



**CBMA. *Chartae Burgundiae Medii Aevi VII.*
« Le “vocabulaire courant” en diplomatique :
techniques et approches comparées »**

Dijon, Maison des sciences de l'homme, 25 janvier 2013

Nicolas Perreaux et Coraline Rey



Édition électronique

URL : <https://journals.openedition.org/cem/12984>
DOI : 10.4000/cem.12984
ISSN : 1954-3093

Éditeur

Centre d'études médiévales Saint-Germain d'Auxerre

Référence électronique

Nicolas Perreaux et Coraline Rey, « CBMA. *Chartae Burgundiae Medii Aevi VII.*
« Le “vocabulaire courant” en diplomatique : techniques et approches comparées » », *Bulletin du centre
d'études médiévales d'Auxerre | BUCEMA* [En ligne], 17.1 | 2013, mis en ligne le 03 juin 2013, consulté le
22 septembre 2022. URL : <http://journals.openedition.org/cem/12984> ; DOI : [https://doi.org/10.4000/
cem.12984](https://doi.org/10.4000/cem.12984)

Ce document a été généré automatiquement le 22 septembre 2022.



Creative Commons - Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions
4.0 International - CC BY-NC-SA 4.0
<https://creativecommons.org/licenses/by-nc-sa/4.0/>

CBMA. *Chartae Burgundiae Medii Aevi* VII.

« Le “vocabulaire courant” en diplomatie : techniques et approches comparées »

Dijon, Maison des sciences de l’homme, 25 janvier 2013

Nicolas Perreaux et Coraline Rey

- 1 « Si les numérisations ne cessent de s’accumuler depuis trois décennies, y compris les reproductions des manuscrits, leur exploitation demeure encore en deçà des changements potentiels importants qu’on pourrait présager. Ce décalage provient à la fois des questions posées par les historiens à cette documentation, qui n’ont pas elles-mêmes évolué, et, par conséquent, des moyens donnés pour l’interroger. En d’autres termes, c’est la portée des opérations intellectuelles qui conditionnent les opérations techniques ¹. »
-
- 2 À ce jour, les chartes bourguignonnes réunies au sein du corpus CBMA offrent le schéma d’interrogation en ligne le plus flexible pour une base de données diplomatique de cette taille : plus de 13 000 actes, sous le logiciel de fouille de données Philologic ², désormais accessibles via un serveur hébergé par le TGE Adonis (<http://www.artehis-cbma.eu/>) ³. Ce nouvel hébergement, effectif depuis le mois de septembre 2012, apporte une structure technique non négligeable, permettant également le stockage de toutes les données, aussi bien en mode image qu’en mode texte. À moyen/long terme, il signifie une pérennisation pour l’outil. Plus précisément, la journée d’étude de janvier 2012 avait permis de discuter de l’intérêt de celui-ci et de partager des expériences en

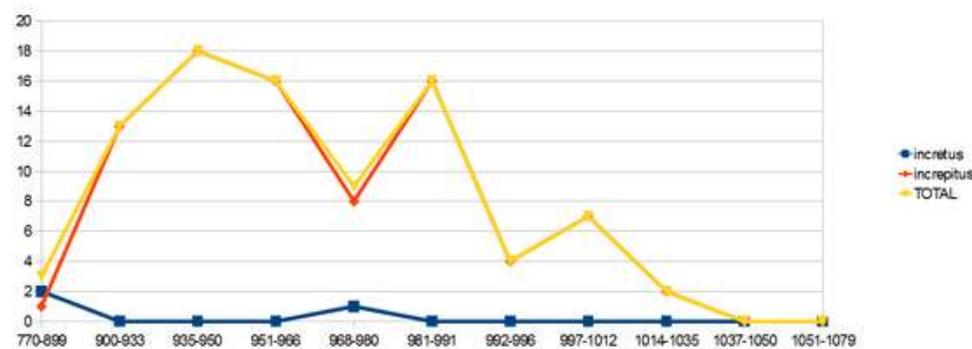
cours, tout en insistant sur la prise en main – de plus en plus sensible – du logiciel par les chercheurs. Grâce à cette base et à cet outil émerge en effet la possibilité de réaliser une diplomatique comparée, thème ancien, mais largement délaissé dans les dernières décennies au profit d'études ciblées⁴. Cette approche, s'appuyant sur une lecture renouvelée du matériau, renforce la mise en cause du « mythe de l'anarchie documentaire »⁵, en montrant que les usages de termes ou de formules souvent considérés comme allant de soi sont en fait typiques d'un moment, d'un espace, d'une écriture.

- 3 Cette journée souhaitait se consacrer aux expériences dans le domaine des termes ou des formules considérés comme évidents/courants, c'est-à-dire dont la forte récurrence pourrait faire penser a priori à des habitudes ou à des emplois globalement partagés. L'objectif était de mettre en lumière quelle part du vocabulaire « commun » répondait à ce paradigme d'une écriture s'inscrivant dans un moment, mais surtout dans un espace, en s'intéressant notamment à la sémantique de ces termes, justement fortement corrélée à leur présence/absence dans des zones à la structuration sociale sans doute différente. S'inscrivant dans une série de travaux, révélateur dans une certaine mesure de l'émergence d'un nouveau paradigme et qu'on a fait le choix de considérer ici comme un tout. On pense bien entendu aux travaux précurseurs d'Alain Guerreau et d'Anita Guerreau-Jalabert sur la question⁶, mais aussi à ceux d'Isabelle Rosé sur *decima*, de Didier Méhu sur *locus*, de Joseph Morsel sur *Brief* et *Schrift*, de Julien Demade et Joseph Morsel sur les *Egeinleute* de Franconie, d'Eliana Magnani sur le vocabulaire du don et sur *uxor/femina*, de Bruno Bon sur *planus*, de Ludolf Kuchenbuch sur *textus* ou *opus*, de Michel Lauwers sur *cimiterium/terra cimiteriata* ou encore *circuitus*, de Benoît-Michel Tock sur *altare*, de Didier Panfili sur *campus*, de Nicolas Perreaux sur *aqua* et plus récemment sur *villa* ou *terra*⁷. Cette journée d'études fut ainsi l'occasion de mesurer combien ce champ de recherche pouvait être riche, à la fois en termes d'approches et de résultats. Les cinq communications présentées, leur apport, aussi bien au plan théorique, technique qu'historique, montrent en effet que la question posée en préambule – « existe-t-il un vocabulaire courant en diplomatique ? » – pousse à s'interroger à la fois sur les structures – fréquentielles mais aussi spatiales – du vocabulaire et sur le sens de l'écriture médiévale. Comme chaque année, dans une optique holiste – allant de la création des fichiers à l'analyse historique proprement dite –, la journée fut aussi l'occasion de présenter les ajouts apportés à la base des CBMA au fil des mois.

Fréquences et spatialisation du vocabulaire

- 4 Pour mieux cadrer et réfléchir au propos de la journée, Nicolas Perreaux s'est d'abord interrogé sur le sens et l'usage actuel des expressions « vocabulaire courant »/« vocabulaire fréquent ». Partant d'une réflexion sur le sens de ces qualificatifs, il a relevé que les occurrences de ces syntagmes étaient rares en diplomatique, et, plus globalement, en médiévistique – en ayant recours à la base Persée, donc aux volumes de la *Bibliothèque de l'École des chartes*, aux volumes des congrès de la Société des historiens médiévistes de l'enseignement supérieur, aux revues *Médiévales* et *Le Moyen Âge*, etc. Rarement employées, malgré quelques occurrences notables – par exemple chez Marc Bloch –, ces expressions renvoient néanmoins à une conception qui voudrait que le vocabulaire fréquent soit transparent

Fig. 2 – Nombre d'occurrences des lemmes *incretus/incretus* dans les CBMA. Afin d'éviter les biais, le corpus a été découpé en 11 paquets d'environ 440 documents (à +/- 4%), pour la période 770-1079.



- 7 Pour rendre compte de ce phénomène et définir la frange du vocabulaire fortement régionalisée, Nicolas Perreaux avance le qualificatif d'« endémique », emprunté à la biogéographie ¹⁵.
- 8 Le retour sur deux autres termes, *villa* ¹⁶ et *campus* ¹⁷, bien plus fréquent qu'*incretus*, a permis de conclure l'exposé par une analyse non seulement géographique, mais aussi et surtout chronogéographique, avec une réflexion sur la dynamique variable du lexique, en fonction des régions et espaces considérés.
- 9 Poussant plus loin cette analyse statistique et abstraite des problèmes posés par le vocabulaire des chartes, Alain Guerreau a, quant à lui, traité des possibilités et des difficultés engendrées par la numérisation de vastes corpus. Il montre ainsi qu'un objectif central reste à atteindre : « mettre au point des méthodes formelles semi-automatisées d'analyse sémantique, dans un cadre historique, en jouant sur les évolutions, et en écartant, autant que possible, tout lien avec un sens actuel ». Des étapes ont déjà été franchies dans cette perspective, une partie du cadre théorique ayant été fourni par les travaux de Jost Trier ¹⁸ ; quant aux outils informatiques pour l'analyse des textes, une série de tests commencés depuis maintenant trois ou quatre décennies tendent à se concrétiser ¹⁹. C'est dans cette dynamique scientifique que s'insère le lemmatiseur développé par l'équipe de l'ANR OMNIA, désormais efficace à plus de 97 %.
- 10 Parmi les points développés, Alain Guerreau est en particulier revenu sur le problème épineux et souvent sous-estimé de l'instabilité structurelle des fréquences du vocabulaire. Il rappelle ainsi que les mots rares, s'ils sont très peu fréquents et occupent, au plan fréquentiel, peu de place dans les textes, abondent et emplissent la plupart des pages de nos dictionnaires. À partir d'exemples choisis – *sum*, *flos*, *formalis* –, il montre que la forme de la distribution fréquentielle de ces lemmes dans la *Patrologie Latine* varie en fonction de leur fréquence relative (rang ; fig. 3 et 4).

Fig. 3 – Évolution du nombre d'occurrences du lemme *flos* dans la *Patrologie latine*, découpée en parties contenant un nombre équivalent de mots/*tokens* (1/100). On note la concentration du terme chez les auteurs les plus anciens présents dans la *Patrologie*.

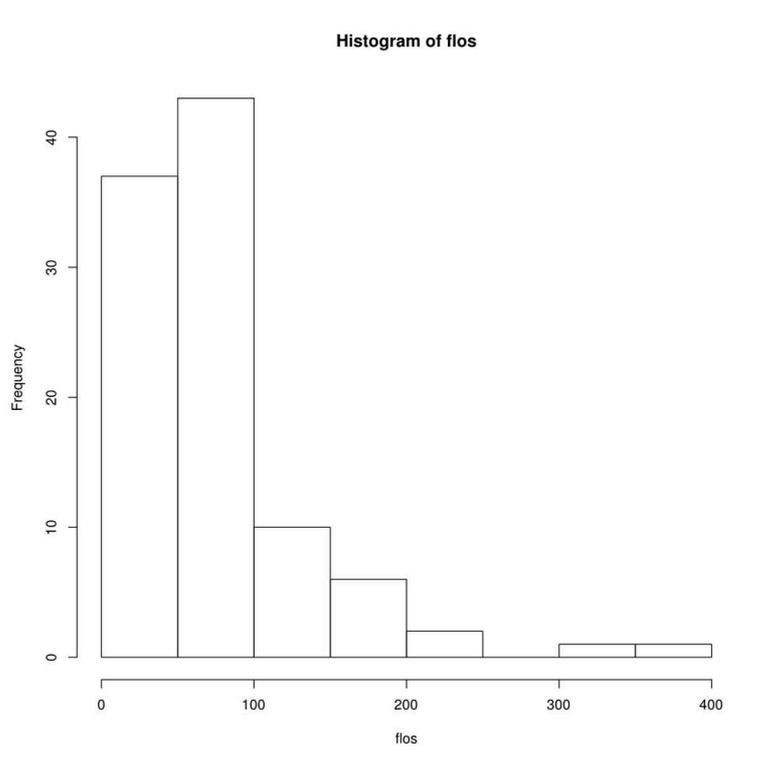
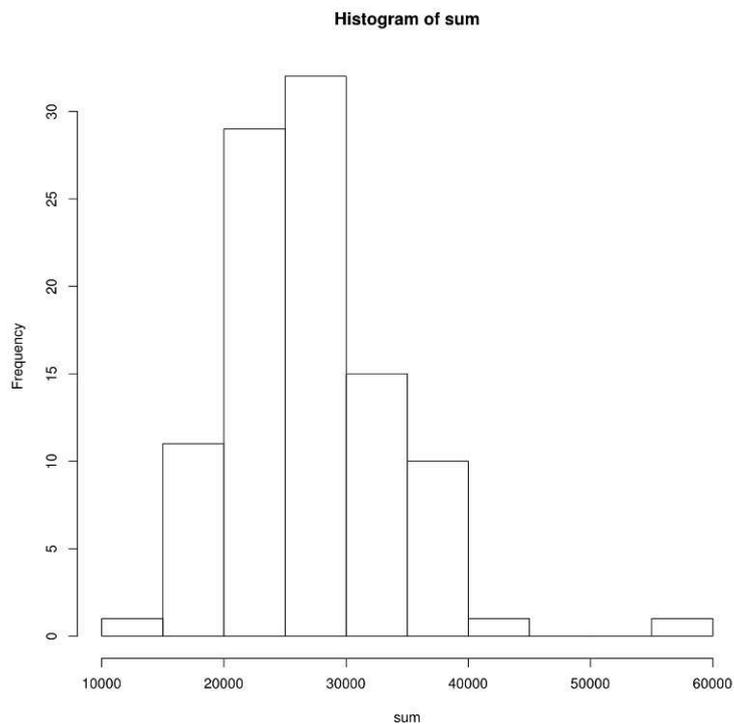


Fig. 4 – Évolution du nombre d'occurrences du lemme *sum* dans la *Patrologie latine*, découpée en parties contenant un nombre équivalent de mots/*tokens* (1/100). La répartition diffère sensiblement de celle observée pour *flos*.



- 11 Autrement dit, les problèmes qui se posent, au niveau des fréquences et donc de l'analyse historique qu'on peut en faire, ne sont pas identiques si les lemmes considérés sont courants, moyennement fréquents ou même rares. Concrètement, la taille d'un vocabulaire augmentant au fil de la lecture d'un texte, ces termes connaîtront, au sein de l'ensemble, des évolutions relatives divergentes, difficilement prévisibles, caractérisées par leur « typologie fréquentielle ».
- 12 On en déduit qu'il est impossible de comparer, statistiquement parlant, les répartitions chronologiques, typologiques ou géographiques, de termes possédant des taux d'occurrences radicalement différents. Surtout, on observe que les fréquences faibles (mots dits « rares ») sont plus sensibles à ces phénomènes d'« irrégularité statistiques ». Dans ce cadre, l'emploi de la notion de « fréquence moyenne » est particulièrement risqué voire fausse.
- 13 Une solution à ce problème consiste à comparer soit des textes de même taille, soit des « tranches » de textes possédant un nombre équivalent de tokens. À partir de la bibliothèque ZipfR²⁰, plusieurs explorations envisageables sont présentées : un graphique rang-fréquence, un spectre de fréquences, des courbes d'augmentation du vocabulaire. Dans le même ordre d'idées, relevant à la fois de la méthode et de la réflexion sur les structures du vocabulaire, Alain Guerreau a rappelé que les recherches sur les cooccurrences présentaient des perspectives prometteuses. À partir d'un travail sur les différents coefficients d'association, il retient en particulier celui de Dice²¹, et expose ainsi des évolutions autour de certains termes « courants » de la Patrologie – par exemple *species*, *panis*, *vinum*, etc.

Désigner, dénommer, décrire : l'écriture des stratégies sociales

- 14 Parallèlement à l'examen des structures fréquentielles et spatiales du vocabulaire des chartes, la journée a donné l'occasion de s'intéresser à des termes/lexiques précis – *dominus*, *ego*, champ lexical de l'écrit – relevant de cette vaste typologie qu'est le « vocabulaire courant ». Ces expériences montrent qu'il est tout à fait possible d'articuler, au sein de cette problématique, approche technique, réflexion abstraite et hypothèses historiques ; ces différentes échelles se combinant alors dans une approche sémantique.
- 15 Afin de comprendre les usages du mot *dominus* – terme « courant » s'il en est – dans les *scriptoria* méridionaux du VIII^e au XIII^e siècle, Didier Panfili s'est appuyé sur vingt-trois cartulaires, émanant principalement d'abbayes bénédictines, d'évêchés ou de chapitres canoniaux, du sud de la France actuelle. Cette lecture globale permet tout d'abord de dégager une caractéristique commune à toutes les compilations : le *dominus*, c'est avant tout Dieu ou le Christ. À partir de ce constat, Didier Panfili propose trois grandes phases chronologiques, ponctuées d'évolutions et de ruptures :
- de la fin du VIII^e siècle aux années 1050, il note une période de progressive diffusion du terme *dominus* pour désigner des membres de la société laïque. Dans un premier temps, néanmoins, le terme reste réservé, systématiquement et exclusivement, à l'empereur. C'est seulement après l'an Mil, que *dominus* se diffuse parmi l'aristocratie locale ;
 - entre les années 1060 et 1150, on observe l'existence d'un lien fort entre l'emploi du terme *dominus* et la cession de *res sacrae* (souvent des églises). Les laïcs ne sont alors

plus désignés par ce vocable, excepté lorsqu'ils cèdent des biens ecclésiastiques. Malgré cette restriction hiérarchique, l'usage du mot est de plus en plus fréquent et désigne, dès lors, souvent un personnel ecclésiastique élargi ;

- après 1155 environ, *dominus* semble à nouveau utilisé pour désigner certains laïcs, mais de façon parcimonieuse, et surtout sans que ceux-ci soient associés à des *res sacrae* : c'est cet usage qui semble se fixer au XIII^e siècle. On entrevoit ainsi une rupture de près d'un siècle dans l'emploi du terme. Le phénomène est d'autant plus remarquable que le féminin *domina* semble connaître une évolution parallèle tout à fait divergente, presque insensible à ces tentatives ecclésiastiques de prétention hégémonique à la titulature et à l'exercice du *dominium*.

- 16 En outre, cette situation tranche radicalement avec ce qui est observé dans le cartulaire des Trencavel, un cartulaire laïque, dans lequel près de 50 % des personnages laïques sont qualifiés de *dominus*. Ces constats permettent d'étayer l'hypothèse suivante : l'emploi du terme *dominus*, loin d'être idiomatique, suit de près les contours des tensions sociales, des ambitions idéologiques des uns et des autres, et, finalement, se montre révélateur de la volonté des établissements religieux de construire un monde nouveau, de leur prétention à la seigneurie universelle, concomitante des combats menés par l'Église. Après cette période que l'on qualifie de « réforme grégorienne », on assiste donc à une nouvelle dynamique laïque, phase qui voit s'affirmer la volonté des princes territoriaux et des monarques de renouveler leur domination territoriale, dans le cadre d'une compétition complexe, menée parfois avec, parfois contre l'*Ecclesia*. Les discussions autour de cette présentation furent nourries, amenant l'ensemble des participants à s'interroger sur les spécificités sociales des zones analysées, sur l'impact de ces spécificités sur l'écriture des actes et des différences dans le traitement des genres.
- 17 La question du genre des mots (masculin ou féminin) dans la désignation des disposants était au centre de l'analyse du lemme *ego* proposée par Eliana Magnani. Ses premières observations sont « issues d'une expérience de “lecture à distance” d'une masse de documents numérisés » à partir de l'analyse des cooccurrences des lemmes désignant les femmes, en particulier *uxor* et *femina* (fig. 5 et 6) ²².

Fig. 5 – Principales cooccurrences de *uxor* dans un corpus de 6348 actes bourguignons et méridionaux des IX^e-XI^e siècles.

<i>uxor</i>* : 5832 occurrences, dans 2597 unités documentaires (~ 41% du corpus)	
Principales cooccurrences de <i>uxor</i>* (5832)	
<i>ego</i> (1232)	
<i>infantibus</i> (352)	
<i>fili</i> (326)	
<i>filiis</i> (170)	
<i>filius</i> (167)	
<i>filio</i> (125)	
<i>filiorum</i> (72)	
<i>frater</i> (59)	
<i>infantes</i> (55)	
<i>fratris</i> (42)	
<i>filiabus</i> (40)	
<i>fratribus</i> (178)	
<i>domino</i> (161)	
	Principales cooccurrences de <i>uxor</i>* (5832)
	<i>nomine</i> (784)
	<i>nomen</i> (72)
	<i>servum</i> (94)
	<i>servo</i> (38)

Fig. 6 – Principales cooccurrences de *femina* dans un corpus de 6348 actes bourguignons et méridionaux des IX^e-XI^e siècles.

Principales cooccurrences de <i>femin</i>* (342)
<i>ego</i> (80)
<i>nomine</i> (37)
<i>homo</i> (44)
<i>homines</i> (20)
<i>filius</i> (17)
<i>filio</i> (12)
<i>infantibus</i> (8)
<i>dono</i> (22)
<i>donamus</i> (9)
<i>donationem</i> (17)
<i>vendo</i> (5)
<i>amore</i> (8)
<i>remedio</i> (10)
<i>anime</i> (5)

- 18 Son enquête a porté sur 6 348 actes du IX^e au XI^e siècle, contenus dans la base des CBMA, et dans plusieurs cartulaires de la vallée du Rhône et de la Provence²³. Elle a alors recherché le vocabulaire pouvant désigner les femmes, même si 90 % des disposants sont des hommes. En dressant un tableau d'occurrences et de cooccurrences, il apparaît que le terme *uxor* est le plus usité : parmi les disposants, la femme est avant tout cantonnée à son statut d'épouse. De la même manière, les termes désignant les

relations de parenté (*filia, soror, mater*) ou d'alliance (*jugalis, vidua*) ont été interrogés : les documents qui les contiennent font surtout état de transferts de biens ou de droits. Les cooccurrences les plus fréquentes des formes indiquant la parenté (*filia, soror, mater*) se réfèrent tout d'abord au tissu relationnel autour de *ego*. D'autres cooccurrences fréquentes – *infantibus, filius, frater, infantes, filiabus* – indiquent que *uxor* se combine avec les termes de la cellule conjugale et de la fratrie, une fratrie qui est pensée également dans le cadre de la fraternité chrétienne réunissant tous les hommes et les femmes, idée traduite par les formules telles *Domino fratribus*.

- 19 Grâce au tableau de cooccurrences de *uxor*, Eliana Magnani a également montré que l'apposition du lemme *nomen*, attribué aussi bien aux anthroponymes féminins qu'aux anthroponymes masculins des *servi*, était l'une des manières d'indiquer la relation de dépendance dans laquelle celles-ci se trouvaient. Quant au terme *femina*, il apporte des informations supplémentaires : parmi les cooccurrences les plus fréquentes de *femina*, on retrouve en effet *ego* et *nomine*. Concernant ce dernier, il s'agit d'un emploi dans les mêmes cadres déjà dégagés pour *uxor* : dans la formule *in Dei nomine* et en apposition à l'anthroponyme féminin. En revanche, les cooccurrences révèlent la jonction systématique de *femina* à *ego*. *Femina* semble donc être une manière de spécifier le genre d'*ego*, tandis que, lorsque le disposant est masculin, aucune spécification n'est requise. Parmi les autres cooccurrences les plus fréquentes de *femina* se retrouvent justement *homo*, dans des expressions qui énumèrent homme(s) et femme(s), soit dans des clauses de garantie, soit dans les mentions des dépendants qui sont l'objet d'un transfert. Une autre norme se dégage : les hommes sont toujours énumérés en premier.
- 20 Toutes ces observations permettent de montrer qu'il semble exister des marqueurs de catégorisation de personnes, par genre ou par relation de dépendance. De fait, l'anthroponyme seul semble ne pas pouvoir désigner les femmes. Ces mots disent bien plus que « nom » et « femme », ils disent aussi, selon les mots d'Eliana Magnani « les ressorts d'une société émaillée par des rapports de subordination ».
- 21 Subordination sociale, mais aussi subordination documentaire : la communication de Coraline Rey a permis d'évoquer la circulation ou la non-circulation du lexique entre les différentes typologies de documents.

Fig. 7 – Tablette de cire contenant les comptes du boursier de Cîteaux (1321-1325). Archives départementales de la Côte-d'Or, 11 H 1165.

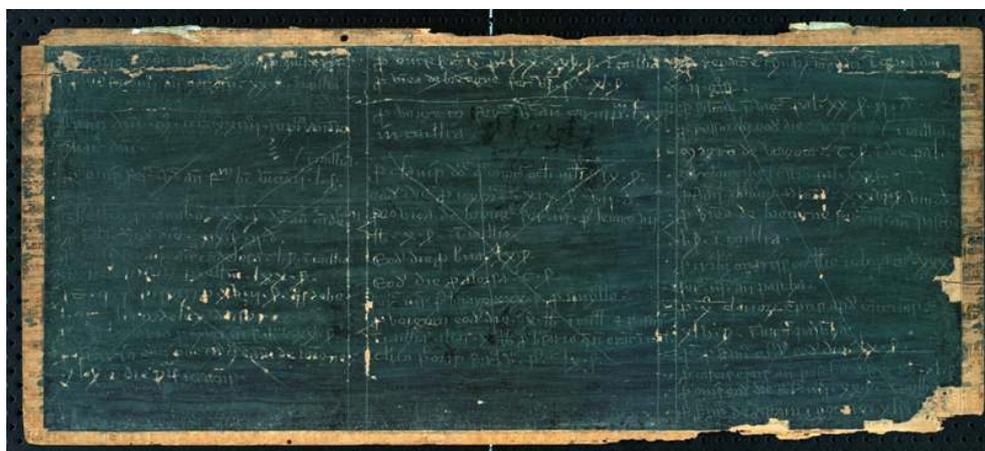
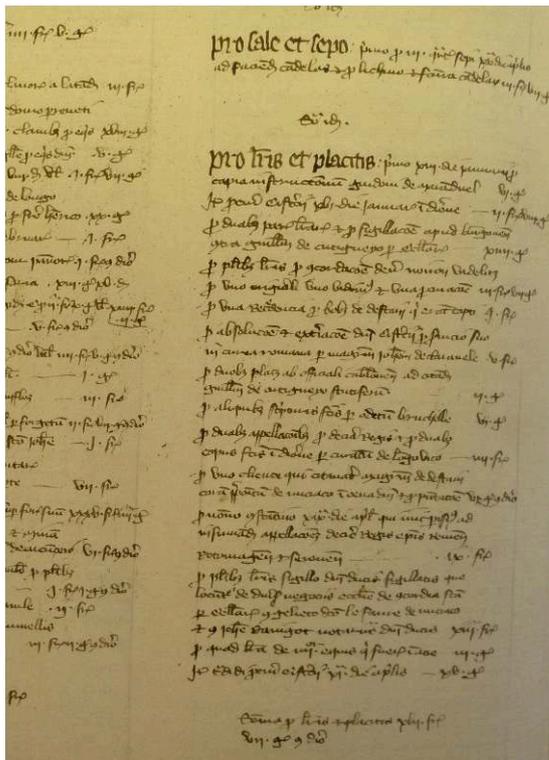


Fig. 8 – Rubrique "Pro litteris et placitis" dans les comptes du boursier (1337-1402). Archives départementales de la Côte-d'Or, 11 H 1160, f. 47v.



- 22 En prenant comme points d'observation non seulement les chartes ²⁴, mais également les comptes des XIV^e et XV^e siècles (fig. 7 et 8) ²⁵ et le catalogue de la bibliothèque du XV^e siècle ²⁶, elle a présenté les premiers éléments d'une enquête sur le lexique des actes et sur le vocabulaire du livre et de l'écriture. Pour apporter des points de comparaison sur le champ lexical des actes, les recherches ont été réalisées aussi bien dans les chartes de Cîteaux que dans la base des CBMA. Quatre termes ont notamment retenu l'attention : *carta*, *mandatum*, *instrumentum* et *littera*. Il est ainsi apparu que *carta* n'a été employée que jusqu'au XIII^e siècle, aussi bien dans la base des CBMA que dans les chartes de Cîteaux ²⁷. Ce constat amène à poser la question suivante : si le mot *carta* tend à disparaître partout au XIII^e siècle, par quel(s) autre(s) terme(s) a-t-il été remplacé ? *Mandatum* et *instrumentum* sont peu usités dans les chartes, mais apparaissent de manière régulière, jusqu'au XV^e siècle. Quant au mot *littera*, utilisé de manière récurrente jusqu'au XV^e siècle, il désigne des documents de natures différentes : la charte à proprement parler, une missive, des lettres de légation ou des lettres pontificales. La comparaison avec les termes et expressions employés dans la comptabilité a permis d'élargir la réflexion sur le lexique et d'envisager le domaine de la pratique de l'écrit. En effet, des mentions éparses d'activités scripturaires ont été retrouvées dans diverses rubriques des comptes : *pro pensionariis* – cette rubrique contient les noms des personnes recevant une pension, y compris les scribes et notaires –, *pro familia abbatie* – dépenses effectuées à l'abbaye même, pour le fonctionnement du bâtiment et des offices –, *pro litteris et placitis* (fig. 8), *pro diversis* – comprenant toutes les dépenses n'entrant pas dans les autres catégories, comme les achats de papier –, *pro capitulo generalis* – les dépenses pour le chapitre général sont tenues à part, on y trouve

donc les mentions de fournitures pour les activités scripturaires –, *pro emptione et reparacione librorum, pro factura cartularum et terrariorum...*

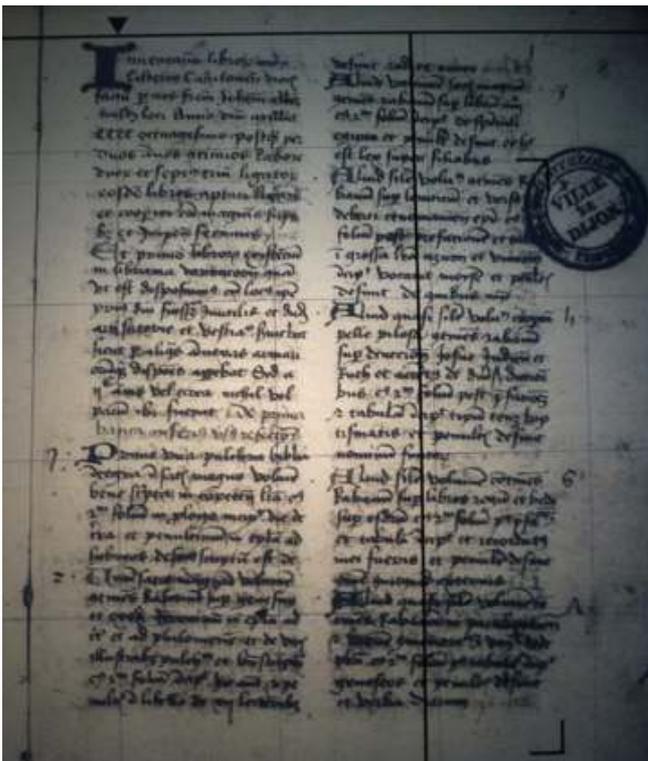
- 23 D’une manière générale, on observe, entre le XIV^e et le XV^e siècle, dans les comptes une diversification du vocabulaire désignant les actes (fig. 9).

Fig. 9 – Vocabulaire désignant les actes dans les comptes de l’abbaye de Cîteaux (XIV^e-XV^e siècle).

Liste non exhaustive des termes désignant les actes dans les comptes du XIV ^e siècle	Preceptum, copia, instructio, concordia, memoriale, supplicatio, vidimus, scriptura, quittancia, mandatum, littera
Liste non exhaustive des termes désignant les actes dans les comptes du XV ^e siècle	Grossatura, grossatione, grossa, mandatum, mandamentum, papirus, sententia, tractatum, littera, carta

- 24 Tous ces termes se trouvent au sein de rubriques « juridiques » – *pro litteris et placitis, pro scripturis* –, et chacun d’eux semble désigner des réalités différentes : par exemple, *littera* est une missive ; *mandatum* est utilisé dans le cadre de procès, sans doute pour convoquer les parties ; *papirus* est utilisé pour désigner une liste des personnes devant des cens à l’abbaye.
- 25 En ce qui concerne le vocabulaire du livre et de l’écriture, l’enquête menée dans les comptes et dans le catalogue de la bibliothèque a montré la spécialisation des annotations propres à ces différents types de registres (fig. 10).

Fig. 10 – Catalogue de la bibliothèque de Cîteaux, dit « de Jean de Cirey » (1480-1482). Dijon, Bibliothèque municipale, ms 610 [MIC 213].



- 26 Dans les comptes, comme il fallait s'y attendre, sont consignées les mentions d'achat de fournitures, essentiellement de papier, parchemin et encre. Cependant, certaines mentions donnent des renseignements sur le fonctionnement des ateliers scripturaires. Par exemple, pour l'année 1386-1387, il est écrit : *pro filo empto pellibus vicibus pro officio pelliparie*, ou encore *pro forpicibus fratris Johanni pelliparii actuendi* ²⁸. Dans le catalogue de la bibliothèque, on retrouve essentiellement le lexique du livre. Les notices descriptives des 1 200 ouvrages sont stéréotypées, sur le modèle : « un volume modeste contenant le livre X, dans une belle écriture, dont le second feuillet commence par... et l'avant-dernier finit par... ». Le vocabulaire reste diversifié quand il vise à qualifier les ouvrages, leur contenu ou leur aspect : *volumen, liber, biblia, concordia, summa, expositio, textus*... Il est à noter que dans le catalogue, à l'inverse de ce qu'on trouve dans les actes ou les comptes, *littera* est toujours employé dans le sens d'écriture.
- 27 Le vocabulaire utilisé à Cîteaux, mais aussi le niveau de langue mobilisé, semble ainsi dépendre du type documentaire. La diversification du lexique, assez marquée dès le XIII^e siècle, paraît liée en particulier au développement de l'appareil juridique et à la mise en place des nouvelles formes de mise en registre normées. Les documents recensant les livres utilisent un vocabulaire plus « littéraire », propre à rendre compte des artefacts désignés, tandis que les listes comptables contiennent un lexique plus « quotidien », pragmatique, peut-être plus proche du langage parlé ²⁹. Il s'agit donc d'un vocabulaire « cumulatif », car même si l'usage d'un mot se raréfie au fil du temps (carta par exemple), il ne disparaît jamais complètement. Ces observations incitent à pousser plus en avant la comparaison entre différents genres d'écrit et de prendre la mesure, en termes de recherche lexicale, des décalages introduits par la catégorie documentaire analysée.

Documentation

- 28 Comme chaque année, cette journée d'études fut aussi l'occasion de dresser le bilan de l'enrichissement de la base documentaire CBMA en 2012 : 1 200 nouveaux actes ont ainsi été enregistrés. Ils comprennent le cartulaire du chapitre cathédral de Sens ³⁰, les neuf premiers volumes des chartes de Saint-Étienne de Dijon ³¹, les chartes de l'abbaye de Cîteaux de 1098 à 1182 ³², le recueil des actes du prieuré de Saint-Symphorien d'Autun ³³. En outre, 14 nouveaux manuscrits en mode image sont désormais disponibles dans la base : 9 cartulaires de l'abbaye de Cîteaux, 3 cartulaires de Saint-Étienne de Dijon, le second cartulaire de Molesme et le cartulaire des dominicains de Dijon ³⁴.
- 29 Il faut également signaler que le projet CBMA est de mieux en mieux connu, indexé et soutenu. On mesure cette notoriété grâce aux remerciements toujours plus nombreux – en particuliers ceux qui arrivent de l'étranger suite à la mise en disposition de nouveaux corpus – ainsi que grâce à la réaction de sites Web qui proposent de signaler notre documentation et nos outils. Plus techniquement, ISIDORE, moteur d'accès aux données SHS, moissonne désormais les pages Web des CBMA ³⁵, permettant de les interroger en une seule requête, de rendre visibles, de diffuser ou de relier les données numériques entre elles.
- 30 Ainsi, le projet CBMA est désormais présent à quatre adresses différentes :
- le site Web d'Artehis et en particulier la page documentaire : [\n_blank](http://www.artehis-cnrs.fr/page-documentaire-CBMA) ;

- le serveur CBMA, hébergé par le TGE Adonis : <http://www.artehis-cbma.eu/> \n _blank ;
 - la collection CBMA sur Revues.org : <http://cem.revues.org/12172> \n _blank ;
 - le blog Laetus Diaconus : <http://laetusdiaconus.hypotheses.org/> \n _blank.
- 31 Indissociable de l'activité analytique à proprement parler – un point sur lequel l'équipe des CBMA tient de nouveau à insister –, le développement de la base se poursuivra au cours des années à venir, en particulier avec la numérisation de documents narratifs (chroniques et hagiographies). Nous sommes plus que jamais conscients que ce travail mené sur le contenu des actes repose sur une documentation longuement éditée par nos devanciers, qui se trouve ainsi, grâce à leur numérisation et à leur exploitation automatisée, (ré)valorisée et (ré)utilisée dans le cadre de la construction de nouveaux paradigmes. C'était déjà ce que présageait Robert Fossier en 1970, et à qui, en guise de conclusion, nous rendons hommage : « [...] le jour viendra vite où il sera permis d'opérer des relevés systématiques de termes, de les comparer, de les apprécier, et ainsi d'atteindre, avec toutes les réserves qu'imposent d'irréparables lacunes documentaires, le stade d'une exploitation quantitative des sources économiques ou sociales. Tant que la recherche historique n'y sera pas parvenue, tout essai restera timide et lacunaire, au niveau des hypothèses de travail [...] ³⁶. »

Rappel du programme de la journée d'étude

32 Matin

- Eliana Magnani et Marie-José Gasse-Grandjean : « Bilan des CBMA, 2011-2012. »
- Didier Panfili : « User du mot *dominus* dans les *scriptoria* méridionaux du VIII^e au début du XIII^e siècle. »
- Nicolas Perreaux : « Vocabulaire courant, vocabulaire fréquent, vocabulaire endémique ? Lecture spatiale d'un phénomène fréquentiel. »

Après-midi

- Alain Guerreau : « L'observation et l'analyse statistique de la fréquence et de la répartition d'un groupe de *n tokens*. »
- Eliana Magnani : « Le genre d'*Ego* : la désignation des disposants dans les documents diplomatiques bourguignons. »
- Coraline Rey : « Le vocabulaire de l'écriture et du livre à Cîteaux. Pour une approche comparée : chartes, comptes et catalogue de bibliothèque. »

Conclusions

Reçu : 2 mars 2013 – Accepté : 24 avril 2013

NOTES

1. E. MAGNANI et M.-J. GASSE-GRANDJEAN, « CBMA. *Chartae Burgundiae Medii Aevi*. VI. Les chartes bourguignonnes sous PhiloLogic », *Bucema* (<http://cem.revues.org/12595>; consulté le 26/02/2013). Nous remercions vivement Eliana Magnani et Marie-José Gasse-Grandjean de nous

avoir offert l’opportunité remarquable d’organiser cette journée d’étude, montrant à cette occasion, s’il en était encore besoin, l’ouverture et la dynamique du projet CBMA. Nous leur exprimons également toute notre gratitude pour leur relecture attentive.

2. Le logiciel est en cours de refonte complète, principalement grâce au travail de Richard Whaling, en langage Python : <https://github.com/rwhaling/libphilologic> (version 4 de Philologic, actuellement en phase alpha). Autres sites relatifs au projet Philologic : <https://sites.google.com/site/philologic3/> ; <http://www.lib.uchicago.edu/efts/ARTFL/philologic/> (consultés le 26/02/2013).

3. Le TGE Adonis est le Très Grand Équipement pour les sciences humaines et sociales, dont la mission principale est d’assurer l’accès et la préservation des données numériques : <http://www.tge-adonis.fr/> (consulté le 26/02/2013).

4. Ce constat avait déjà été réalisé par Robert-Henri Bautier, en 1971 : « Certains aussi concevaient qu’au-delà de l’étude des chancelleries particulières devait naître, ou plutôt renaître, une diplomatique comparative, condition de nouveaux développements de secteurs entiers des sciences historiques », in R.-H. BAUTIER, « La Commission internationale de diplomatique. Sa genèse. Son organisation. Son programme de travail », *Bibliothèque de l’École des chartes*, 129 (1971), p. 421-425, ici p. 421.

5. O. GUYOTJEANNIN, « “*Penuria scriptorum*” : le mythe de l’anarchie documentaire dans la France du Nord (x^e-première moitié du xi^e siècle) », *Bibliothèque de l’École des chartes*, 155 (1997), p. 11-44 ; autres développements à partir de ce point de réflexion initial : L. MORELLE, « Histoire et archives vers l’an Mil : une nouvelle “mutation” ? », *Histoire et archives*, 3 (1998), p. 119-141 ; *id.*, « Instrumentation et travail de l’acte : quelques réflexions sur l’écrit diplomatique en milieu monastique au xi^e siècle », *Médiévales*, 56 (2009), p. 41-74.

6. A. GUERREAU, « Pourquoi (et comment) l’historien doit-il compter les mots ? », *Histoire et mesure*, 4 (1989), p. 81-105 ; *id.*, « *Vinea* », in *Les historiens et le latin médiéval*, Paris, 2001, p. 67-73 ; *id.*, « “*Textus*” chez les auteurs latins du xii^e siècle », in L. KUCHENBUCH et U. KLEINE (éd.), “*Textus*” im *Mittelalter*, Göttingen, 2005, p. 149-178 ; A. GUERREAU-JALABERT, « *Spiritus et caritas*. Le baptême dans la société médiévale », in F. HÉRITIER-AUGÉ et E. COPET-ROUGIER, *La parenté spirituelle*, Paris, 1995, p. 133-205 ; *EAD.*, « Parole/Parabole : analyse d’un champ lexical et sémantique », in R.-M. DESSI et M. LAUWERS (éd.), *La parole du prédicateur*, Nice, 1997, p. 311-339 ; A. GUERREAU-JALABERT et B. BON, « *Pietas* : réflexions sur l’analyse sémantique et le traitement lexicographique d’un vocable médiéval », *Médiévales*, 42 (2002), p. 73-88.

7. I. ROSÉ, « Enquête sur le vocabulaire et les formulaires relatifs à la dîme dans les chartes bourguignonnes (ix^e-xiii^e siècle) », in M. LAUWERS (éd.), *La dîme, l’église et la société féodale*, Turnhout, 2012, p. 191-233 ; D. MÉHU, « *Locus, transitus, peregrinatio*. Remarques sur la spatialité des rapports sociaux dans l’Occident médiéval (xi^e-xiii^e siècle) », in *Construction de l’espace au Moyen Âge : pratiques et représentations*, Paris, 2006, p. 275-293 ; J. MORSEL, « Brief und Schrift. Überlegungen über die sozialen Grundlagen schriftlichen Austauschs im Spätmittelalter am Beispiel Frankens », in L. KUCHENBUCH et U. KLEINE (éd.), “*Textus*” im *Mittelalter...*, *op. cit.*, p. 285-321 ; J. DEMADE et J. MORSEL, « Les *Eigenleute* de Franconie aux xiii^e-xv^e siècles. Essai d’appréhension spatiale et sémantique d’une catégorie sociale malmenée », in P. H. FREEDMAN et M. BOURIN (dir.), *Forms of servitude in Northern and Central Europe*, Turnhout, 2005, p. 75-113 ; E. MAGNANI, « L’échange dans la documentation diplomatique bourguignonne : autour du vocabulaire des transferts », in P. DEPREUX et I. FEES (dir.), *L’acte d’échange, du viii^e au xii^e siècle/Tauschgeschäft und Tauschurkunde vom 8. bis zum 12. Jh.*, Limoges 11-13 mars 2010, à paraître en 2013 (numéro spécial d’*Archiv für Diplomatik*) ; *EAD.*, « *Uxor et femina*. Enquête sur la désignation des femmes dans les documents diplomatiques bourguignons (ix^e-xi^e siècle) », in J.-P. RENARD (dir.), *La place et le rôle des femmes dans l’histoire de Cluny : en hommage à Ermengarde de Blesle, mère de Guillaume le Pieux*, Brioude, 2013, p. 125-138 ; B. BON, « L’influence des langues romanes sur le latin médiéval : l’exemple de *planus* », in M. PÉREZ

GONZÁLEZ et E. PÉREZ RODRÍGUEZ (dir.), *Influencias léxicas de otras lenguas en el Latín medieval*, Valladolid, 2011, p. 21-32 (pour cet auteur, voir aussi ses collaborations notables avec Anita Guerreau-Jalabert, par exemple sur *pietas*, note 6) ; pour Ludolf Kuchenbuch, se reporter à l'importante réédition de certains de ses articles, récemment parue : L. KUCHENBUCH, *Reflexive Mediävistik : Textus - Opus - Feudalismus*, Frankfurt, 2012 ; M. LAUWERS, *Naissance du cimetière : lieux sacrés et terre des morts dans l'Occident médiéval*, Paris, 2005 (en particulier les pages 115 à 139) ; B.-M. TOCK, « Altare dans les chartes françaises antérieures à 1121 », in J. HAMESSE (dir.), *Roma, magistra mundi. Itineraria culturae medievalis. Mélanges offerts au Père L. E. Boyle à l'occasion de son 75^e anniversaire*, t. 2, Louvain-la-Neuve, 1998, p. 901-926 ; N. PERREAUX, « L'eau, l'écrit et la société (IX^e-XII^e siècle). Étude statistique sur les champs sémantiques dans les bases de données [C.B.M.A. et autres] », *Bucema*, 15 (2011), p. 439-449 (<http://cem.revues.org/index12062.html> ; consulté le 26/02/2013) ; *id.*, « Dynamique sociale et écriture documentaire (Cluny, X^e-XII^e siècle). Observations statistiques sur le champ sémantique de l'eau », in D. IOGNA-PRAT, M. LAUWERS, F. MAZEL, D. RUSSO, C. SAPIN et I. ROSÉ (éd.), *Cluny, le monachisme et la société au premier âge féodal (880-1050)*, Rennes, à paraître en 2013. Il va de soi que ces références bibliographiques ne sont que des exemples d'études portant sur ce « vocabulaire courant », et que la liste pourrait beaucoup s'allonger, y compris avec la bibliographie exhaustive de ces mêmes auteurs.

8. Cette approche se fonde bien entendu sur la lecture proposée maintenant depuis plusieurs décennies par Alain et Anita Guerreau, qui consiste à considérer que les termes les plus fréquents – en 2001, Alain Guerreau parle de « vocabulaire concret élémentaire » –, trop souvent considérés comme transparents, sont en fait porteurs d'un sens massif pour la société médiévale. A. GUERREAU, *L'avenir d'un passé incertain. Quelle histoire du Moyen Âge au XXI^e siècle ?*, Paris, 2001.

9. Quelques titres classiques parmi une abondante bibliographie : J. VIELLIARD, *Le latin des diplômes royaux et chartes privées de l'époque mérovingienne*, Paris, 1927 ; J. MONFRIN, *Études de philologie romane*, Genève, 2001 ; A. UDDHOLM, « Les traits dialectaux de la langue des actes mérovingiens et le formulaire de Marculf », *Archivum latinitatis medii aevi*, 25 (1955), p. 47-69 ; P. AEBISCHER, *Études de stratigraphie linguistique*, Bern, 1978 : c'est sans doute ce dernier auteur qui a le plus marqué la mise en place du questionnaire scientifique de cette journée, son approche singulière mêlant étude de termes dits « courants » et lecture spatiale (cf. ses travaux sur *bosco*, *ecclesia*, *molendinum*, etc.).

10. O. GUYOTJEANNIN (dir.), *La langue des actes*, Paris, 2004.

11. <http://www.glossaria.eu/> (consulté le 26/02/2013). B. BON, « OMNIA – Outils et Méthodes Numériques pour l'Interrogation et l'Analyse des textes médiolatins », *Bucema*, 13 (2009), p. 291-292 (<http://cem.revues.org/11086>) ; 14 (2010), p. 251-252 (<http://cem.revues.org/11566>) ; 15 (2011), <http://cem.revues.org/12015> (consultés le 26/02/2013).

12. Un pourcentage précis n'est pas indiqué ici car le vocabulaire suit une distribution de type LNRE (*Large Number of rare Events* ; voir H. BAAYEN, *Word Frequency Distributions*, Dordrecht, 2001), évoluant en fonction de la longueur du texte, c'est-à-dire de la taille du stock de lemmes considéré.

13. Ce corpus a été réalisé dans le cadre d'une thèse de doctorat en cours à l'université de Bourgogne, sous la direction d'Eliana Magnani et de Daniel Russo, intitulée *L'écriture du monde. Perception, catégorisation et appropriation de l'environnement dans les sources numérisées du Moyen Âge (VII^e-XII^e siècle)*. Les tentatives pour extraire les lemmes « typiques » des zones ainsi formées n'ayant pas donné de résultats probants, il faut conclure que ce sont les associations entre les termes/lemmes qui forment les spécificités régionales, et non pas les lemmes isolés (cf. ci-dessus l'hypothèse 3).

14. Afin d'éviter les biais (voir le résumé d'Alain Guerreau ci-dessous), le corpus a été découpé en 11 paquets d'environ 440 documents (à ± 4 %), pour la période 770-1079.

15. H. ELHAÏ, *Biogéographie*, Paris, 1968.

16. Résumé de la communication dans E. MAGNANI et M.-J. GASSE-GRANDJEAN, « CBMA. *Chartae Burgundiae Medii Aevi*. VI. Les chartes bourguignonnes sous PhiloLogic. Dijon, Maison des sciences de l'homme, 27 janvier 2012 », *Bucema*, 16 (2012) : <http://cem.revues.org/12312> (consulté le 26/02/2013).
17. Didier Panfili a d'ores et déjà organisé deux réunions de réflexion commune autour du terme *campus*, l'une à Paris, au LAMOP, en octobre 2010 ; l'autre à Llérída, Catalogne, en novembre 2012.
18. J. TRIER, *Der Deutsche Wortschatz im Sinnbezirk des Verstandes : Die Geschichte eines sprachlichen Feldes*, t. 1 (*Von den Anfängen bis zum Beginn des 13 Jahrhunderts*), Heidelberg, 1931 ; *id.*, « Über Wort- und Begriffsfelder », in L. SCHMIDT (éd.), *Wortfeldforschung : zur Geschichte und Theorie des sprachlichen Feldes*, Wissenschaftliche Buchgesellschaft, Darmstadt, 1973, p. 1-38.
19. En particulier dans le domaine des logiciels dédiés à la fouille de données textuelles, avec deux projets qui, outre Philologic, se démarquent nettement : TXM (<http://textometrie.ens-lyon.fr/> ; consulté le 27/02/2013) et CWB/CQP (<http://cwb.sourceforge.net/> ; consulté le 27/02/2013).
20. Bibliothèque R, développée par Stefan Evert et Marco Baroni, disponible à l'adresse suivante : <http://zipfr.r-forge.r-project.org/> (consulté le 26/01/2013). S. EVERT et M. BARONI, « ZipfR : Word frequency distributions », in R. *In Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics, Posters and Demonstrations Session*, Prague, 2007.
21. Le coefficient de Dice est une mesure de similarité, dont le but initial est de comparer la proximité de deux échantillons statistiques (http://en.wikipedia.org/wiki/Dice%27s_coefficient \t "_blank", consulté le 28/02/2013). Compris entre 0 et 1, il n'est pas très différent de l'indice de Jaccard (http://en.wikipedia.org/wiki/Jaccard_index \t "_blank", consulté le 28/02/2013).
22. E. MAGNANI, « *Uxor et femina...* », *op. cit.*
23. Les chartes de l'abbaye de Cluny représentent cependant 60 % de l'ensemble du corpus considéré ici.
24. J. MARILLIER, *Chartes et documents concernant l'abbaye de Cîteaux (1098-1182)*, Rome, 1961 : transcription de 250 chartes et autres documents, des origines à 1182. C. REY, *L'entreprise archivistique de Jean de Cirey, abbé de Cîteaux (1476-1501). Le dossier documentaire de la seigneurie de Villars, en Côte-d'Or*, mémoire de master sous la direction d'Elia Magnani et de Marie-José Gasse-Grandjean, Dijon, 2009 : transcription de 62 actes concernant la localité de Villars, dont 61 originaux, de 1182 à 1495.
25. Dijon, Archives départementales de la Côte-d'Or [ADCO], 11 H 1165, comptes du boursier sur tablettes de cire (1321-1325) ; 11 H 1160, comptes de l'Ordre et comptes du boursier (1337-1402) ; 11 H 1166, comptes du boursier (1489-1493). Ces dernières années, les recherches sur les comptabilités ont également porté sur les documents eux-mêmes et sur les conditions de leur production, constituant ainsi une codicologie, une « diplomatique des comptabilités » : O. GUYOTJEANNIN, « Préface », in P. BECK, *Archéologie d'un document d'archives. Approche codicologique et diplomatique des chartes de feux bourguignonnes (1285-1543)*, Paris, 2006, p. 5. La revue *Comptabilité(S)* a d'ailleurs consacré deux numéros à la comptabilité médiévale, l'un sur l'approche codicologique de ces documents : *Approche codicologique des documents comptables du Moyen Âge*, in *Comptabilité(S)*, n° 2, 2011 (<http://comptabilites.revues.org/364> ; consulté le 26/02/2013) ; l'autre sur le vocabulaire des comptabilités médiévales, *Le vocabulaire et la rhétorique des comptabilités médiévales*, in *Comptabilité(S)*, n° 2, 2012 (<http://comptabilites.revues.org/838> ; consulté le 26/02/2013).
26. DIJON, *Bibliothèque municipale*, ms. 610 [MIC 213], 1480-1482, catalogue dit « de Jean de Cirey ». Ce catalogue a été publié dans le *Catalogue général des manuscrits des bibliothèques publiques de France*, t. 5, Dijon, Paris, 1889, p. 339-452. Yolanta Załuska et Marie-Françoise Damangeot ont étudié le scriptorium et la bibliothèque de Cîteaux, donnant également quelques analyses de ce catalogue : Y. ZAŁUSKA, *L'enluminure et le scriptorium de Cîteaux au XII^e siècle*, Cîteaux, 1989, plus

particulièrement p. 17-20 ; Y. ZAEUSKA (dir.), *Manuscrits enluminés de Dijon*, Paris, 1991 ; M.-F. DAMONGEOT, « La bibliothèque de l'abbaye au Moyen Âge », in M. PLOUVIER et A. SAINT-DENIS (dir.), *Pour une histoire monumentale de l'abbaye de Cîteaux*, 1098-1998, p. 232-245, plus particulièrement p. 236-242.

27. Des analyses statistiques s'avèrent nécessaires afin de vérifier ces hypothèses puisque la base des CBMA reste soumise à divers biais : surreprésentation de certains corpus (tel celui des chartes de l'abbaye de Cluny), limites chronologiques d'autres corpus (les chartes de l'abbaye Saint-Bénigne de Dijon, antérieures à 1300, pour ne citer qu'un exemple), majorité d'actes en latin.

28. Cette mention d'un office de pelleterie et d'un pelletier (qualifié aussi de maître pelletier l'année suivante) atteste l'existence de cette activité à l'abbaye. Cependant, les mentions de ce type étant rares, il n'est pas possible de dire, dans l'état actuel des recherches, si cet office fonctionnait de manière régulière.

29. À titre d'exemple, les cisterciens consomment régulièrement des *clavi de girofle* ainsi que du *gizimbre*

30. *Cartulaire du chapitre de Sens*, éd. E. CHARTRAIRE, Sens, 1904. Nous remercions à nouveau Gérard Moysse, directeur des archives départementales de la Côte-d'Or, et Frédéric Petot, photographe, pour leur collaboration très efficace.

31. Éditées dans des thèses de droit de l'université de Bourgogne, Dijon, 1905-1914.

32. J. MARILIER, *Chartes et documents...*, *op. cit.*

33. *Recueil des actes du prieuré de Saint-Symphorien d'Autun de 696 à 1300*, éd. A. DÉLÉAGE, Autun, 1936.

34. Cartulaires de Cîteaux : ADCO, 11 H 64 (cartulaire 166), 11 H 69 (cartulaire 172), 11 H 68 (cartulaire 171), 11 H 70 à 11 H 78 (cartulaires 185 à 193) ; cartulaires de Saint-Étienne de Dijon : ADCO, G 126 (cartulaire 22), G 128 (cartulaire 36), G 137 (cartulaire 24) ; le second cartulaire de Molesme : ADCO, 7 H 7 (cartulaire 143) ; cartulaire des Dominicains de Dijon : ADCO, H 53 (cartulaire 221). Tous ces cartulaires sont disponibles en ligne : <http://www.artehis-cnrs.fr/page-documentaire-CBMA> \n _blank (consulté le 26/02/2013). Nos remerciements vont également à Anthony Dumontet, infographiste Artechis, pour son travail de mise en ligne rigoureux et rapide.

35. <http://www.rechercheisidore.fr/> (consulté le 28/02/2013).

36. R. FOSSIER, *Histoire sociale de l'occident médiéval*, Paris, 1970, p. 12.

AUTEURS

NICOLAS PERREAUX

Université de Bourgogne, UMR Artechis

CORALINE REY

Université de Bourgogne, UMR Artechis