



HAL
open science

Histoire et perspectives du Novum Glossarium Mediae Latinitatis. Un “ Nouveau Du Cange ”

Bruno Bon

► **To cite this version:**

Bruno Bon. Histoire et perspectives du Novum Glossarium Mediae Latinitatis. Un “ Nouveau Du Cange ”. *Archivum Latinitatis Medii Aevi*, 2015, 73, pp.297-308. halshs-01895053

HAL Id: halshs-01895053

<https://shs.hal.science/halshs-01895053v1>

Submitted on 24 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



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

BRUNO BON

Histoire et perspectives du *Novum Glossarium Mediae Latinitatis**

Un « Nouveau Du Cange »

Le Novum Glossarium en 1924

A l'issue de la Première guerre mondiale, quand l'Europe cherchait à se donner les moyens de préserver la paix, la fondation de l'Union Académique Internationale (UAI) en 1919 fut un acte politique autant qu'intellectuel. Cette fédération d'académies était destinée à encourager la coopération scientifique et intellectuelle entre les peuples déchirés par le conflit le plus meurtrier de l'Histoire. Depuis sa première assemblée générale en 1920, de nombreuses entreprises collectives ont été régulièrement lancées, l'UAI en dirigeant encore aujourd'hui plus de quatre-vingts.

La cinquième entreprise de l'UAI, proposée dès 1920, fut inscrite en janvier 1924 sous le titre : « Dictionnaire du Latin Médiéval ». Il s'agissait de reprendre une idée lancée en 1913 au Congrès historique international de Londres, visant la refonte du vénérable (et non remplacé) *Glossarium Mediae et Infimae Latinitatis* de Charles Du Cange, paru en 1678 et considérablement augmenté par la suite. Au terme de trois années de discussions – pour le détail desquelles on consultera avec profit les divers comptes rendus¹ publiés dans la revue *Archivum Latinitatis Medii Aevi (ALMA)* associée dès l'origine à l'entreprise – il fut finalement décidé de doter les études médiévales, non d'un « Nouveau Du Cange » à proprement parler, mais plutôt d'un instrument lexicographique comparable au

* Cette étude a bénéficié du soutien à la mobilité internationale du CNRS.

1 Voir en particulier Langlois (1924), repris par Lefèvre (1975).

*Thesaurus Linguae Latinae*² en cours de rédaction à Munich pour les études classiques. Au fil du temps, le futur *Novum Glossarium Mediae Latinitatis* (NGML – Blatt *et al.* 1957³) est ainsi devenu un véritable dictionnaire de la langue latine médiévale, à vocation historique et généraliste, élaboré à nouveaux frais selon des critères scientifiques modernes, et se donnant pour objectif de rendre compte de tous les usages du latin dans l'ensemble de l'Europe des IX^e-XII^e siècles.

Première étape indispensable au développement d'un tel projet lexicographique, la campagne de dépouillement des sources fut organisée selon les principes fondateurs de l'UAI, par la coopération internationale entre les divers Comités nationaux, coordonnés par le « Comité Du Cange » à Paris. Chaque académie partenaire était chargée de rassembler le matériau issu de la documentation médiolatine produite sur son territoire, selon une amplitude chronologique qui reflétait l'histoire locale de l'usage de cette langue, et qui ne recouvrait parfois que très partiellement la période centrale retenue pour le NGML⁴. Pour tenir compte de cette double diversité géographique et chronologique, les différents pays furent enfin invités à rédiger leur propre dictionnaire à partir de leurs propres sources. Il s'est ainsi progressivement constitué un véritable réseau international de lexicographie médiolatine, autour du NGML et de l'UAI⁵.

Le Novum Glossarium en 2014

Le « Dictionnaire du latin médiéval » de l'UAI n'est donc pas un objet unique. Il est désormais constitué de la publication parallèle d'un dictionnaire international (le NGML), et de l'ensemble des

2 Derniers fascicules parus : *N-Navalis* et *R-Rarus*.

3 Dernier fascicule paru : *Plana-Plego*.

4 A titre d'exemple, les dictionnaires d'Italie (Arnaldi/Smiraglia 1970), des Pays-Bas (Fuchs *et al.* 1977) et de Suède (Westerbergh/Odelman 1968) couvrent respectivement les périodes 476-1022, 800-1500 et 1150-1530.

5 Ce réseau regroupe aujourd'hui l'Allemagne, la Belgique, la Castille, la Catalogne, le Danemark, la France, la Grande-Bretagne, la Hongrie, l'Irlande, l'Italie, les Pays-Bas, la Pologne, la Suède et la République Tchèque.

dictionnaires nationaux⁶, tous également dotés d'une vocation « universelle » : le premier dans la fourchette chronologique de 800 à 1200, les seconds sur la zone géographique qu'ils recouvrent. La décision de rédiger des dictionnaires nationaux, qui aurait pu remettre en cause l'existence du dictionnaire international, s'est inscrite au contraire dans un souci de complémentarité. Le travail de rédaction promettant d'être particulièrement long, les instruments nationaux étaient invités à commencer par la lettre A, pendant que le *NGML* allait passer directement à la lettre L : ainsi, espérait-on, l'alphabet serait-il couvert plus rapidement par l'un ou l'autre des dictionnaires.

Plus de cinquante ans après la parution des premiers fascicules, la situation est très contrastée, tant en raison d'une répartition inégale de la documentation entre les différents pays partenaires de l'entreprise, que des écarts de moyens mis en œuvre. Si plusieurs dictionnaires nationaux sont achevés⁷, les ouvrages les plus importants sont toujours en cours de rédaction, à l'exception notable du *Dictionary of Medieval Latin from British Sources* dont le dernier fascicule est paru en décembre 2013. Ainsi, sur une grande partie de l'Europe carolingienne, la couverture de l'alphabet, encore disparate⁸, repose sur la complémentarité entre le *Mittellateinisches Wörterbuch*⁹ (*MLW*) pour le début et le *NGML* pour la fin de l'alphabet. Pour autant, le cadre commun de l'entreprise de l'UAI n'a pas empêché les différentes équipes de rédaction d'en infléchir notablement l'esprit, en vertu des intérêts particuliers de chacun des responsables. Les fascicules parus depuis le milieu du siècle dernier reflètent ainsi des

6 Cf. *infra* notes 7 et 8.

7 Finlande, ante 1530 (Hakamies 1958) ; Grande-Bretagne, a. 500-1600 (Latham *et al.* 1975) ; Pays-Bas, a. 800-1500 (Fuchs *et al.* 1977) ; Suède, a. 1150-1500 (Westerbergh/Odelman 1968) ; Yougoslavie, a. 800-1526 (Kostrenčić 1969).

8 Allemagne, a. 500-1280 (Lehmann *et al.* 1959, lettres A-I) ; Castille, a. 700-1230 (González 2010, *imperfectum*) ; Catalogne, a. 800-1100 (Climent *et al.* 1960, lettres A-D, F-G) ; Danemark, ante 1560 (Due/Terkelsen 1987, lettres A-R) ; Hongrie, a. 1000-1526 (Boronkai/Szovák 1983, lettres A-I) ; Irlande, a. 400-1200 (Harvey 1979, lettres A-H) ; Italie, ante 1022 (Arnaldi/Smiraglia 1970, *imperfectum*) ; Pologne, a. 1000-1600 (Plezia *et al.* 1958, lettres A-S) ; République Tchèque a. 800-1500 (Silagiová 1977, lettres A-M).

9 Dernier fascicule paru : *Hospitalarius-Illibezus*.

choix éditoriaux qui varient non seulement d'un dictionnaire à l'autre, mais aussi d'un fascicule à l'autre à l'intérieur d'un même dictionnaire¹⁰, certaines décisions originelles donnant lieu à un traitement différencié : équivalences graphiques et organisation alphabétique, référence au *Totius Latinitatis Lexicon* de Forcellini¹¹, maintien ou élimination des mots et des sens attestés en latin classique, *etc.* Il en résulte que, malgré leur complémentarité fondamentale, les instruments qui relèvent du « Dictionnaire du latin médiéval » sont profondément différents. Enfin, la décision d'assortir le *NGML* de dictionnaires nationaux a permis à son équipe de rédaction d'en adapter les objectifs pratiques aux circonstances : des priorités ont été définies au sein de la documentation médiolatine des IX^e-XII^e siècles, destinées à combler les principales lacunes identifiées par ailleurs. La rédaction du *NGML*, sans renier sa vocation internationale, fait désormais la part belle aux documents italiens et espagnols des XI^e-XII^e siècles, partiellement pris en compte dans les dictionnaires correspondants, et aux textes français de l'ensemble de la période, en l'absence de tout dictionnaire national pour ce pays¹².

Outre les autres dictionnaires scientifiques du réseau de l'UAI, le *NGML* ne dispense pas de consulter le *Glossarium Mediae et Infimae Latinitatis* de Du Cange (*DuC*), dont il n'est pas la refonte, et avec lequel il entretient des rapports ambigus. Il n'y a pas lieu de s'étendre ici sur la différence de nature entre un ouvrage du XVII^e siècle, qui relève autant du glossaire que de l'encyclopédie, et un instrument scientifique moderne. Nous attirerons seulement l'attention sur le fait que le *DuC* rend compte d'une documentation extrêmement riche, en grande partie manuscrite et parfois disparue de nos jours, tandis que le *NGML* se limite aux éditions imprimées, à de rares exceptions près¹³. Ensuite, le *DuC* est le seul instrument généraliste

10 On se reportera avec profit aux observations de Stotz (2002 : 193-242).

11 « Le 'Dictionnaire du latin médiéval' est la suite naturelle du Forcellini – De Vit et l'on ne perdra pas de vue ce point essentiel » (Instructions 1924 : 66).

12 L'attribution tardive du *NGML* à la France s'est soldée par la suspension (puis l'abandon) du projet de dictionnaire français du latin médiéval.

13 Il s'agit surtout de glossaires médiévaux majeurs, longtemps inédits, comme le *Glossarium alphabeticum e codice Bruxellensi* (Bruxelles, BR. II 1049) ou le *Liber derivationum* d'Ugutio (Paris, BnF. Lat. 14090). « Doch beschränken

disponible pour la langue latine de la fin du Moyen âge, et en particulier pour l'Europe occidentale des XIV^e-XV^e siècles. En effet, si les dictionnaires de l'Europe centrale et orientale dépassent généralement cette date, le *MLW* et le *NGML* s'arrêtent au début ou au cours du XIII^e siècle. Enfin, le *DuC* présente l'immense avantage d'être achevé, les mots de la fin de l'alphabet y trouvant une rare occasion d'être relevés. Pour toutes ces raisons, le *NGML* et le *DuC* sont des instruments fondamentalement complémentaires.

Le « Dictionnaire du latin médiéval » s'inscrit naturellement au nombre des entreprises modernes de lexicographie latine. Il entretient des rapports étroits avec les dictionnaires de latin antique, à commencer par le *Thesaurus Linguae Latinae*, dont il prétend prendre la suite et qui sert de modèle à certaines équipes de rédaction, et avec les lexiques spécialisés du latin médiéval, parmi lesquels le *Lexicon Minus* de Niermeyer pour les textes diplomatiques, et le *Dictionnaire des auteurs chrétiens* de Blaise pour le latin patristique. Pour autant, on ne mettra jamais assez en garde contre l'utilisation d'un dictionnaire de latin classique pour tenter de comprendre un texte médiéval. Car derrière la continuité formelle apparente de la langue, se cache une évolution sémantique profonde, à la mesure du bouleversement des structures de la société qui l'emploie¹⁴.

Le renouveau du *Novum Glossarium Mediae Latinitatis*

La rédaction du Novum Glossarium

Pour mener à bien le grand projet dont elle a la charge, l'équipe de

sich die meisten W[örterbuch]unternehmungen aus guten Gründen auf die Auswertung gedruckter Texte. Viele Belege aus handschriftlichen Texten, namentlich Archivalien, finden sich bei DuC [...]. Beim NG[ML] war lange Zeit die einzige ins Gewicht fallende Ausnahme das lexikographische Werk des Ugutio » (Stotz 2002 : 259).

14 Voir à ce sujet l'analyse des mots *Pietas* (Bon/Guerreau-Jalabert 2002), *Thesaurus* (Guerreau-Jalabert/Bon 2010) et *Planus* (Bon 2011).

rédaction du *NGML* est confrontée depuis l'origine à deux difficultés majeures. Rendre compte de l'ensemble de la production latine européenne sur une période de quatre siècles supposait d'abord la capacité à en extraire le matériau lexicographique. Sans moyens propres, et dans le cadre d'une coordination scientifique perturbée par les événements politiques de la première moitié du XX^e siècle, l'identification et la répartition des sources, puis leur dépouillement ont donné lieu à l'empirisme le plus total. Il faut dire que le latin médiéval, dont prétendait s'occuper le Comité Du Cange, n'était pas encore reconnu par la communauté scientifique comme une langue à part entière¹⁵. Il s'ensuit que les instruments manquaient pour ordonner sérieusement les choses, à commencer par une bibliographie générale des œuvres considérées¹⁶. Le soin d'établir les listes de sources à dépouiller fut donc confié à chaque Comité national, ce qui conduisit inévitablement à des doublons (plutôt qu'à des lacunes), dans un contexte politique favorable aux réflexes nationalistes. Quant au dépouillement lui-même, qui visait à transformer les textes en matériau lexicographique, il fut principalement réalisé dans l'Entre-deux Guerres par une multitude de collaborateurs le plus souvent bénévoles, selon des principes décidés à Paris en 1924¹⁷. Le résultat de ce travail, qui se trouve en grande partie rassemblé dans le fichier central de Paris¹⁸, et qui sert de base historique à la rédaction du *NGML*, est à la mesure de cette organisation : chaque dépouilleur étant invité à s'arrêter, au cours de la lecture, sur les mots et les sens qui lui semblaient intéressants, sans aucune contrainte préalable, il en ressort évidemment que le nombre et la nature des attestations relevées dépendent étroitement tant des connaissances et des intérêts, que de l'investissement et de la disponibilité de chacun. Un fichier constitué dans ces conditions ne permet à peu près aucune autre conclusion que

15 « De manière assez continue depuis le XIX^e siècle, les observations sur le latin médiéval se ramènent à un 'discours de la décrépitude' » (Alexandre *et al.* 2013 : 3).

16 Dont l'attribution, la datation, ou la localisation sont pourtant souvent difficiles à reconstituer.

17 Cf. Instructions (1924).

18 Conservé par l'Académie des Inscriptions et Belles-Lettres, à l'Institut de France, où il peut être consulté auprès du Comité Du Cange.

celle-ci : un mot qui s'y trouve a probablement existé¹⁹. En revanche, rien ne peut être déduit de tous les manques qu'on y relève : le traitement des mots très fréquents est à cet égard édifiant²⁰.

Étant entendu que le corpus des textes médiolatins, clos par nature, est en pratique ouvert, de nouveaux textes et nouvelles éditions paraissant tous les ans, l'équipe du *NGML* doit aujourd'hui faire face à plus de dix mille références bibliographiques, de tous types de textes, dont sont extraites plus de deux millions de fiches lexicographiques. Pour s'en occuper, et c'est la deuxième difficulté, les forces humaines sont depuis longtemps notoirement insuffisantes : un à deux rédacteurs sont régulièrement affectés par la France à l'élaboration de ce dictionnaire²¹, pendant que le nombre des collaborateurs du *MLW* dépasse la dizaine. Dans ce contexte, l'équipe du *NGML* a été amenée à suivre de près le développement des nouvelles technologies, dans l'espoir d'y trouver le moyen de compenser la faiblesse de ses effectifs. Bien sûr, l'informatisation précoce de la rédaction a considérablement amélioré le travail de bureautique. Dans le souci permanent d'enrichir le fichier de base du *NGML*, les rédacteurs ont même commencé à réaliser des fiches informatisées, qui étaient en réalité imprimées, pour être ensuite traitées manuellement. Mais après ce premier changement de support, est bientôt venu celui de la consultation des sources, à l'origine d'un bouleversement plus important encore. Les premiers dictionnaires et corpus de textes médiolatins numérisés furent développés par des éditeurs privés, dont l'objectif était purement commercial. En sont issus de grands ensembles de données accessibles contre une redevance élevée, et assortis de simples outils de recherche bibliographique traditionnelle et d'intérêt scientifique très limité. Pour autant, la disponibilité croissante de textes interrogeables relevant du *NGML* semblait devoir

19 Sous la réserve d'une confusion parfois attestée entre commentaire moderne et texte médiéval, notamment dans la grande collection des *Acta Sanctorum*.

20 Issu d'un dépouillement tendant à privilégier les mots et les sens inattendus, le fichier présente (en proportion, mais parfois absolument) très peu d'attestation des mots les plus fréquents, considérés (à tort) comme « évidents », et d'autant plus difficiles à traiter que le matériau manque.

21 La rédaction du *NGML* est actuellement confiée à la Section de Lexicographie latine de l'Institut de Recherche et d'Histoire des Textes (CNRS).

permettre de s'affranchir progressivement du fichier central et de tous ses défauts. L'idéal rêvé d'un dépouillement « exhaustif » des dix mille références bibliographiques allait devenir réalité. Malheureusement, il faut bien admettre que l'indigence des corpus textuels proposés à grands frais par les éditeurs privés ont rapidement douché cet espoir. Car si les outils d'interrogation de ces corpus offrent effectivement la possibilité de piocher un mot donné (ou plus exactement une forme donnée²²) dans tous ses contextes, le résultat n'en est pas toujours exploitable par le rédacteur d'un dictionnaire : que faire des soixante mille occurrences de la forme *pater*²³ dans la *Patrologie Latine*²⁴ ? En l'état actuel de leur développement, ces corpus propriétaires peuvent donc éventuellement servir, sous certaines conditions, à compléter le fichier central, mais absolument pas à le remplacer.

Révolution (numérique) au Comité Du Cange

Persuadée que, malgré l'absence de débouchés commerciaux, les nouvelles technologies pouvaient servir autant à l'étude du latin médiéval qu'à celle des langues contemporaines, l'équipe du *NGML* s'est finalement engagée dans sa propre révolution numérique, et dans deux directions. La première préoccupation, la plus facile à concevoir, concernait le traitement des articles déjà publiés. La numérisation rétrospective d'une vingtaine de fascicules du *NGML*, imprimés dans la seconde moitié du XX^e siècle sur un papier de qualité, devait poser d'autant moins de difficultés que la puissance des matériels progressait rapidement. En revanche, il semblait indispensable de la compléter par un instrument qui permît de couvrir l'intégralité de l'alphabet. C'est ainsi que fut décidée la numérisation parallèle du

22 Cf. *infra* note 33.

23 Il s'agit ici des nominatif et vocatif singulier ; pour traiter le lemme *Pater* (que nous différencions par une capitale), il conviendrait de pouvoir tenir compte de toutes les formes de ce mot, soit plus de deux cent mille occurrences !

24 Version électronique (payante) : <<http://pld.chadwyck.co.uk>>. Les bases de données de la plate-forme éditoriale <<http://www.brepolis.net>> présentent les mêmes caractéristiques.

NGML et du *DuC*, tombé depuis longtemps dans le domaine public, mais dont les caractéristiques techniques étaient beaucoup moins ordonnées. Car l'objectif de cette opération ne consistait pas en un simple changement de support, pour lequel des images indexées auraient pu suffire. Il s'agissait au contraire de fournir un matériau exploitable par des procédures automatisées, ce qui impliquait à la fois la reconnaissance des caractères (OCR) et l'encodage des données lexicographiques. Mais le *NGML* n'était pas seul : dès la réunion du réseau de l'UAI à Bruxelles en 2000, le Comité Du Cange attirait l'attention de ses collègues sur le double intérêt d'une numérisation de l'ensemble du « Dictionnaire du latin médiéval », et d'une forme d'harmonisation en permettant l'interrogation croisée. Quatre ans plus tard, à Barcelone, il proposait de s'accorder sur une liste minimale de huit champs de requête communs, et sur le principe d'un encodage XML²⁵. Les bases d'une éventuelle collaboration à venir étant ainsi posées, il restait encore un instrument à traiter : l'*Index* des sources²⁶. En effet, dans un dictionnaire qui accumule les sigles d'auteurs et de textes, leur liaison automatisée avec la bibliographie est un progrès majeur, que le seul encodage des articles ne permet pas d'atteindre. Il est donc revenu à l'équipe du *NGML* de convertir la liste de ses dix mille références en base de données SQL. A cette occasion, une réflexion s'est engagée sur l'intérêt d'en compléter l'indexation, jusque-là limitée à de parcellaires indications chronologiques. Le transfert technologique rendait ainsi possible, au-delà des méthodes d'interrogation, un enrichissement des données scientifiques elles-mêmes²⁷. Sur ce point, en revanche, il n'a pas été possible de rapprocher le *DuC* et son successeur, car l'ouvrage ancien n'est pas assorti d'un index bibliographique à proprement parler. Les sources y sont identifiées par des indications très variables d'un article à l'autre,

25 Comptes rendus de la réunion de Bruxelles dans Bon (2000), et de celle de Barcelone : « Les champs communs à prévoir [...] seront les suivants : lemme, variantes graphiques [...], catégorie grammaticale, étymon, étymologie, domaine [...], traduction moderne, exemples, références [...], bibliographie » (Heid 2004 : 331-332).

26 La liste des éditions citées dans le *NGML* est régulièrement révisée pour tenir compte des nouveaux textes ; dernier supplément : Bon (2005).

27 Cf. *infra* note 37.

et censées se suffire à elles-mêmes. Qui plus est, beaucoup d'entre elles correspondent à des manuscrits, dont la cotation a pu changer. La numérisation rétrospective du *NGML* et de son prédécesseur n'ont donc pas abouti à un résultat strictement homogène²⁸.

En même temps qu'elle s'occupait de valoriser les résultats du passé, l'équipe du *NGML* s'est préoccupée de l'avenir de son activité. Si la publication électronique s'est finalement imposée comme une perspective naturelle et souhaitable pour les dictionnaires médiolatins en cours de rédaction, le maintien d'une version imprimée est encore très largement défendu. Il faut reconnaître que l'exemple hybride et original du *Dictionary of Medieval Latin from Celtic Sources* (Harvey 1979)²⁹ n'a pas particulièrement incité les collègues à se lancer dans la publication numérique. D'autre part, la maîtrise des nouvelles technologies est encore peu répandue dans les instances responsables de nos programmes scientifiques, au contraire d'un fort attachement à la tradition et à l'imprimé. Pourtant, la question de l'abandon du papier n'est pas neutre, car elle implique de nombreux paramètres, qui ne valent pas pour tous les cas. Il convient en effet de différencier entre deux états d'avancement de nos dictionnaires : certains, comme le *MLW* ou le *NGML*, sont encore très loin d'être théoriquement achevés, pendant que d'autres, comme le *Lexicon Mediae et Infimae Latinitatis Polonorum (LMILP – Plezia et al. 1958)*, couvriront l'ensemble de l'alphabet dans un avenir relativement proche. Pour ces derniers, la poursuite de la publication imprimée, sans préjudice d'une édition numérique parallèle, est éminemment souhaitable, puisque l'ouvrage aujourd'hui très avancé sera bientôt complet. Pour les premiers, la situation dépend des forces et délais mis en œuvre pour leur achèvement. Le cas du *MLW* est ambigu, dans la mesure où l'échéance du financement fixée par les autorités de tutelle est trop proche pour permettre d'en achever la rédaction. Dans ces conditions, même si le temps consacré à la mise en page pourrait être économisé dans le cadre d'une publication numérique, le maintien de l'édition

28 Consultable librement sur le site : <<http://www.glossaria.eu>>.

29 Le dictionnaire irlandais a été rédigé à partir de la base de données des textes irlandais *Archive of Celtic Latin Literature*, dont il reprend les codes bibliographiques (et non les textes !), classés par sens. La version imprimée ne peut donc être utilisée qu'en consultant parallèlement la base textuelle.

imprimée est de nature à renforcer l'unité d'un projet inachevé. Quant au *NGML*, qui souffre de la démesure entre ses objectifs et ses moyens, la question mérite d'être sérieusement posée. La position actuelle de l'équipe tente de ménager l'avenir : il convient de prévoir, lors de la préparation des articles, la possibilité d'une double publication. Or cet art de l'entre-deux demeure très théorique, car les contraintes de l'impression papier sont telles, tant pour le fond (taille des fascicules, architecture des articles, ordre alphabétique) que pour la forme (jeux de caractères limités, impression en noir & blanc), qu'il est impossible pour le rédacteur de s'en abstraire réellement. Au contraire, l'encodage XML, indispensable à la version numérique, est inutile à la publication imprimée, et donc confié *a posteriori* au soin de filtres XSLT. Pour développer une édition électronique du *NGML* digne de ce nom, fondée sur l'encodage natif des données lexicographiques, capable d'intégrer les remords et (partiellement) déliée des contraintes de l'ordre alphabétique, il faudra donc en passer, sinon par une suspension (plutôt qu'un abandon) de la version imprimée, du moins par une refonte de cette dernière, qui devra présenter une photographie de la version numérique, et non plus l'inverse. Et pour ce faire, le choix de l'interface de rédaction XML n'est pas très important, pourvu qu'elle soit efficace et que son produit en soit libre et gratuit. Car le *NGML*, instrument de recherche réalisé sur fonds publics, est ainsi livré sans restriction à la communauté scientifique, à charge pour elle, en l'utilisant et en l'amendant, de lui donner un sens.

Vers un Nouveau « Nouveau Du Cange »

Une opportunité à saisir

Si, comme nous l'avons vu et comme son nom l'indique, le *NGML* était un instrument de la modernité en 1924, il fallait trouver des moyens pour qu'il le demeurât en 2014. Le développement européen

du financement de la recherche sur appels à projets en a donné l'occasion, grâce à la conjonction de plusieurs opérations. Le programme européen COST – Medioevo Europeo³⁰, qui s'est donné pour objectif de mettre en œuvre un réseau scientifique au service des études médiévales, comprend un groupe de travail sur les dictionnaires et corpus textuels, au sein duquel se retrouvent plusieurs équipes du « Dictionnaire du latin médiéval » de l'UAI, et particulièrement le *NGML* et le *LMILP*. Mais, le financement direct de la recherche n'étant pas prévu dans le cadre des projets européens COST, il convenait, pour ne pas en rester indéfiniment au stade de la discussion, de s'appuyer sur d'autres sources de financement. En France, l'équipe du *NGML* avait ainsi obtenu le soutien de l'Agence Nationale de la Recherche pour le programme « Omnia »³¹, consacré à la numérisation du *DuC* et du *NGML*, et au développement d'outils d'analyse statistique du latin médiéval. De leur côté, les collègues du *LMILP* ont remporté deux appels à projets nationaux, pour la numérisation du dictionnaire et la création d'un corpus textuel du latin médiéval de Pologne³². Forte de ces contributions, la collaboration entre les deux équipes n'a dès lors cessé de se renforcer, au service d'un idéal commun : une lexicographie médiolatine profondément renouvelée grâce à l'exploitation statistique des textes numérisés.

Le latin médiéval, comme le latin classique, est une langue déclinée à forte variation morphologique, à laquelle s'ajoute, bien davantage encore que dans l'Antiquité, une très grande variabilité purement graphique³³. Toute étude lexicale et sémantique sérieuse en latin médiéval implique donc le regroupement préalable des diverses formes d'un même mot sous son lemme. Or, aussi surprenant que cela puisse paraître, et sans doute à cause de l'ampleur de la tâche à

30 Coordonné par l'Institut de Recherche et d'Histoire des Textes (CNRS, Paris) et la Sismel (Firenze) : <<http://www.medioevoeuropeo.org>>.

31 Outils et Méthodes Numériques pour l'Interrogation et l'Analyse des textes médiolatins : <<http://www.glossaria.eu>>.

32 'Elektroniczny słownik łaciny średniowiecznej na ziemiach polskich', et 'Korpus języka łacińskiego na ziemiach polskich (1000-1550)' : <<http://www.scriptores.pl>>.

33 Pour une étude sur la variabilité graphique et morphologique du latin médiéval, voir Alexandre *et al.* 2013.

réaliser, les corpus textuels propriétaires évoqués plus haut³⁴ ont systématiquement manqué cette étape. Voici donc de très grands ensembles de données lexicales brutes, certes interrogeables par proximité ou troncature, mais parfaitement inadaptés à la moindre enquête statistique. Il faut reconnaître que, si la mise au point d'un outil de lemmatisation pour l'anglais ou le français pouvait trouver un public suffisant pour justifier un investissement, le latin médiéval était perdu. C'est ainsi que l'équipe du *NGML* s'est lancée dans la réalisation de paramètres adaptés à cette langue pour le logiciel libre *TreeTagger*³⁵. La mise au point parallèle d'outils d'analyse statistique des cooccurrences permet d'envisager le dégageant prochain d'un horizon encore inexploré en lexicographie médiolatine : le traitement « scientifique » des mots les plus fréquents³⁶ !

Outre ce profond renouvellement à venir des méthodes de rédaction, la collaboration, dans le cadre du projet COST – Medioevo Europeo, entre le *NGML* et le *LMILP*, rejoints par le *Glossarium Mediae Latinitatis Cataloniae* (Climent *et al.* 1960), a également abouti à la mise en œuvre d'un prototype d'encyclopédie interactive. Le projet de Wiki-Lexicographica³⁷, qui fait l'objet d'une communication conjointe avec Krzysztof Nowak pendant cette conférence, offre l'occasion de repenser les conditions de consultation des dictionnaires historiques. L'exploration des données lexicographiques, en s'affranchissant du cadre formel des articles traditionnels, peut suivre désormais des repères chronologiques, géographiques ou typologiques. Nouvelle composition, nouvelle exposition : le « Dictionnaire du latin médiéval » se transforme, dans le fond comme dans la forme, en quelque chose de vraiment « Nouveau ».

34 *Cf. supra* note 24.

35 Développé pour le marquage morphosyntaxique, ce logiciel permet aussi de lemmatiser : <<http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger>>.

36 Il deviendrait ainsi possible de reconstituer les rapports qu'entretiennent entre eux des mots considérés comme synonymes, tels *Pluvia* et *Imber* (pluie) ou *Polus* et *Celum* (ciel).

37 En cours de développement à l'adresse : <<http://www.scriptor.es/wiki>> ; *cf.* Bon/Nowak (2013).

Indications bibliographiques

- Alexandre, Renaud / Bon, Bruno / Guerreau-Jalabert, Anita 2013. Variations graphiques, variations morphologiques et lemmatisation du latin médiéval. In Elfassi, Jacques *et al.* (éd.) *Amicorum Societas. Mélanges offerts à François Dolbeau*. Firenze : Sismel, 3-18.
- Arnaldi, Francesco / Smiraglia P. 1970. *Latinitatis Italicae Medii aevi Lexicon imperfectum*. Bruxelles ; ²2001. Firenze.
- Blaise, Albert 1954. *Dictionnaire latin-français des auteurs chrétiens*. Strasbourg.
- Blatt, Franz / Lefèvre, Yves / Monfrin, Jacques / Dolbeau, François / Guerreau-Jalabert, Anita (dir.) 1957-2011. *Novum Glossarium Mediae Latinitatis*. Copenhague-Bruxelles-Genève ; disponible en ligne gratuitement : <<http://www.glossaria.eu/ngml>>.
- Bon, Bruno 2000. Compte rendu de la Session 'Dictionnaires du Latin Médiéval' réunie à Bruxelles les 18 et 19 juin 2000. *Archivum Latinitatis Medii Aevi* 58, 309-327.
- Bon, Bruno 2005. *Index scriptorum novus mediae latinitatis, Supplementum (1973-2005)*. Genève.
- Bon, Bruno 2011. L'influence des langues romanes sur le latin médiéval : l'exemple de 'Planus'. In González, Maurilio Pérez *et al.* (éd.) *Influencias léxicas de otras lenguas en el latín medieval*. Valladolid-León, 21-32.
- Bon, Bruno / Guerreau-Jalabert, Anita 2002. 'Pietas' : réflexions sur l'analyse sémantique et le traitement lexicographique d'un vocable médiéval. *Médiévales* 42, 73-88.
- Bon, Bruno / Nowak, Krzysztof 2013. Wiki Lexicographica. Linking Medieval Latin Dictionaries with Semantic MediaWiki. In Kosem, Iztok *et al.* (éd.) *Electronic Lexicography in the 21st century: Thinking outside the paper. Proceedings of the eLex 2013 Conference*. Tallinn-Ljubljana, 407-420 ; disponible en ligne gratuitement : <<http://eki.ee/elex2013/proceedings>>.
- Boronkai, Iván / Szovák, Kornél 1983-1999. *Lexicon Latinitatis Medii aevi Hungariae*. Budapest.
- Climent, Mariano Bassols de / Parera, Joan Bastardas / Quetglas, Pere

- J. 1960-2006. *Glossarium Mediae Latinitatis Cataloniae*. Barcelona.
- Du Cange, Charles Du Fresne 1678-1887. *Glossarium Medie et Infime Latinitatis*. Paris-Niort ; disponible en ligne gratuitement : <<http://ducange.enc.sorbonne.fr>>.
- Due, Otto Steen / Terkelsen, Peter (éd.) 1987-2008. *Lexicon Mediae Latinitatis Danicae*. Aarhus.
- Forcellini, Aegidius / De Vit, Vincentius 1858-1875. *Totius Latinitatis Lexicon*. Prato.
- Fuchs, Johanne W. / Weijers, Olga / Gumpert-Hepp, Marijke (éd.) 1977-2006. *Lexicon Latinitatis Nederlandicae Medii aevi*. Amsterdam-Leiden.
- González, Maurilio Pérez (dir.) 2010. *Lexicon Latinitatis medii aevi regni Legionis imperfectum*. Turnhout.
- Guerreau-Jalabert, Anita / Bon, Bruno 2010. Le trésor au Moyen âge : étude lexicale. In Burkart, Lucas *et al.* (éd.) *Le trésor au Moyen âge*. Firenze : Sismel, 11-31.
- Hakamies, R. 1958. *Glossarium Latinitatis Medii aevi Finlandicae*. Helsinki.
- Harvey, Anthony 1979-2005. *The Non-Classical Lexicon of Celtic Latinity*. Dublin-Turnhout.
- Heid, Caroline 2004. Table ronde « Lexicographie et informatique » (Barcelone, 2 juin 2004). *Archivum Latinitatis Medii Aevi* 62, 327-332.
- Instructions techniques destinées aux collaborateurs immédiats du Dictionnaire du latin médiéval. *Archivum Latinitatis Medii Aevi* 1, 66-76.
- Kostrenčić, Marko 1969-1978. *Lexicon Latinitatis Medii aevi Jugoslaviae*. Zagreb.
- Langlois, Charles-Victor 1924. Historique sommaire de l'entreprise de 1920 à 1924. *Archivum Latinitatis Medii Aevi* 1, 5-15.
- Latham, R. E. / Howlett, David R. / Ashdowne, Richard K. (éd.) 1975-2013. *Dictionary of Medieval Latin from British Sources*. Oxford.
- Lefèvre, Yves 1975. Les dictionnaires du latin médiéval et l'Union académique internationale. *Comptes Rendus des séances de*

l'Académie des Inscriptions et Belles-Lettres 119^e année, 402-414.

Lehmann, Paul / Gneuss, Helmut / Stotz, Peter (dir.) 1959-2014, *Mittellateinisches Wörterbuch bis zum ausgehenden 13. Jahrhundert*, München.

Niermeyer, Jan Frederik / De Kieft, C. van 1976-2004. *Mediae Latinitatis Lexicon Minus*, Leiden : Brill.

Plezia, Marian / Weyssenhoff-Brożkova, Krystyna / Rzepiela Michał (éd.) 1958-2014. *Lexicon Mediae et Infimae Latinitatis Polonorum*, Warszawa-Kraków.

Silagiová, Zuzana (dir.) 1977-2011. *Latinitatis Medii aevi Lexicon Bohemorum*. Praha.

Stotz, Peter 2002. *Handbuch zur lateinischen Sprache des Mittelalters I*. München : Beck.

Thesaurus Linguae Latinae 1900-2012. Leipzig-Stuttgart-Berlin.

Westerbergh, Ulla / Odelman, Eva 1968-2002. *Glossarium Mediae Latinitatis Suediae*. Stockholm.