHLINGER, C., *Genèse 1-11*, dans RÖMER, T., MACCHI, J.-D. et NIHAN, C., *Introduction à l'Ancien Testament*, 2009, p.22.

3.4.2 Nouveau Testament

Le registre de cooccurrence prend une nouvelle teinte dans la seconde moitié de la Bible. Bien que *seruus* y soit plus discret (quelques 121 occurrences à peine), les mots avec lesquels il évolue changent. Pour commencer, on notifie la disparition du lemme *tuus*, même si *suus* demeure très présent (35 cooccurrences). On remarque aussi la présence de nombreux verbes dans son environnement proche à l’instar de *facio* (34 apparitions), *ueneo* (16), *mitto* (13) ou encore *inuenio* (9). Pour le reste, la paire *seruus* (10) et *dominus* (31) marque l’unique présence de substantifs récurrents observable dépassant la dizaine de cooccurrences. En outre, on peut également mentionner les lemmes *tempus* (8), *domina* (8) et *fructus* (7) dont la présence, bien que légèrement moins fréquente, reste néanmoins significative à cette échelle. On retrouve comme autre adjectif qualificatif le lemme *christus* (16), mais également le mot *amen* (8). Ce dernier est employé au sein d’une structure syntaxique régulière que l’on retrouve sporadiquement dans les différents textes qui composent le Nouveau Testament sous forme d’apposition, à savoir, par exemple : *Amen. Paulus seruus Dei, Apostolus autem Iesu Christi*²¹².

Nonobstant la différence totale d’occurrences, on constate aussi un léger glissement des verbes qui flanquent *seruus*. Bien sûr, il reste difficile de définir son sens si l’on ignore les mots avec lesquels il est effectivement en interaction. Quelques lignes de code permettent d’isoler les verbes qui peuvent être employés avec *seruus*. À titre de comparaison, l’Ancien Testament déploie près de quatre fois plus de verbes que le Nouveau dans les alentours de *seruus* au nominatif, soit 170 contre 40. Le lexique est également très différent. Un nombre considérable d’hapax soutiennent *seruus* et s’accordent avec lui. Parmi les plus fréquents de l’AT, on retrouve *sum*, *aio*, *dico*, *audio* ou encore *oro*, mais mis à part *dico* qui trouve aussi sa place parmi les 5 verbes les plus fréquents, aucun autre ne s’y retrouve vraiment. Par conséquent, il est légitime de conclure que la forme nominative au singulier connaît une légère modification, sinon de sens, du moins d’environnement textuel entre l’Ancien et le Nouveau Testament.

Au-delà des lemmes les plus fréquents, TXM indique aussi les jointoiements les plus pertinents, grâce auxquels d’autres lemmes se distinguent. *Dominus* reste en tête du classement, suivi de près par *suus* et par *seruus*, mais des mots comme *uocatus*, *auricula*, *obedio* ou *paterfamilias*, bien que plus discrets dans le corpus, émergent grâce à cet indicateur.

Le ton a également changé. Finis les déterminants possessifs à la deuxième personne du singulier, le style n’est absolument pas similaire. Néanmoins, comme pour l’Ancien Testament, on retrouve une nouvelle fois l’expression d’une binarité quelque peu oxymo-

212 Deuxième épître à Timothée.

rique au travers de la paire *seruus/dominus*. Tout cela étant dit, il reste difficile de donner une définition de *seruus* à partir de ses collocations et cooccurents au sein de la Bible. Tout ce qu'il est possible d'affirmer, c'est que mises à part quelques singularités comparables que l'on retrouve dans les documents de la pratique, le lexique de cooccurrences se profile comme étant drastiquement divergent. La forme même de la Bible ne s'apparente en rien aux standards des documents diplomatiques. Sans structure répétitive, il devient logique de voir jaillir un tel nombre d'hapax.

L'essentiel demeure que, même sans définition claire, il est possible d'établir le constat suivant : l'emploi du lemme *seruus* diverge nettement entre l'Ancien et le Nouveau Testament. Cela confirme en un sens l'hypothèse de pluralité sémantique précédemment évoquée. Attention toutefois à ne pas oublier les limites de cette comparaison sémantique. En effet, la lecture de ces résultats doit prendre en compte que la taille des échantillons sur lesquels se base la comparaison est inégale, à cause de la baisse d'occurrences au sein du Nouveau Testament, soit cinq fois moins d'occurrences pour seulement trois fois moins de mots (548315 contre seulement 182342).

4 Au coeur de l'acte diplomatique : les types de documents et la formule d'humilité

4.1 Les différences sémantiques au travers de la typologie : chartes, bulles, lettres et notices.

Comme matériaux de base, les *CEMA* disposent des chartes médiévales de tous types, allant de l'acte public²¹³ à la bulle pontificale²¹⁴ en passant par des notices²¹⁵ et des diplômes²¹⁶. Bien qu'une grande majorité de documents ne s'est pas encore vu attribuer de genre, ces chartes représentent l'essentiel des typologies à elles seules. La répartition des formes de *seruus* varient inexorablement selon le type de document dans lequel il se trouve. En cumulé, entre le 9^e et le 13^e siècle, on comptabilise 24760 occurrences de *seruus*, tous genres confondus. Mais si l'on se penche individuellement sur chacun, des différences apparaissent. Lettres et bulles proviennent toutes deux de la chancellerie pon-

213 Toutes les définitions suivantes sont issues du livre de Olivier Guyotjeannin et de Jacques Pycke, *Diplomatique médiévale* (atelier du médiéviste), p.25. Les actes publics sont « des actes émanés d'une autorité publique (pape, souverain, prince territorial, autorité disposant d'une juridiction réputée publique), agissant en vertu de cette autorité ».

214 « une bulle est, au sens large, un acte émanant du pape, et scellé du sceau rond (bulle de plomb) qui caractérise la chancellerie pontificale ».

215 Une notice n'est pas un acte à proprement parler car elle n'a pas de valeur juridique. On y consigne la substance d'une action ou d'un fait juridique pour en conserver la mémoire, généralement sous forme objective. Il s'agit d'un aide mémoire écrit *a posteriori*.

216 « Un diplôme est un acte émané d'un souverain ».

tificale et dénombrent 21673 occurrences. Pour le reste, les diplômes en comptent 1572, les chartes 1298 et les notices à peine 217. Au total, cela représente les deux tiers des occurrences totales répertoriées, pour lesquelles les actes émanant de l'autorité du pape sont on ne peut plus prédominants. Après l'élagage de la formule *seruus seruorum dei*, on descend à seulement 748 occurrences, soit 3% des occurrences totales. Autrement dit, dans les documents pontificaux, cette formule compose à elle seule l'absolue majorité des apparitions de *seruus*. Une grosse distinction est à opérer entre les genres, car en fonction de l'autorité de laquelle émane une charte, les implications sur la sémantique varient.

4.2 Histoire et signification de *Seruus seruorum dei*

Tout au long de ce travail, il a été question d'étudier le sens de *seruus* en faisant abstraction de son occurrence principale. La formule d'humilité de *seruus seruorum dei* est en effet essentielle dans les documents diplomatiques, à tel point qu'il est inconcevable de la laisser indéfiniment de côté. Dans l'absolu, cette titulature est propre à l'usage pontifical, mais une question se pose alors : d'autres auteurs la récupéraient-ils dans leurs propres documents ? Certains évêques, princes ou autres figures d'autorité s'arrogeaient-ils le droit de se présenter selon les mêmes termes que le pape ? En outre, qu'en est-il de l'évolution de la formule, de ses occurrences ? Pour quelles raisons fluctuent-elles ? Les *CEMA* sont, de première intuition, opaques à cette problématique.

Cette dernière section va chercher à comprendre l'histoire de cette formule d'humilité, d'en saisir la symbolique théologique et l'implication qu'elle a dans la perception du monde médiéval. À cette fin, la philosophie et la sociologie seront d'aussi bons alliés que la linguistique et l'informatique. La première partie sera expérimentale, dans le sens où elle abordera ces questions de préambule précédemment formulées. Il y sera question des figures d'autorité se présentant sous cette appellation, et de l'évolution du nombre d'occurrences à travers le temps. Puis, après un bref explicatif de l'origine et de la symbolique de l'expression, nous nous pencherons sur l'opposition quasi binaire qui existe entre le mot *seruus* et *deus* (ou *dominus*) dans la conception médiévale. Que cela révèle-t-il sur la conception du pouvoir, sur la légitimation des institutions ecclésiastiques ?

Pour éclaircir toutes ces questions et s'offrir quelques pistes viables, il est nécessaire de passer par la programmation. Dans un premier temps donc, attardons-nous sur les noms que l'on retrouve joints à *seruus seruorum dei*. Une recherche simple formulée en langage *CQL* dans les *CEMA* en retourne 13497 occurrences. À partir de ce tableau importé, il est possible de séparer le bon grain de l'ivraie et de mettre à l'écart la majorité des mots non pertinents, limitant ainsi la liste des collocations. Ne restent alors que les trois mots précédant l'expression pour chaque texte dans lequel l'expression apparaît. La fonction

Counter employée pour compter les occurrences de chaque mot, jumelée à la méthode *most_common()* en python retourne donc des dictionnaires listant tous les noms. La seule limite de cette méthode provient des documents diplomatiques eux-mêmes. Lorsque les papes se présentent, ils le font par leur nom, sans précision supplémentaire. Plusieurs papes du même nom se présenteront ainsi de la même manière, à quelques exceptions près. Dès lors, chaque siècle sera présenté indépendamment l'un de l'autre pour minimiser la confusion.

Au 9^e siècle, quatre actes sortent du lot. En 812, un certain Adalfrid se présente aussi comme un *seruus seruorum dei*, sans mention d'un statut quelconque. Il n'est pas évêque, mais son nom est mentionné au sein du préambule de la charte, il en est l'auteur. Quarante années plus tard, un certain Honoratus accompagne aussi son nom de la formule, mais cette fois en fin d'acte, comme formule de validation²¹⁷. Ainsi, Honoratus est probablement juste un témoin de la transaction effectuée par l'évêque Bon de Sisteron avec l'évêque d'Apt. Une autre mention d'un certain Agapitus en 872²¹⁸ peut être écartée. Il s'agit là d'un faux, faisant référence au pape Agapet Ier, pape du 7^e siècle ou Agapet II du 10^e siècle. Enfin, il se peut que l'occurrence liée à un certain Uictor soit également anachronique voire fausse et fasse référence au pape Victor II, du 11^e siècle. Dans l'ensemble, au 9^e siècle, peu de personnes se présentent comme des serviteurs des serviteurs de Dieu en dehors des papes.

Par la suite, au 10^e siècle, d'autres personnes se sont présentées de la sorte, à commencer par le comte Acfredus d'Aquitaine en 926 et 927. D'après É. Magnou-Nortier, elles seraient en partie fausses et auraient été rédigées plus tardivement²¹⁹, hypothèse nuancée par Philippe Buc²²⁰. Quoi qu'il en soit, seule la date est remise en question dans ces articles, pas la légitimité de l'auteur. Il reste toutefois interpellant de voir un laïc employer une telle formule. Pourrait-on y voir la volonté d'Acfred, lui qui médite tellement sur le salut de son âme dans la charte, de se le voir accorder par un comportement humble et digne ? Plus tard, en 964 et en 971, c'est le nom Petrus qui ressort, mais cela ne fait *in fine* référence qu'à Saint Pierre. Hormis donc cette charte d'Acfred d'Aquitaine, aucune autre forme d'autorité, aucun autre auteur ne se présente comme le fait le pape.

Tout comme pour les deux siècles précédents, peu d'auteurs se risquent ou pensent à se présenter sous les traits de *seruus seruorum dei*, cette tendance reste marginale. Dans

217 0852_apt_70895

218 0872_stmarcelles_7986

219 MAGNOU-NORTIER, E., *Contribution à l'étude des documents falsifiés : l'acte de fondation du monastère de Sauxillanges*, dans *Chaiers de civilisation médiévale*, t.21, 1978, p.313-328.

220 BUC, P., *Les débuts de Sauxillanges : à propos d'un acte de 927*, dans *Bibliothèque de l'école des chartes*, 1998, t.156, p.545.

les premières années du 11^e siècle, l'archevêque de Ravenne, Frédéric de Saxe, emploie également ce titre dans un acte où il concède des biens situés dans la massa Sancti Illari à des laïcs. Toujours à Ravenne, l'archevêque Arnold de Saxe, une dizaine d'années seulement après son prédécesseur et dans une charte à même finalité, va également se présenter de la sorte, tout comme Gebehardus premier à sa suite, en 1030²²¹. C'est pour ainsi dire la première fois où l'on assiste à une répétition au sein d'une même institution. À l'opposé d'un tel rang dans la hiérarchie ecclésiale, en 1029, le ministre de l'église anglaise de Canterbury en Angleterre, Egelnothus, se présente humblement comme un *peccator servus servorum dei et minister aeclesiae Christi Anglorum*²²². À la fin du siècle, un autre clerc anglais (Wistanus Wigornensis), se présentera comme *indignus servus servorum dei*, accentuant davantage la formule d'humilité.

Contre-intuitivement, la multiplication de la documentation au 12^e siècle n'est pas synonyme d'une recrudescence d'emplois non-pontificaux de *servus servorum dei*. Il n'y a que deux situations dans lesquelles cela se produit, comme une goutte d'eau dans l'océan. Le 13^e siècle est encore plus désertique : il n'y a pas une seule divergence par rapport aux listes papales. On pourrait y reconnaître l'élan d'uniformisation et de spécification de la production écrite mêlée à la croissance de la documentation officielle²²³. En conclusion, en dehors de la machine pontificale, aucune chancellerie, aucun autre acteur ne se sert de façon répétée d'une telle formule. Elle reste donc l'apanage du pape. Même le comportement étrange des archevêques de Ravenne prend fin dans le courant du 11^e siècle, aussi vite qu'il était apparu.

4.2.1 Histoire de *servus servorum Dei*

La toute première mention de cette expression dans le corpus date de la toute fin du 6^e siècle et fut apposée au nom du pape Grégoire I^{er}. Une grande partie de l'historiographie existante à ce sujet (brèves études et entrées d'encyclopédies)²²⁴ accepte cette assertion. Grégoire le Grand est le premier à hériter de cette formule d'humilité. La Vita Gregorii de Jonh Deacon, et plus particulièrement le second livre, composée près d'un siècle plus tard, vers 885, est le premier ouvrage à faire état de la formule annexée à son nom. D'après elle, peu après avoir été élu pape, fort de son caractère humble et universel, il s'est affublé de

221 1002_ravenne-eveques_52901, 1017_ravenne-eveques_52891, 1030_artem-varia_49423

222 1029_canterbury-christ_114560

223 JAMME, A., *Écrire pour le pape du XI^e au XIV^e siècle. Formes et problèmes.*, dans *Mélanges de l'École française de Rome - Moyen Âge*, t.128/1, 2016, p.14.

224 D'après CHAI-ELSHOLZ, R., : Schmitz K., Ursprung und Geschichte der Devotionsformeln bis zu ihrer Aufnahme in die fränkische Königsurkunde, Stuttgart, Ferdinand Enke, 1913, pp. 120-139 ; Delehaye H., *Servus servorum Dei*, in *Strena Buliciana. Commentationes gratulatoriae Francisco Buli ob XV vitae lustra feliciter peracta oblatae a discipulis et amicis a.d. IV non. oct. 1921, Zagrebiae, Aspalathi*, 1924, pp. 377-378 ; Leclercq H., "Servus servorum Dei", *Dictionnaire d'archéologie chrétienne et de liturgie*, 15 (1950), col. 1360-1363 ; Levillain L., "Servus servorum Dei", *Le Moyen Âge*, 1 (1930), pp. 5-7.

cette formule qui par la suite fut transmise à tous ses successeurs²²⁵. L'humilité est vraiment le maître-mot de ce second livre. Elle est illustrée tant sur le plan du langage, avec *seruus seruorum dei*, que d'un point de vue de l'apparence, par ses vêtements modestes. Ces reliques vestimentaires sont donc une démonstration visible et palpable, tangible, de sa vertu. Il y a donc une connexion évidente entre la notion d'humilité et la formule.

Le lien entre l'humilité et le service n'est pas neuf pour autant. Depuis les premiers temps de l'Église, les officiels empruntaient déjà les termes de *seruus dei/seruus Christi*, attribués précédemment dans l'Ancien Testament à de nombreuses personnalités majeures²²⁶. La volonté de s'émanciper du sens péjoratif hérité de l'Antiquité, marqué par l'oppression, la dépendance et la servitude se présente dès les premiers temps de la religion chrétienne à travers ses textes fondateurs. La véritable rupture naît au travers du Nouveau Testament et de la scène du lavement des pieds (Jn 13 : 12-15) où Jésus exemplifie le sens du nouveau paradigme du service. Il en inverse le sens initial, passant d'un modèle hiérarchique vertical à un modèle horizontal de commun service. Malgré tout, l'expression *seruus seruorum dei* n'apparaît pas dans la Bible. Elle n'est *in fine* que le catalyseur, l'outil d'inspiration d'une telle démarche. *Seruus* n'est pas un cas isolé de renouvellement sémantique entre l'époque antique des païens et l'époque médiévale, catholique. Dans de vieilles éditions de la Bible (*Vetus Latina*), ou par la plume d'auteurs comme Tertullien (3^e siècle), le terme *pontifex* change de sens. Ce dernier désigne par ce mot non plus l'empereur romain, mais le Christ lui-même²²⁷. Il existe donc, si ce n'est une volonté, du moins le résultat d'un changement sémantique qui atteste d'une évolution de la réalité sociale.

Dans le courant des 7^e et 8^e siècles, la formule croît lentement mais son importance ne peut être réellement mesurée au vu de la documentation lacunaire. Plus les siècles avancent, plus les documents pullulent, plus la formule apparaît. Ce n'est vraiment qu'à partir de la moitié du 11^e siècle et du pontificat de Léon IX qu'elle va réellement s'envoler. Précisons tout de même que l'expression en soi ne prend pas davantage de poids au fur et à mesure des siècles, mais ses occurrences sont juste intrinsèquement liées à la masse documentaire disponible. Or, le tournant documentaire provient lors de la réorganisation de l'Église catholique au cours de ce que l'on nomme aujourd'hui la « réforme grégorienne », bien qu'elle fut initiée déjà par Léon IX. Les mentions de *seruus seruorum dei* connaissent un véritable premier pic en 1067, où 119 chartes la renseignent. C'est bien là une anomalie créant un précédant dans l'histoire, jamais autant de chartes n'avaient compris la formule.

225 CHAI-ELSHOLZ, R., *Servus servorum Dei in Context : Historical and Literary Intertextuality to c. 875*, dans FRANK, T. et RANDO, D., *Ernst Kantorowicz 1895-1963* : *Storia politica come scienza culturale*, p.75

226 DON N., H.-Jr, *Servants of the servants : a biblical Theology of Leadership*, p.7.

227 VAN HAEPEREN, F., *Des pontifes païens aux pontifes chrétiens. Transformation d'un titre : entre pouvoirs et représentations*, dans *Revue belge de philologie et d'histoire*, t.81, fasc.1, 2003, p.142.

Et il faudra attendre douze ans avant de voir rejaillir un échantillon de taille presque égale, sous le pontificat de Grégoire VII (90). Clément II a aussi, en 1093, atteint des sommets avec 238 bulles/lettres. L'un des constats que l'on peut poser, c'est que les dates au plus haut taux de profusion ne sont jamais des années charnières entre deux pontificats. Elles sont plutôt le symptôme d'une législature en bonne santé²²⁸. Officiellement, la formule fut ajoutée très tardivement aux titulatures pontificales officielles. Il a fallu attendre Vatican II, sous Paul VI, en 1962, soit plus d'un millénaire après ses premières attestations dans la fonction²²⁹, pour officialiser sa présence.

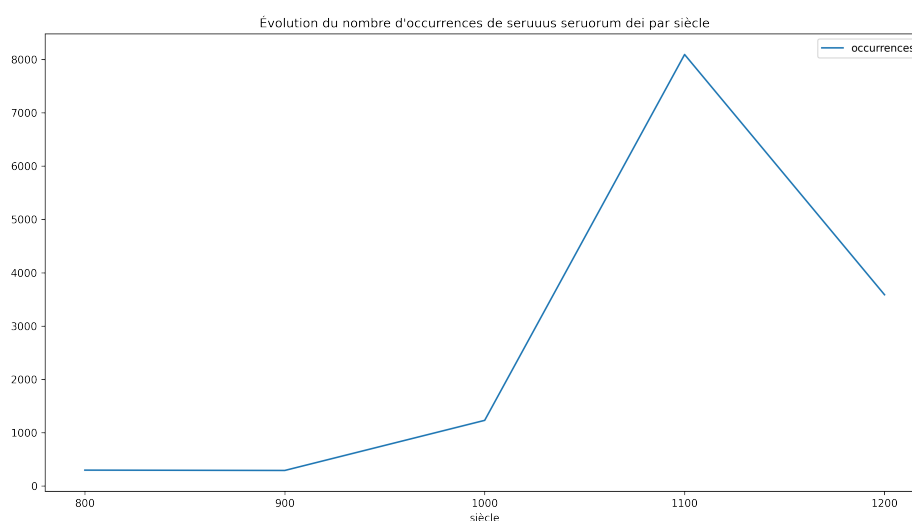


Figure 50 – *Occurrences de seruus servorum dei par siècle.*

4.2.2 L'importance du titre d'évêque de Rome

Parmi les chartes issues de la chancellerie pontificale, celles dans lesquelles apparaissent la formule d'humilité appartiennent aux constitutions apostoliques. Elles sont « toujours rédigé[e]s en latin, le plus souvent scellées, [...] et portent en tête le nom du pape, suivi de la mention « pape » (PP) ou « évêque » (episcopus) [...]. La deuxième ligne porte la mention « serviteurs des serviteurs de Dieu » (servus servorum Dei) »²³⁰.

La gestion de l'évêché de Rome prédomine sur la fonction de pape, si bien que son

228 Les années les plus prolifiques en bulles et lettres sont les suivantes : 1051 (3e année de pontificat de Léon IX) ; 1067 (6e année de pontificat d'Alexandre II) ; 1079 (7e année de pontificat de Grégoire VII) ; 1093 (6e année de pontificat d'Urbain II) ; 1108 (10e année de pontificat de Pascal II) ; 1121 (3e année de pontificat de Calixte II) ; 1127 (4e année de pontificat d'Honorius II) ; 1136 (7e année de pontificat d'Innocent II) ; 1149 (5e année de pontificat d'Eugène III) ; 1156 (4e année de pontificat d'Adrien IV) ; 1163 (6e année de pontificat d'Alexandre III) ;...

229 DICKÈS, C., *Dictionnaire du Vatican*, p.480.

230 *Idem*.

prestige provient presque exclusivement de son titre d'évêque. L'une des explications résiderait dans la croyance en la présence des tombeaux de Saint-Pierre et Saint-Paul dans la ville²³¹. Le pape tire de cette filiation une primauté vis-à-vis des autres évêques. Dans les premiers siècles, le titulaire du siège romain ne parvient pas toujours à imposer son point de vue, pas avant le 4^e siècle en tout cas. Le concile de Nicée change la donne et attribue au pape la primauté théorique, autrement dit, la supériorité sur l'ensemble des clercs et sur toute l'Église²³². Jusqu'au 9^e siècle, le titulaire du siège romain était élu par le peuple et le clergé de la ville de Rome. Marin I^{er}, en 882, sera le premier à être élu sur le siège de Pierre alors qu'il détenait déjà des charges en tant qu'évêque auparavant. Par la suite, le pape devient évêque parce qu'il est choisi pour occuper le trône de Pierre²³³.

Il existe un grand nombre d'expressions pour désigner le chef de l'Église catholique, au-delà des deux susmentionnées. Chacune d'entre elle renvoie à une réalité juridique ou historique. En 2012, le pape Benoît XVI en retenait huit : « Évêque de Rome, vicaire de Jésus-Christ, successeur du prince des apôtres, souverain pontife de l'Église universelle, primat d'Italie, archevêque et métropolitain de la province romaine, souverain de l'État de la cité du Vatican, serviteur des serviteurs de Dieu »²³⁴. Dans les faits, entre le 9^e et le 13^e siècle, sitôt le pape se présente comme évêque [de Rome], sitôt il accompagne cette mention de celle du *servus servorum dei*, ce qui peut paraître quelque peu paradoxal en un sens, voire presque oxymorique. Il se pare ainsi du plus haut titre de l'Église catholique, mais postule également qu'il est le serviteur de tout homme. Au sein de la hiérarchie structurale inhérente à l'institution ecclésiale, il est au sommet, mais sa place parmi les hommes reste modeste.

Cette perception de la gouvernementalité des hommes où le pouvoir est renversé, inversé même, n'est en apparence pas la plus à même à structurer la société et à stabiliser l'ordre social. Pourtant, il s'agit paradoxalement du modèle le plus pérenne et le moins critiqué du Moyen Âge. Basé sur la doctrine du « les derniers seront les premiers », ce système laisse perplexe quant à l'explication de son efficacité²³⁵. Cela amène tout naturellement à s'interroger sur cette double implication, à savoir le principe d'une gouvernementalité à la hiérarchie intervertie et à la place qu'occupe la notion de service dans la société médiévale. Les deux sont étroitement imbriquées.

Tout fidèle, tout croyant, peut être désigné par l'étiquette de *servus dei*. Le pape s'envisage comme le serviteur de ces serviteurs. Comme évoqué précédemment, son titre se

231 <https://fr-academic.com/dic.nsf/frwiki/1531755>

232 BASDEVANT-GAUDEMET, B., *Le pape évêque de Rome*, dans *Pouvoirs*, 2017, n°162, p.115.

233 *Ibid*, p.116.

234 DICKÈS, C., *Dictionnaire du Vatican*, p.1365.

235 DALARUN, J., *Gouverner c'est servir*, 2004, p.15.

veut l'expression de l'humilité, mais il le fait surtout par le renvoi au service. À l'image de la souveraineté du Christ qui « prend visage et s'exerce dans le service. [Son] autorité n'est pas une autorité qui asservit, mais une autorité qui se met au service d'autrui et l'entoure de sollicitude »²³⁶, le pape cherche non pas à diriger mais à servir. Car c'est ainsi que les évangiles le tournent également. Dans l'évangile de Saint Matthieu (20-28) : « C'est ainsi que le Fils de l'Homme n'est pas venu pour être servi, mais pour servir et donner sa vie en rançon pour une multitude »²³⁷. Et dans les Admonitions de Saint-François d'Assise :

« Ceux qui ont été établis au-dessus des autres, qu'ils se glorifient de cette prélature autant que s'ils avaient été députés à l'office de laver les pieds des frères. Et s'ils se troublent plus de la perte de la prélature que de la perte de l'office de laver les pieds, c'est comme s'ils amassaient un magot au péril de leur âme ». ²³⁸

Pour citer un exemple de l'inversion du pouvoir médiéval, prenons l'essor des ordres mendiants (franciscains, dominicains, carmes, guillelmites,...) à partir du début du 13^e siècle. Dalarun retrace dans son livre "Gouverner c'est servir" l'histoire de l'ordre franciscain, autant pour les hommes (figure de Saint-François), que pour les femmes (figure de Sainte-Claire). Ceci n'est pas sujet de digression, mais d'exemple et d'explication. Dans ce passage, François rappelle les bases de l'autorité chrétienne. Par cela, il établit le constat d'une conception pervertie de laquelle il faut s'extraire. L'essor des groupes mendiants au 13^e siècle provient d'un désir profond à un retour aux valeurs apostoliques et à une relation saine avec les valeurs évangéliques. La notion de service y est absolument centrale. Claire incarne cette volonté, cet assainissement, cette vertu :

« Parmi les servantes du Christ, [elle voulut] plus volontiers servir qu'être servie. Mais, contrainte par le bienheureux François, elle finit par accepter le gouvernement des dames : de là naît en son coeur la crainte, non pas l'arrogance ; et la liberté n'augmente pas, mais la servitude. Car plus on la voit désormais élevée par l'apparence de quelque supériorité, plus on la trouve vile dans l'estime de soi, plus prompte au service, plus méprisable dans l'ap-prêt »²³⁹.

Le fait qu'elle refuse la charge et se voit contrainte de l'adopter témoigne d'une interprétation du pouvoir et de l'autorité qui ne s'accorde pas avec la théorie des évangiles. La refuser, c'est déjà la comprendre comme n'étant pas l'incarnation de ces valeurs. Il y a donc un écart entre l'idée sous-jacente de l'autorité, qui doit suivre la vision représentée

236 ZUMSTEIN, J., *Le lavement des pieds (Jean 13, 1-20) : un exemple de la conception Johannique du pouvoir*, dans *Revue de Théologie et de Philosophie*, vol.132, n°4, 2000, p.356.

237 Bible de Jérusalem, p.938; https://livres-mystiques.com/partieTEXTES/Bibliques/Bible_de_Jerusalem.pdf

238 DALARUN, J., *Gouverner c'est servir*, 2004, p.71.

239 DALARUN, J., *Gouverner c'est servir*, 2004, p.28.

par le Christ, et la réalité des choses. Le clergé régulier du 13^e siècle n'est pas le seul à manifester un rejet de la conception d'autorité.

Au sein du clergé séculier, prendre les offices, c'était déjà participer au sauvetage des âmes des fidèles, c'était une manière de servir la communauté. Mais comme toute entité structurée, il est nécessaire de hiérarchiser les fonctions pour répondre à des besoins de logistique et d'efficacité. Cela étant, chaque échelon gagne en autorité et en pouvoir, ce qui est contraire à la doctrine d'abnégation et de service défendue par l'institution. En se présentant non pas sous un simple titre, mais sous l'appellation de « serviteurs de Dieu », ou « serviteur des serviteurs de Dieu », les évêques, archevêques et papes légitiment leur autorité par le service qu'ils rendent à la communauté. *Seruus seruorum dei* définit donc les qualités propres à l'individu qui s'en revendique, mais peut être également perçu comme une formule de légitimation de la structure institutionnelle. Lorsque Grégoire le Grand emploie pour la première fois cette formule, la primauté de Rome sur les autres diocèses n'est pas encore reconnue²⁴⁰, les Églises régionales adressent peu de demandes à Rome. Pour Charlemagne, le pape a rôle d'informateur, il enseigne sur la tradition romaine sans pour autant gouverner l'Église. Que les évêques Romains aient été les seuls à l'employer au cours des siècles peut laisser penser qu'elle est l'incarnation de la volonté romaine de devenir le centre, la charnière d'une Église universelle.

D'un point de vue global, moins les papes avaient d'influence, moins ils produisaient de documents. Au moins ils produisaient de documents, au moins apparaissait la formule. Le « désert » d'occurrence des 9^e, 10^e et début du 11^e siècles est donc le témoin, non pas d'une chancellerie pontificale malade et inefficace, mais de l'autorité incomplète ou vacillante du pape.

L'association *episcopus seruus seruorum dei* fait donc naître un premier oxymore par la succession des titres latins, que l'on peut interpréter comme la jonction d'un statut (d'une position sociale) et d'une vertu. Le fait de lier les deux revient à exprimer l'opposition de deux situations qui sont de prime abord très distinctes. Comment peut-on à la fois dominer la hiérarchie ecclésiale et simultanément paraître pour le serviteur de tout un chacun ? La formule d'humilité explique à elle seule en quoi il n'y a guère d'opposition. Car si l'autorité est synonyme de service, l'évêque n'est *in fine* qu'un homme vertueux au service de tous, qui ne les domine pas mais qui les sert.

Le contraste entre *seruus* et *episcopus* n'est pas le seul qui puisse être pointé du doigt. Dans la formule *seruus seruorum dei*, *seruus* et *deus* (qui peut aussi être *dominus* ou *Christus*) symbolisent une dichotomie plus fondamentale encore. L'anthropologie moderne a

240 Elle a été décrétée dès le 4^e siècle, sans pour autant persuader les autres décideurs occidentaux.

mis sur pieds des outils pour schématiser ce type de mode de pensées, à l'image de Philippe Descola dans son ouvrage *Par-delà nature et culture* : les matrices d'analogie générale. En histoire, ce concept a notamment été récupéré par Anita Guerreau-Jalabert, qui résume toute dichotomie (homme-femme, homme-Dieu, ordre-anomie,...) par un lien constant avec le duo *caro-spiritus*²⁴¹. Cette paire s'applique à bon nombre de réalités et joue un rôle essentiel dans la conception des rapports sociaux. Dans la conception des médiévaux, tout ce qui appartient à l'univers "se laisse répartir entre ces deux pôles, ou plutôt se caractérise par une modalité particulière, positive ou négative, d'articulation du corporel et du spirituel"²⁴²

Comme Anita Guerreau le souligne, « il apparaît en tout cas assez clair que c'est entre saint Paul et saint Augustin que le christianisme a construit, sur la base d'une représentation de l'homme, les fondements d'une matrice d'analogie généralisée qui a servi de cadre à la pensée occidentale durant plus d'un millénaire. »²⁴³. La paire *caro-spiritu*, et toutes les autres aussi, sont pensées comme deux valeurs distinctes, mais aussi et avant tout complémentaires, unies. Il ne s'agit pas là d'un dualisme, mais bien d'une manière de penser l'unité de l'homme, et par extension, l'unité de l'*ecclesia* et du monde. Le sens des mots n'est pas à trouver dans les objets de la matrice, mais dans le rapport qui les unit. *Seruus*, dans son lien avec *dominus*, où *deus* est du côté de *caro* et se teinte de la charge morale négative attribuée à la chair face aux vertus de l'esprit que l'on associe à Dieu. Par association, le *dominus* détient les valeurs du *spiritus*.

Ainsi en est-il en tout cas dans l'interprétation analogique de *seruus* prise dans le contexte de la formule *seruus seruorum dei* et, *in fine*, pour toutes les occurrences de *seruus* dans sa dichotomie avec Dieu. Comment interpréter alors la relation entre le *seruus* que l'on décrit dans le coeur des actes, avec le *dominus*/prince/seigneur ? Ces *serui*-là s'inscrivent aussi dans une matrice binaire. Tout comme le terme *seruus*, *dominus* possède ce double sens qui réfère d'une part au Seigneur (Jésus-Christ) et en même temps à un statut, celui de seigneur. Difficile de croire que l'analogie puisse s'inverser, elle est donc probablement transposée à cette seconde réalité. Les liens entre *seruus* et *caro* découlent de cette perception analogique, et justifient alors leur servitude par leur imperfection. Pourquoi se plaindre de son sort lorsque l'on sait que l'on n'est pas vertueux ?

Enfin, être *seruus* à l'époque médiévale, c'est être fidèle, c'est adhérer à la croyance, c'est choisir d'obéir. Son obéissance est le fruit de sa volonté, de son libre-arbitre. Mais elle n'est pas connotée négativement, l'obéissance spirituelle est le choix d'écouter et

241 GUERREAU-JALABERT, A., *Occident médiéval et pensée analogique : le sens de spiritus et caro*, dans GENET, J.-P., *La légitimité implicite*, p.457-476.

242 BASCHET, J., *La civilisation féodale : de l'an mil à la colonisation de l'Amérique*, p.598.

243 *Idem*.

de se confier à Dieu. « Progresser dans l'obéissance, c'est toujours à nouveau choisir de mettre Dieu en premier »²⁴⁴. Un *seruus* est donc lié au *dominus* par une obéissance qui est avant tout morale et chrétienne. Pour les Chrétiens, d'ailleurs, l'obéissance rend libre. Tout comme le peuple d'Israël qui, après sa sortie de l'esclavage en Égypte, a été « Libéré pour servir »²⁴⁵, les *serui* ne désirent que la vraie liberté, celle du choix de l'obéissance, celle qui cherche à construire la communion, la solidarité, la liberté solidaire, pour le bien et la vérité. Le *seruus* ne se voit pas imposer l'obéissance, il la vit comme un choix intégré et assumé.

244 HOEGGER, M., *Une obéissance qui libère*, p.36.

245 *Ibid*, p.15.

Conclusion

Pour mener une recherche, il faut poser des choix. La méthodologie dépend du chercheur et s'exprime donc explicitement, parfois implicitement, au travers des corpus et des décisions relatives à son exploitation. Choisir, c'est renoncer, et il a fallu renoncer à de nombreuses pistes en raison d'un timing trop serré. Tout commence par le choix des documents, réalisé de manière réfléchie, mais qui reste toutefois empreint d'une part d'arbitraire. *A posteriori*, une étude diachronique d'une telle ampleur n'était peut-être pas la meilleure idée, en cela qu'elle ne permet pas de se plonger en profondeur dans les réalités synchroniques de chaque siècle. Mais, à l'inverse, une étude si précise n'aurait été pertinente qu'en cas de foisonnement documentaire, ce qui n'est *in fine* le cas que pour les 11^e et 12^e siècles, et dans une moindre mesure, le 13^e siècle également. Tout cela pour dire que nos choix ont amené à certains résultats qui se doivent d'être remis en contexte.

4.3 L'inutilité étymologique et les pistes de la lexicographie

La première décision de ce travail a été de se reposer sur les techniques classiques d'étude sémantique en se penchant sur les définitions préétablies du mot *seruus* dans les dictionnaires d'antan et d'aujourd'hui. Même si des auteurs renommés comme Bruno Bon et Anita Guerreau-Jalabert le déconseillaient fortement, nous avons tout de même opté pour une brève introduction étymologique. Bien entendu, nous étions conscients des lacunes d'une telle démarche, mais elle demeurait souhaitable dans la mesure où elle permettait d'établir un premier jalon temporel. Les multiples analyses menées par la suite ont conforté cette idée que l'étymologie n'est pas un outil adéquat dans l'étude sémantique. Pour cause, qu'il s'agisse de l'étude par cooccurrence vaste ou précise, d'une analyse matricielle ou d'une comparaison sémantique avec le lexique biblique, aucun indice n'est venu conforter l'idée que *seruus* ait pu se rapprocher de la notion de *gardien*.

L'étude lexicographique était, quant à elle, déjà plus justifiée, bien qu'incomplète. La liste des manuscrits employés est loin de l'exhaustivité, mais offre néanmoins un profil précis de la perception des médiévaux. Pour beaucoup, ils se référaient à cette notion étymologique, sans en compléter la définition. Une véritable rupture qu'il fait bon de mentionner est la transition entre les lexicographies médiévales et les premiers dictionnaires franco-latins. Dans ces derniers, on attribue à *seruus* l'étiquette de *serf*, qui lui est accolée subrepticement. Le concept de servage est en lui-même discutable et découle d'un prisme biaisé des premiers savants des Lumières, d'après Jérôme Baschet. Pour construire une identité propre à ce mouvement novateur des idées, cela impliquait de se définir en rapport avec quelque chose et « pour la bourgeoisie, le Moyen-Âge constitu[ait] le parfait repoussoir », comme le souligne Baschet. Mais le cœur du problème repose sur ce que Guerreau appelle « la double fracture conceptuelle », à savoir l'emploi de concepts pos-

térieurs à l'époque médiévale pour analyser sa dynamique et son évolution. C'est à cause de l'éclatement des notions d'*ecclesia* et de *dominium*, remplacés par les concepts d'économie (Adam Smith) et de religion (Jean-Jacques Rousseau), que la réalité nous échappe et que la discipline, au mieux stagne et au pire s'éloigne de la vérité²⁴⁶.

Ce que nous disent les textes modernes, à la lumière de la théorie de Bashet, c'est qu'on relie les anciens *serui* à une servitude unique qui au 16^e siècle pouvait même se rapprocher de la notion d'esclavage. Les définitions sans nuance de l'époque moderne s'insèrent bien dans une perspective de simplification et de contradiction avec les fonctionnements médiévaux. Là encore, l'étude était limitée et aurait nécessité davantage de définitions.

En ce qui concerne les dictionnaires contemporains, les résultats sont nettement plus probants et critiques, même si les définitions paraissent éclatées. En fonction du dictionnaire que l'on consulte, le sens prend une teinte quelque peu différente, et, le Du Cange mis à part, aucun ne fait état de la variété d'usages de *seruus* sur le millénaire qui compose le Moyen Âge. Les contemporains sont déjà plus lucides, ou simplement plus loquaces, quant à la liaison intrinsèque de *seruus* avec le domaine spirituel. Il reste cependant possible, comme l'a montré Carl Ginzburg au travers de son ouvrage *Le fromage et les vers*, que la non-mention d'un sens propre à l'*ecclesia* soit en réalité l'exemple d'un "exceptionnel-normal"²⁴⁷. Que la notion soit tellement répandue et implicite que l'idée de définir l'évidence ne traversait pas l'esprit des auteurs médiévaux.

4.4 Quels résultats pour les outils numériques

Cette première analyse lexicographique était instructive, nonobstant ses limites évidentes. L'objet même de ce mémoire était toutefois de prolonger les méthodes traditionnelles et d'opérer une sorte de mise à jour. À cette fin, nous avons attaqué la question du sens de *seruus* par trois axes que nous allons désormais justifier. D'abord, il a été question d'isoler le mot au sein du champ lexical auquel il est habituellement rattaché en comparant les différences dans l'emploi des cas ou au moyen d'indices de similarité entre leurs lexiques de cooccurrences. Puis, de le définir au moyen de ses propres cooccurents. Enfin, en comparant les résultats avec ceux initiés dans l'ouvrage majeur de la période médiévale, à savoir la Bible.

Comme annoncé dans l'introduction, nous tenions à dire qu'il est possible que certains concepts, certaines fonctions aient été mal comprises, voire mal interprétées. Malgré tout, la démarche nous semblait pertinente car même si les résultats en pâtissent, la méthode,

246 BASCHET, J., *La civilisation féodale, de l'an mil à la colonisation de l'Amérique*, 2018, p.28.

247 LE COQ, A., Compte-rendu de la journée d'études *Microhistoire et pratiques historiennes. Échelles, acteurs, formes narratives*, 2012, dans *Memini*, n°15.

elle, s'enrichit.

4.4.1 La différenciation sémantique : une piste caduque ?

La liste des lemmes avec lesquels comparer *seruus* a été établie par la lecture de différents articles ou ouvrages. D'aucuns la qualifieraient de réductrice, car oui, la comparaison ne touche *in fine* que sept lemmes initiaux. Mais il faut prendre en compte deux choses : d'abord, les mots les plus fréquemment cités dans la littérature lorsqu'il est question de la servitude médiévale s'y retrouvent tous (*colonus*, *ancilla*, et *mancipium*), mais aussi et surtout que l'analyse est aisément extensible grâce aux fonctions mises en place.

Qu'en est-il de l'analyse des cas ? Comme expliqué dans le premier chapitre, il est difficile de faire la distinction, dans certaines déclinaisons, entre plusieurs cas. Le degré de précision est donc bridé par les limites que pose le latin lui-même. Mais cela n'empêche pas de mettre en exergue les tendances générales. Ce qu'on peut dire, c'est que *seruus* se distingue, il sort du lot. Sa distribution des cas est sans pareille parmi la liste de lemmes. Ses proportions à l'accusatif pluriel et au génitif pluriel ne trouvent aucun compétiteur. Il est dès lors justifié de dire, avec mesure, bien entendu, que l'emploi spécifique de *seruus* dans les textes témoigne d'un sens particulier.

À cela, on peut ajouter d'autres indicateurs, à commencer par la fréquence d'apparitions dans une seule charte. En moyenne, *ancilla* et *mancipium* sont davantage répétés que *seruus*. La différence est maigre mais notable. Pour le reste, cette approche n'a pas été des plus fructueuses, à l'exception faite de l'indice TF-IDF. Comme les résultats l'indiquent, *seruus* est sensiblement moins structurant de la charte que le sont les autres lemmes. Cela implique que son utilisation est plus formatée, qu'il apparaît, par exemple, dans des formules toutes faites et répétées.

La dernière étape consistait à évaluer la similarité de chacun des lemmes de la liste de base avec *seruus* au moyen de calcul de distances. Pour cette approche, les critiques sont nombreuses. Tout d'abord, l'indice de Jaccard et ses résultats déconcertants pour la seconde moitié du 13^e siècle (une similarité de près de 80%), indique un dysfonctionnement éventuel dans la programmation de la fonction que nous n'avons pas réussi à expliciter, ni à corriger. Cela peut éventuellement remettre en question la validité des autres distances, à l'exception de la fonction personnelle.

Si l'on fait abstraction de cet accroc, les résultats générés sont sans équivoque : *seruus* se rapproche plus de *serua* et d'*ancilla* que des autres. Pour le premier, cela peut s'expliquer une nouvelle fois par la possible mauvaise attribution des formes au datif/ablatif pluriel, et pour le second, il est clair que la formule *seruus et ancilla* a joué un rôle dans

ce rapprochement.

En conclusion, il est possible de tirer de tout cela que *seruus* est un cas à part, qu'il se différencie des autres lemmes, bien que cette différenciation soit difficilement quantifiable. Pour autant, l'hypothèse initiale se valide tout de même : un usage différent et un contexte différent donne au mot un sens différent. Le seul bémol, c'est qu'il est impossible de donner de valeur absolue aux résultats de ces recherches. Par conséquent, malgré les reproches qu'on peut lui faire, la méthode reste intéressante. Tous les résultats ne sont pas pertinents, ni n'apportent même de contribution à la tâche, mais pour autant, globalement, elle est fiable et utile.

4.4.2 La définition par cooccurrence

Une fois passée l'étape de distinction, fallait-il encore essayer de trouver ce sens particulier qui le définit. Les outils mis en place ont été nombreux, à commencer par une séparation des catégories grammaticales, qui, chacune à leur manière, ont apporté des éléments de réponse. Même s'il reste impossible de définir *stricto sensu* le mot *seruus*, le concept est désormais mieux établi.

Tout d'abord, ce que les adjectifs révèlent, c'est l'usage massif et croissant des déterminants possessifs entre le 9^e et le 13^e siècle. Ce que l'on constate d'autant plus, c'est que ces déterminants sont principalement de la première personne, ou de la troisième (*meus*, *noster* et *suus*). Ces cas de figure indiquent dès lors une forme d'appartenance légale, à l'inverse des déterminants à la deuxième personne qui réfèrent quant à eux à un lignage moral, spirituel. D'autres adjectifs, liés à un vocabulaire plus laïc, tendent à s'évincer fortement avec le temps. Rappelons encore que les listes d'adjectifs n'ont pas été analysées dans leur entièreté, et qu'il a fallu procéder à une sélection d'adjectifs à traduire. Il est possible que la traduction et la catégorisation exhaustive des adjectifs auraient modifié les résultats.

Verbes et noms voient émerger la même dichotomie sémantique, une opposition flagrante entre un sens lié à Dieu et un sens plus terre-à-terre. En outre, un autre problème transparaît : une forte augmentation de documentation à partir du 12^e siècle, entraînant une hausse d'occurrences de la formule *seruus seruorum dei*, avec une chute antithétique du nombre d'occurrences de *seruus* au sein du document diplomatique. Avec toutes ces informations en tête, on se doit de comprendre différemment les occurrences de *seruus* au 9^e siècle, et celles des 11^e-13^e siècles. En ce qui concerne la bipolarité de *seruus*, elle n'est intrinsèquement pas étonnante. Mais est-elle possible dans un langage formaté autour du langage biblique ?

4.4.3 Une dichotomie sémantique dans la Bible

Pour les moines, clercs, ou autres scribes auteurs des chartes médiévales, était-il concevable d'écrire *seruus* plusieurs fois dans une même charte tout en lui apposant deux sens distincts ? Pour vérifier cette hypothèse, il faut se plonger dans la Bible. Après quelques recherches, un double-sens apparaît effectivement, même s'il est difficile de le cerner pleinement. En effet, dans un texte littéraire, il y a moins de structure, de formules. Le paysage des cooccurrents est d'autant plus éclaté. Cette différence sémantique s'illustre une nouvelle fois par des emplois différents de verbes, noms et adjectifs dans le giron de *seruus*. Un scribe, au Moyen Âge, était donc déjà au courant de cette polysémie. Mais était-elle pour autant antithétique ?

4.5 Ce que révèle la formule d'humilité

Les théories récentes en anthropologie historique, au travers notamment d'Anita Guerreau-Jalabert, donnent un cadre qui structure la pensée médiévale, qui relativise chaque mot dans son emploi au Moyen Âge. Avec un tel prisme, la perception de *seruus* prend une autre dimension. La formule d'humilité est indicatrice d'un paradoxe peu commun qui emploie deux termes que tout oppose. Tandis que l'un renvoie aux bienfaits de l'esprit, du *spiritus* et détient une certaine supériorité morale, le second baigne dans l'incomplétude, l'imperfection, les défauts. Car le *Dominus* ne peut être que bienveillant et moralement juste, à l'inverse du *seruus* qui est tenté par le vice.

L'analogie est évidente si on l'aborde du point de vue de ce premier pôle. La relation entre Dieu et ses serviteurs, ses fidèles, se complète par cette dichotomie et désigne tous ceux qui ne sont pas Dieu comme ses serviteurs, enclins au péché. Dieu apparaît alors comme un idéal à atteindre, la vertu à imiter. Mais ce vocabulaire n'est pas propre au monde ecclésial, comme nous l'avons constaté. Il est transposé pour aborder des sujets plus terre-à-terre, moins éthérés. Comment faire vivre cette analogie dans une relation entre deux hommes ? Attribue-t-on toujours le lien du *spiritus* au *dominus* ? Si tel est le cas, les *serui* doivent voir en leur seigneur un homme moralement plus juste. Les services qu'ils leur rendent leur sont faits non sous le poids de la contrainte, mais plutôt justifiés par leur supériorité morale.

L'analogie permet de relier ces emplois distincts constatés tout au long de nos recherches. Les *serui* sont des fidèles, qui par opposition aux vertus de leur *dominus*, se plongent dans l'obéissance, non pas servile, mais chrétienne, autrement dit : délibérée. Les différences lexicales renvoient à une réalité simple. Les *serui* servent par cette analogie. Ils acceptent l'autorité de leur *dominus* en leur attribuant les qualités de l'esprit qu'ils n'ont pas. L'analogie est donc le ciment qui enlace deux réalités constatées, et qui permet

de les conceptualiser non sous forme de dualisme, ce que les médiévaux n'aiment pas, mais plutôt comme une unité, une continuité.

4.6 Les résultats face à l'historiographie

Ce mémoire n'a jamais eu la prétention de remettre en cause l'historiographie, mais plutôt de partir d'une base de données riche en sources et, au moyen d'une méthodologie qui nous est propre, d'obtenir des résultats à mettre en parallèle avec plusieurs assertions des historiens qui nous ont précédés. Pour répondre à la brève présentation historiographique du premier chapitre, nous tenions à évoquer trois thématiques principales : que disent les résultats du prolongement de l'esclavage antique au Moyen Âge ? S'accordent-ils à l'idée de Marc Bloch et de Fustel de Coulanges ? Et comment aborder la théorie mutationniste autour de l'an mil avancée par Georges Duby à la lumière de la sémantique diachronique ?

Pour reprendre les mots de Marc Bloch, "En rompant avec l'esclavage, le Moyen Age, dont les mœurs sociales n'étaient pas tendres, n'a certes pas détruit, ni prétendu détruire, les inégalités de fait ou de naissance ; mais il leur a donné, si l'on peut dire, une tonalité plus humaine". Il y a en un sens une continuité dans les faits entre l'époque antique et le Moyen Âge. Des hommes et des femmes continuent d'obéir à d'autres, mais les modalités de leur asservissement ont changé. Nos données ne permettent pas d'émettre un avis sur la transition entre l'Antiquité et le Moyen Âge, mais permettent néanmoins de plussoyer la plus grande humanité du traitement à l'époque médiévale. Parmi les cooccurrents de *seruus*, aucun ne mentionnait d'acte de violence. Tout au plus était-il cédé d'un seigneur à un autre, pour cause de son appartenance à la terre. Il est clair néanmoins que les chartes ne sont probablement pas le type de document le plus propice à une analyse sémantique du traitement des *serui*. En outre, on observe également une évolution progressive, siècle après siècle, un coulissement lent de son sens qui va de pair avec une baisse d'occurrences. Une moindre apparition dans les documents diplomatiques pourrait résulter d'un meilleur traitement également. Il s'agirait d'une sorte de témoin silencieux d'une amélioration des conditions serviles. Entre le 12^e et le 14^e siècle, les paysans sont effectivement parvenus à racheter leur liberté, ou tout du moins à améliorer considérablement leur sort, au travers de la disparition de certaines marques évidentes de leur condition servile²⁴⁸. Le silence des occurrences pourrait l'expliquer *a posteriori*. En outre, le christianisme a changé le paradigme de la servitude. De la contrainte de l'esclavage que l'on peut métaphoriser comme un bâton, le christianisme se présente quant à lui comme une carotte. L'obéissance n'est plus contrainte de la même manière, et, interprétée au travers du prisme chrétien, devient même l'apanage des *serui*.

248 FELLER, L., *Paysans et seigneurs au Moyen-Age*, p.166.

Les auteurs mutationnistes (Duby, Bonnassie, Poly) font remonter l'apparition du servage au début du 11^e siècle, et soutiennent qu'une servitude de type antique s'est prolongée jusque-là. Difficile de pleinement soutenir une telle hypothèse lorsque l'on constate la raréfaction de documents au cours du 11^e siècle et, par conséquent, la baisse du nombre d'occurrences. Cette baisse s'effectue également dans le courant du 12^e siècle et se marque encore plus pour le 13^e siècle. Ce changement soudain pourrait renforcer les théories initiées par Marc Bloch et Georges Duby quant à l'apparition d'un nouveau servage²⁴⁹, même si l'appellation de servage est à nouveau contestable. Toujours est-il que l'on peut effectivement constater un changement radical qui s'opère à cette époque dans l'emploi du terme *servuus*.

Finalement, les conclusions historiques restent maigres, faut-il encore le rappeler? Mais l'objectif premier de ce mémoire était de s'essayer à une méthodologie nouvelle par l'usage des technologies numériques. Les conclusions à ce sujet sont claires.

4.7 Les pistes d'amélioration et de prolongement

D'un point de vue historique, il serait nécessaire de se pencher individuellement sur le fonctionnement de chacun des lemmes empruntés à Alice Rio. Mais plus important encore d'inclure pour les périodes de rupture (12^e et 13^e) les autres mots desquels se sont inspirés nos prédécesseurs, à l'instar de *rusticii*, *libellarii*, *massarii*, *liberi manentes*, *villani*²⁵⁰ et bien d'autres encore.

Il serait encore plus intéressant de creuser les pistes de formulaires et de collecter, pour toutes ces formules que nous avons relevées, leur sens à travers les siècles. En ce qui concerne le 9^e siècle en tout cas, il va sans dire que cela serait nécessaire, puisque que la majorité des occurrences de *servuus* renvoient à elles.

Mais là où notre recherche aurait le plus à gagner, c'est à comparer les bases de données diplomatiques avec d'autres corpus d'un type différent. Par exemple, le *Corpus Corporum*²⁵¹ qui, outre le fait qu'il se compose de l'équivalent de 130 millions de lettres²⁵², propose une panoplie de textes latins assez vaste, soit près de 6000, entièrement nettoyés, lemmatisés et disponibles librement. Pour faciliter l'exploitation, le site offre le choix du

249 CARRIER, N., "Nouveaux servages" et société en Europe (XIII^e-XX^e siècle), dans *Bulletin du centre d'études médiévales d'Auxerre*, n°12, p.1.

250 PANERO, F., *Le nouveau servage et l'attache à la glèbe aux XII^e et XIII^e siècles : l'interprétation de Marc Bloch et la documentation italienn*, dans *Mélanges de l'école française de Rome*, t.112/2, 2000, p.551.

251 <http://www.mlat.uzh.ch/home>

252 L'équivalent de 75 millions de mots.

format d'extraction : en XML/TEI ou en document .txt. En outre, le projet ne mobilise que peu de ressources, financières ou humaines, et fonctionne partiellement à partir de participations externes. En dehors de cela, le corpus se définit surtout comme étant une forme de synthèse, de répertoire de bases de données déjà existantes telles que la *Patrologia latina*, les *Auctores scientiarum varii* et beaucoup d'autres. On y retrouve donc un échantillon non exhaustif mais néanmoins représentatif des productions littéraires médiévales, y compris une version du texte biblique : la Vulgate latine. Tel qu'il est conçu pour le moment, le site offre la possibilité de réaliser des recherches de lemmes ou de mots au niveau du corpus global, ou bien d'un sous-corpus spécifique. Bien entendu, après l'extraction de toutes ces (méta-)données, il faudrait les lisser, les harmoniser avant de les passer au crible d'un lemmatiseur.

Une autre piste serait de voir dans quelle mesure il est possible d'inclure une autre typologie d'écriture, à savoir les textes hagiographiques et les sources narratives en règle générale. Les vies de Saints, narrées tout au long du Moyen-Âge, peuvent se révéler des plus pertinentes lorsqu'il est question d'analyse sémantique, tout comme les textes historiques. Quelques tentatives diverses de numérisation ont ainsi vu le jour, mais aucune n'a encore pu égaler les résultats du *Corpus corporum*. Citons à titre d'exemple le site de l'Unamur qui propose humblement des liens renvoyant à d'autres pages web contenant, quant à elles, des données hagiographiques²⁵³. Pointons plus particulièrement le site de Narrative Sources²⁵⁴, un projet mené parallèlement à celui des Diplomata Belgica. Néanmoins, malgré un répertoire à caractère transnational, le site ne dispose pas proprement de fichiers textes ou de fichiers XML qui faciliteraient l'intégration à notre projet. Il existe également une ébauche similaire réalisée à l'échelle locale, celle de la Normandie, répondant au nom de *Repertorium Textuum Hagiographicorum in Normannia conscriptorum* et « qui propose une analyse de tous les textes hagiographiques écrits sur le territoire du duché de Normandie entre 911 et 1204 »²⁵⁵. En outre, l'un des principaux éléments que devrait comporter le corpus est les *Acta Sanctorum*, pilier incontournable de l'hagiographie européenne. Mais il n'existe, à notre modeste connaissance, pas de version de cette collection lemmatisée. Il serait donc nécessaire de procéder à l'océrisation des fichiers .jpg disponibles en libre accès sur Internet²⁵⁶ et à la lemmatisation de ces documents afin d'avoir un corpus des plus homogènes.

In fine, même si l'on peut reconnaître une certaine abondance de corpus numériques,

253 <https://www.unamur.be/philolettres/histoire/a19.html>

254 DEPLOIGE, J., CALLENS, C., DEMONTY, P. et De TRÉ, G., *Remedying the obsolescence of digitised surveys of medieval sources. Narrative Sources and Diplomata Belgica* dans *BULLETIN DE LA COMMISSION ROYALE D'HISTOIRE : ACADEMIE ROYALE DE BELGIQUE*, p.160

255 VANGONE, L., "Le ReTeHNor : un répertoire numérique de l'hagiographie latine de Normandie" dans *Tabularia*

256 <http://www.patristique.org/Acta-sanctorum.html>

comparativement aux périodes qui encadrent le Moyen Âge, il faut aussi admettre que la plupart du temps, ces tentatives importantes sont restées isolées. Cela justifierait l'intérêt de la démarche. Centraliser la documentation pour optimiser l'extraction d'informations. C'est pourquoi, à cette fin, croiser les typologies revêt une importance primordiale. Si l'on se réfère à l'historiographie récente de la discipline, des auteurs comme Nicolas Perreaux et son étude du terme *aqua* ou Alain Guerreau et son analyse de la sémantique de *vinea* se limitent principalement aux textes diplomatiques. Il serait possible de se rapprocher d'autant plus d'une vérité historique avec cette transversalité supplémentaire.

Table des figures

1	<i>Évolution des occurrences de seruus par demi-siècle dans les CEMA.</i>	3
2	<i>Évolution de la proportion de serf et servage dans Google Ngram Viewer</i>	14
3	<i>Évolution de la proportion de serf et servage dans Gallicagram</i>	15
4	<i>Différence d'occurrences entre les lemmes.</i>	34
5	<i>Exemple de l'algorithme de suppression de seruus seruatorum dei.</i>	35
6	<i>Fonction python pour écarter l'expression seruus seruatorum dei.</i>	36
7	<i>Heatmap des cas avec la formule seruus seruatorum dei.</i>	37
8	<i>Heatmap des cas sans la formule seruus seruatorum dei.</i>	38
9	<i>Fonction d'évincement de seruus seruatorum dei.</i>	43
10	<i>Fonction de préparation au TF-IDF.</i>	43
11	<i>Tableau TF-IDF de seruus.</i>	44
12	<i>Fusion des tableaux</i>	44
13	<i>Scatterplot des indices TF-IDF et de la Fréquence de seruus.</i>	45
14	<i>Dataframe comparatif des occurrences par charte de chaque lemme.</i>	46
15	<i>Dataframe comparatif du nombre de mots par charte pour chaque lemme.</i>	47
16	<i>Dataframe comparatif de la fréquence de chaque lemme par charte.</i>	48
17	<i>Dataframe comparatif de l'indice TF-IDF pour chaque lemme.</i>	49
18	<i>Formule de Dice</i>	52
19	<i>Indice de similarité de Dice (E. NEGRE)</i>	52
20	<i>Fonction de similarité de Dice.</i>	52
21	<i>Distribution des indices de Dice par lemme entre 800 et 1300 (SUB).</i>	53
22	<i>Distribution des indices de Dice par lemme entre 800 et 1300 (VBE).</i>	54
23	<i>Distribution des indices de Dice par lemme entre 800 et 1300 (QLF).</i>	55
24	<i>Formule de dépendance entre l'indice de Jaccard et l'indice de Dice</i>	55
26	<i>Indice de similarité cosinus (E. NEGRE)</i>	56
25	<i>Indice de similarité de Jaccard (E. NEGRE)</i>	56
27	<i>Fonction indice de Jaccard</i>	56
28	<i>Jaccard (QLF)</i>	57
29	<i>Jaccard (SUB)</i>	57
30	<i>Jaccard (VBE)</i>	58
31	<i>Fonction de similarité cosinus.</i>	60
32	<i>Corrélation de Pearson (E. NEGRE)</i>	61
33	<i>Comparaison de lexiques de cooccurrences</i>	62
34	<i>Comparaison des lemmes en fonction des indices.</i>	65
35	<i>Comparaison des indices de similarités pour seruus (QLF).</i>	66
36	<i>Comparaison des indices de similarités pour seruus (VBE).</i>	67
37	<i>Comparaison des indices de similarités pour seruus (SUB).</i>	67

38	<i>Diagramme comparatif des types d'adjectifs entre le 9^e et le 13^e siècle.</i>	71
39	<i>Pie charts en bâtonnets des différentes catégories d'adjectifs.</i>	73
40	<i>Pie charts de la distribution des cas des adjectifs accordés à seruus au 9^e siècle.</i>	74
41	<i>Pie charts de la distribution des cas des adjectifs accordés à seruus au 13^e siècle.</i>	75
42	<i>Nested Pie Chart des adjectifs cooccurents de seruus entre le 9^e et le 13^e siècle</i>	76
43	<i>Dataframe des occurrences par pays au 13^e siècle (QLF)</i>	77
44	<i>Dataframe des occurrences par pays du 9^e au 13^e siècle (QLF)</i>	77
45	<i>Évolution des cooccurrences de verbes avec seruus</i>	81
46	<i>Évolution du nombre d'occurrence de seruus par demi-siècle sans seruus seruorum dei.</i>	82
47	<i>Substantifs les plus fréquemment employés avec seruus dans la première moitié du 9^e siècle.</i>	84
48	<i>Substantifs les plus fréquemment employés avec seruus dans la première moitié du 10^e siècle.</i>	86
49	<i>Substantifs les plus fréquemment employés avec seruus dans la première moitié du 12^e siècle.</i>	87
50	<i>Occurrences de seruus seruorum dei par siècle.</i>	98

Bibliographie

— Ouvrages de synthèse

- (1) AMBRISIO, A., BARRET, S. et VOGELER, G., *Digital diplomatics : the computer as a tool for the diplomatist ?*, 2014.
- (2) BANNOUR, S., AUDIBERT, L. et NAZARENKO, A., *Mesures de similarité distributionnelle entre termes*, 2011.
- (3) BASCHET, J., *La civilisation féodale, de l'an mil à la colonisation de l'Amérique*, 2018.
- (4) BLAISE, A., *Dictionnaire latin-français des auteurs chrétiens*, 1954.
- (5) BOIS, G., *La mutation de l'an mil. Lournand, village mâconnais de l'Antiquité au féodalisme*, Paris, 1989.
- (6) BONDUE, D., *De servus à slavus : la fin de l'esclavage antique (371-918)*, 2011.
- (7) BRÉAL, M. et BAILLY, A., *Dictionnaire étymologique latin 1918*, 2017.
- (8) CECCHINI, E., *Uguccione da Pisa : Derivationes*, 2004.
- (9) CORRIOL, V., *Les serfs de Saint-Claude* 2009.
- (10) DAHAN, G., *Lire la Bible au Moyen Âge : essais d'herméneutique*, Genève, 2009.
- (11) DALARUN, J., *Gouverner c'est servir*, 2004.
- (12) DE COULANGES, F., *Histoire des institutions de l'ancienne France : la monarchie franque*, 1888.
- (13) DE SAINLIENS, C., *The Treasurie of the French Tong : Teaching the ways to Varie All Sortes of Verbes*.
- (14) DICKÈS, C., *Dictionnaire du Vatican*, 2013.
- (15) DON N., H.-Jr, *Servants of the servants : a biblical Theology of Leadership*, 2003.
- (16) DUBY, G., *Guerriers et paysans VIII-XII^e siècles, premier essor de l'économie européenne*, 2003.
- (17) E. BROWN, R., *Que sait-on du Nouveau Testament ?*, 2011.
- (18) ERNOUT, A. et MEILLET, A., *Dictionnaire étymologique de la langue latine*, 2001.
- (19) EVERT, S., *The Statistics of Word Cooccurrences : Word pairs and Collocations*, 2004.
- (20) FEAR, A. et WOOD, J., *Isidore of Seville and his reception in the Early Middle Ages : transmitting and transforming knowledge*, 2016.
FELLER, L., *Paysans et seigneurs au Moyen-Age*
- (21) GRIESER, H., *Sklaverei in spantiken und früe mittelalterlichen Gallien*, 1997.

- (22) GRONDEUX, A. et CINATO, F. (éd), *Liber Glossarum digital*, Paris, 2016.
- (23) GUERREAU, A., *L'avenir d'un passé incertain*, 2001.
- (24) GUERREAU-JALABERT, A., *Occident médiéval et pensée analogique : le sens de spiritus et caro*, dans GENET, J.-P., *La légitimité implicite*, p.457-476.
- (25) GUYOTJEANNIN, O., *Diplomatique médiévale*, 2006.
- (26) HANCIL, S., *Histoire de la langue anglaise*, 2013.
- (27) HARTOG, F., *le XIX^e siècle et l'histoire : le cas Fustel de Coulanges*, 1988.
- (28) LOBRICHON, G., *La Bible au Moyen Age*, 2003.
- (29) MAGNOU-NORTIER, É., *Servus - servitium : une enquête à poursuivre*, dans *Media in Francia. Recueil de mélanges offert à Karl Ferdinand Werner à l'occasion de son 65^e anniversaire par ses amis et collègues français*, 1989.
- (30) MEILLASSOUX, C., *Anthropologie de l'esclavage*, Paris, 1986
- (31) NEGRE, E., *Comparaison de textes : quelques approches...*, 2013.
- (32) NIERMEYER, J.-F., *Medieval Latinatis Lexicon Minus*, 1954.
- (33) PARISSE, M., , *Un lexique latin médiéval-français ?*, dans PARISSE, M., *Les Historiens et le latin médiéval*, p.315-319.
- (34) PERREAUX, N., *De l'accumulation à l'exploitation ? Expériences et propositions pour l'indexation et l'utilisation des bases de données diplomatiques*, dans AMBRSIO, A., BARRET, S. et VOGELER, G., *Digital diplomatics : the computer as a tool for the diplomatist ?* , p.187-210.
- (35) POLY, P. et BOURNAZEL, E., *La mutation féodale, X^e-XII^e siècles*, Paris, 2004.
- (36) REYDELLET, M., *Isidore de Séville : Étymologies*, t.IX, 1984.
- (37) RIO, A., *Slavery after Rome (500-1100)*, 2017.
- (38) URBANCZYK, P., *Europe around the year 1000*, 2001.
- (39) VAN LIER, F., *An introduction to the Medieval Bible*, 2014.
- (40) VAN OLBERG, G., *Freie, Nachbarn Gefolgsleute*, 1983.
- (41) VERLINDEN, C., *L'esclavage dans l'Europe médiévale*, t.1, 1955.
- (42) VERLINDEN, C., *L'esclavage dans l'Europe médiévale*, t.2, *Italie, colonies italiennes du Levant, Levant latin, empire byzantin*, 1977.
- (43) VERLINDEN, C., *L'origine de sclavus = esclave*, 1943.
- (44) VICTOR, S., *Les fils de Canaan*, Paris, 2019.
- (45) YANOSKI, J., *De l'abolition de l'esclavage ancien au Moyen Âge et de sa transformation en servitude*, 1860.

— **Articles**

- (1) ANDRIEUX, P., *Les usages de la sevitute. Seigneurs et paysans dans le royaume de Bourgogne (VI^e-XV^e siècle)*, compte rendu dans *Revue historique du droit français et étranger*, vol.92, n°2, p.321-324.
- (2) BARTHÉLÉMY, D., *La mutation féodale a-t-elle eu lieu ?*, dans *Annales, Économies, Sociétés, Civilisations*, 47^e année, n°3, 1992, p.767-777.
- (3) BASDEVANT-GAUDEMET, B., *Le pape évêque de Rome*, dans *Pouvoirs*, 2017, n°162, p.115-126.
- (4) BENOIT, J.-L., *Clef du texte, clef du royaume. La lecture de la Bible au Moyen Âge comme paradigme de la littérature*, dans *Les Clefs des textes médiévaux*, 2006, p.303-320.
- (5) BERTRAND, P., *À propos de la révolution de l'écrit (X^e-XIII^e siècle). Considérations inactuelles*, dans *Médiévales*, n°56, p.75-92.
- (6) BIONDI, L., *Sur quelques mots et leur parcours dans la lexicographie latine du Moyen Âge et de l'Humanisme*, dans *Histoire Épistémologie Langage*, t.39, fasc.2, 2017, p.153-176.
- (7) BLOCH, M., *Comment et pourquoi finit l'esclavage antique ?*, dans *Annales*, p.30-44.
- (8) BLOCH, M., *Liberté et servitude personnelles*, dans *Mélanges historiques*, p.296 et seq..
- (9) BOGAERT, P.-M., *La Bible latine des origines au Moyen Âge*, dans *Revue Théologique de Louvain*, t.19/2, 1988, p.137-159.
- (10) BON, B. et GUERREAU-JALABERT, A., *L'apport des dictionnaires dans une recherche sémantique, l'exemple du trésor*, dans *Listy Filologické/Folia Philologica*, vol.131, n°1-2, 2008, p.85-101.
- (11) BON, B. et GUERREAU-JALABERT, A., *"Pietas" : réflexions sur l'analyse sémantique et le traitement lexicographique d'un vocable médiéval*, dans *Médiévales*, n°42, 2002, p.73-88.
- (12) BONNASSIE, P., *Survie et extinction du régime esclavagiste dans l'Occident du haut moyen âge (IV^e-XI^e s.)*, dans *Cahiers de civilisation médiévale*, 28^e année, n°112, 1985, p.307-343.
- (13) BRUNNER, T., *Le passage aux langues vernaculaires dans les actes de la pratique en Occident*, dans *De Boeck Supérieur : Le Moyen Âge*, 2009, t.115, p.29-72.
- (14) BUC, P., *Les débuts de Sauxillanges : à propos d'un acte de 927*, dans *Bibliothèque de l'école des chartes*, 1998, t.156, p.537-545.

- (15) BULLY, P., *Zipf, créateur de la linguistique statistique*, dans *Communication et langages*, n°2, 1969, p.23-28.
- (16) CARRIER, N., *De l'esclavage au servage : pour une étude des dynamiques de la servitude*, dans *Médiévales*, n°81, 2022, p.179-196.
CARRIER, N., "Nouveaux servages" et société en Europe (XIII^e-XX^e siècle), dans *Bulletin du centre d'études médiévales d'Auxerre*, n°12, p.1.
- (17) CHAI-ELSHOLZ, R., *Servus servorum Dei in Context : Historical and Literary Intertextuality to c. 875*, dans FRANK, T. et RANDO, D., *Ernst Kantorovicz 1895-1963) : Storia politica come scienza culturale*, p.69-90.
- (18) CORRIOL, V., *Sandrine Victor, Les fils de Canaan : l'esclavage au Moyen Âge*, dans *Cahiers de civilisation médiévale*, t.249, 2020, p.95-96.
- (19) DÉLÉAGE, A. et BLOCH, M., *Les caractères originaux de l'histoire rurale française. Compte rendu de OSLO, H.*, dans *Annales du midi : revue archéologique, historique et philologique de la France méridionale*, t.44, n°174, p.211-219.
- (20) DELESALLE, S., *De l'adjectif : lecture d'une grammaire*, dans *Documentation et recherche en linguistique allemande contemporain - Vincennes*, n°41, p.169-188.
- (21) DEPLOIGE, J., CALLENS, C., DEMONTY, P. et De TRÉ, G., *Remedying the obsolescence of digitised surveys of medieval sources. Narrative Sources and Diplomata Belgica* dans *BULLETIN DE LA COMMISSION ROYALE D'HISTOIRE : ACADEMIE ROYALE DE BELGIQUE*.
- (22) DERVILLE, A. et LOCH, M., *M. Bloch, Rois et serfs et autres écrits sur le servage* dans *Revue du Nord*, t.329, 1996, p.191-192.
- (23) DOEHAERD, R., *Institutions médiévales : un livre, une méthode*, dans *Annales. Économies, Sociétés, Civilisations.*, 4^e année, n°1, 1949, p.23-28.
- (24) DUBY, G., *Sur les voies ouvertes par Marc Bloch : esclavage et servage au Moyen Âge*, dans *Annales*, t.12/1, p.123-126.
- (25) DUFFO, J. et ALLARD, P., *Les origines du servage en France*, dans *Revue d'histoire de l'Église de France*, t.5, n°28, p.512-515.
- (26) FAVEZ, C., *Review of Dictionnaire latin-français des auteurs chrétiens by A. Blaise*, dans *Latomus*, t.14/1, p.135-139.
- (27) FELLER, L., *Alice Rio, Slavery after Rome (500-1100)*, dans *Cahiers de civilisation médiévale*, n°243, 2018, p.318-320.
- (28) GRONDEUX, A., *Le "Liber Glossarum" (VIII^e siècle) : prélogomènes à une nouvelle édition*, dans *Bulletin Du Cange*, n°69, 2011, p.23-51.
- (29) GUERREAU, A., *Pourquoi (et comment) l'historien doit-il compter les mots ?*, dans *Histoire & Mesure*, vol.4, 1989, n°1-2, p.81-105.

- (30) GUERREAU, A., *Textes anciens en série*, dans *Bulletin du centre d'études médiévales d'Auxerre*.
- (31) GUYOTJEANNIN, O., *La diplomatique médiévale et l'élargissement de son champ*, dans *De la diplomatique médiévale à la diplomatique contemporaine (actes du colloque organisé par l'École nationale des chartes et la Bentley historical Library de l'université de Ann-Arbor [Michigan, États-Unis]*, Paris, 1993, p.12-18.
- (32) Heiden Serge, Magué Jean-Philippe, & Pincemin Bénédicte. (2010). TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement. In JADT 2010 : 10th International Conference on the Statistical Analysis of Textual Data (pp. 12 p.). Rome, Italie.
- (33) HETLEN, K. et BERTELS, A., *Sémantique distributionnelle en linguistique de corpus*, dans *Langages*, n°201, 2016, p.51-64.
- (34) HOEGGER, M., *Une obéissance qui libère*.
- (35) JACQUIN, G., *L'oeuvre de Robert Estienne et les dictionnaires de la Renaissance*, dans *Revue 303 arts recherches créations. Patrimoine linguistique*, n°120, 2012, p.28-33.
- (36) JAMME, A., *Écrire pour le pape du XI^e au XIV^e siècle. Formes et problèmes.*, dans *Mélanges de l'École française de Rome - Moyen Âge*, t.128/1, 2016.
- (37) KAYNAR, E., *De la philologie à la sémantique historique en études ottomanes*, dans *European Journal of Turkish Studies*, n°31, 2020.
- (38) LAUWERS, M., *L'Église dans l'Occident médiéval : histoire religieuse ou histoire de la société? Quelques jalons pour un panorama de la recherche en France et en Italie au XX^e siècle* dans *Mélanges de l'école française de Rome* , t.121/2, 2009, p.267-290.
- (39) LAUWERS, M., *Usages de la Bible et institution du sens dans l'Occident médiéval*, dans *Médiévales*, t.55, p.5-18.
- (40) LE GOFF, H., *Droit de la guerre et droits des prisonniers de guerre au XVI^e siècle : le cas de la Ligue en Bretagne (1587-1598)*, dans *Annales de Bretagne et des Pays de l'Ouest*, n°124/2, 2017, p.7-28.
- (41) LE COQ, A., *Compte-rendu de la journée d'études Microhistoire et pratiques historiennes. Échelles, acteurs, formes narratives*, 2012, dans *Memini*, n°15.
- (42) MAGNANI, E., *Uxor et femina. Enquête sur la désignation des femmes dans les documents diplomatiques bourguignons (IX^e-XI^e siècle)*, dans RENARD, J.-P., Jean-Paul RENARD. *La place et le rôle des femmes dans l'histoire de Cluny : en hommage à Ermengarde de Blesle, mère de Guillaume le Pieux.*, Actes du colloque de Blesle des 23 et 24 avril 2010, Editions Créer, pp.125-138, 2013.

- (43) MAGNOU-NORTIER, E., *Contribution à l'étude des documents falsifiés : l'acte de fondation du monastère de Sauxillanges*, dans *Chaiers de civilisation médiévale*, t.21, 1978, p.313-328.
- (44) MAYAFFRE, D., *De l'occurrence à l'isotopie : les co-occurrences en lexicométrie*, dans *Syntaxe et Sémantique*, n°9, 2008/1, p.53-75.
- (45) MICHEL-TOCK, B., *L'étude du vocabulaire et la datation des actes : l'apport des bases de données informatisées*, dans GERVERS, M., *Dating Undated Medieval Charters*, p.80-96.
- (46) MULLER, C., *La statistique lexicale*, dans *Langue française*, n°2, 1969, p.30-43.
- (47) NERLICH, B., *Avant-propos : La sémantique historique en Allemagne, en Angleterre et en France*, dans *Histoire Épistémologie Langage*, t.15, fasc.1, 1993, p.5-30.
- (48) PANERO, F., *Le nouveau servage et l'attache de la glèbe aux XII^e et XIII^e siècles : l'interprétation de Marc Bloch et la documentation italienne*, dans *Mélanges de l'école française de Rome*, t.112/2, 2000, p.551-561.
- (49) PARVULESCU, A., *Lat. Servus*, dans *Indogermanische Forschungen*, vol.115, Berlin, 2010, 190-197.
- (50) PERREAUX, N., *Possibilities, challenges and limits of a European charters corpus (CEMA)*.
- (51) PFISTER, C., BLOCH, M. et DIEHL, C., *Le centenaire de Fustel de Coulanges (1830-1889)*, dans *Revue internationale de l'enseignement*, t.84, 1930, p.161-187.
- (52) POLO DE BEAULIEU, M.-A., *Panorama de la lexicométrie*, dans *Histoire & Mesure*, vol.2, 1989, p.173-197.
- (53) RAJMAN, M. et LEBART, L., *Similarités pour données textuelles*.
- (54) RENOARD, Y., *Charles Verlinden, L'esclavage dans l'Europe médiévale. Tome I. Péninsule Ibérique-France*, dans *Bulletin hispanique*, t.58, n°3, 1956, p.374-376.
- (55) RIO, A., *Les formulaires et la pratique de l'écrit dans les actes de la vie quotidienne*, dans *Médiévales*, p.11-22.
- (56) RYCHLY, P., *Un score d'association favorable aux lexicographes. Actes du deuxième atelier sur les progrès récents dans le traitement du langage naturel slave*, dans *RASLAN*, 2008, p.6-9.
- (57) SCHMITT, J.-C., *L'anthropologie historique de l'Occident médiéval. Un parcours*, dans *L'Atelier du Centre de recherches historiques*, t.6, 2010.
- (58) SHAHZAD, Q. et RAMSHA, A., *Text mining : Use of TF-IDF to Examine the Relevance of Words to Documents*, dans *International Journal of Computer Applications* (0975– 8887), vol.181, 2018, p.25-29.
- (59) SINGHAL, A., *Modern Information : A brief Overview*.

- (60) SOUTI, C., *La lexicométrie ou l'analyse du discours assisté par ordinateur : ce que l'informatique et les mathématiques peuvent apporter à la littérature et la linguistique*, dans *Expressions*, n°1, 2015, p.122-130.
- (61) TURPIN, B., *L'esclavage vu par les dictionnaires français*, dans DUMONT, L. et FREMIN, M., *Esclavage, Esclavages*, UCP, p.31-58.
- (62) UEHLINGER, C., *Genèse 1-11*, dans RÖMER, T., MACCHI, J.-D. et NIHAN, C., *Introduction à l'Ancien Testament*, 2009, p.204-207.
- (63) VAN HAEPEREN, F., *Des pontifes païens aux pontifes chrétiens. Transformation d'un titre : entre pouvoirs et représentations*, dans *Revue belge de philologie et d'histoire*, t.81, fasc.1, 2003, p.137-159.
- (64) VANGONE, L., "Le ReTeHNor : un répertoire numérique de l'hagiographie latine de Normandie" dans *Tabularia*.
- (65) VERHULST, A., *The decline of Slavery anduna*, dans *Past and Present*, n°133, 1991, p.195-203.
- (66) VERLINDEN, C., *Slavenhandel en economische ontwikkeling in midden-, oost-, en noord-Europ gedurende de hoge Middeleeuwen* dans *Mededelingen van de Koninklijke academie voor wetenschappen, letteren en schone kunsten van België. Klasse der letteren*, 1979.
- (67) ZUMSTEIN, J., *Le lavement des pieds (Jean 13, 1-20) : un exemple de la conception Johannique du pouvoir*, dans *Revue de Théologie et de Philosophie*, vol.132, n°4, 2000, p.345-360.

— Webographie

- (1) <http://telma.irht.cnrs.fr/outils/originaux/colophon/>
- (2) https://www.diplomata-belgica.be/colophon_fr.html
- (3) http://deeds_library_utoronto_ca/
- (4) <http://www.cbma-project.eu/>
- (5) <http://ducange.enc.sorbonne.fr/SERVUS#SERVUS-71>
- (6) <https://gaffiot.org/85072>
- (7) <https://gallica.bnf.fr/ark:/12148/btv1b105154441.image#>
- (8) GUERREAU, A., <http://www.glossaria.eu/minorque/cours/Guerreau-5-Semantique.pdf>
- (9) <http://liber-glossarum.huma-num.fr/exist/apps/libgloss/entry-SE612>

- (10) https://devhist.hypotheses.org/898#_ftn23
- (11) [https://support.minitab.com/fr-fr/minitab/18/help-and-how-to/statistics/basic-statistics/how-to/store-descriptive-statistics/interpret-the-statistics/](https://support.minitab.com/fr-fr/minitab/18/help-and-how-to/statistics/basic-statistics/how-to/store-descriptive-statistics/interpret-the-statistics/interpret-the-statistics/)
- (12) https://stringfixer.com/fr/Sorenson_index
- (13) https://link.springer.com/chapter/10.1007/978-3-642-00296-0_5?noAccess=true
- (14) http://elec.enc.sorbonne.fr/cid/cid1992/art_02
- (15) <https://fr-academic.com/dic.nsf/frwiki/1531755>
- (16) Bible de Jérusalem, p.938; https://livres-mystiques.com/partieTEXTES/Bibliques/Bible_de_Jerusalem.pdf
- (17) <http://www.patristique.org/Acta-sanctorum.html>
- (18) https://www.unamur.be/philo_lettres/histoire/a19.html
- (19) <http://www.mlat.uzh.ch/home>

