

COMMUNICATION

LES « DOCUMENTS LINGUISTIQUES » DE LA FRANCE
HISTOIRE, PRÉSENT ET PERSPECTIVES D'UN PROJET CENTENAIRE,
PAR M. MARTIN GLESSGEN

Le projet des *Documents linguistiques* que Paul Meyer a entrepris en 1890 et que nous poursuivons depuis 2001 est emblématique d'une des grandes traditions philologiques de la France. Il est aussi intimement lié à l'Académie des Inscriptions et Belles-Lettres où ses protagonistes ont siégé même pendant plus d'un siècle : Paul Meyer était membre de 1883 à 1917, Clovis Brunel de 1937 à 1971 et Jacques Monfrin de 1983 à 1998.

Nous présenterons par la suite d'abord l'histoire du projet, puis son état éditorial actuel et enfin les volets de l'analyse linguistique que nous avons développés.

1. Histoire du projet

1. 1. LE CADRE GÉNÉRAL

L'écrit documentaire est le grand absent parmi les préoccupations de la philologie actuelle. Peu de linguistes et de philologues se sont passionnés pour cet ensemble textuel dont le rôle en son temps a toutefois été capital. Entre 1300 et 1500, les inventaires et documents comptables, les actes de la pratique juridique, décrets et testaments, ou encore les lettres commerciales et politiques forment plusieurs centaines de millions de pages en français, italien, occitan, catalan, espagnol et gallégo-portugais¹. La masse textuelle de l'écrit documentaire dépasse dans la *Romania* médiévale celle de tous

1. M. Glessgen, « Das altfranzösische Geschäftsschrifttum in Oberlothringen: Quellenlage und Deutungsansätze », dans *Skripta, Schreiblandschaften und Standardisierungstendenzen*, Actes du colloque de Trier, 16-18 septembre 1998, K. Gärtner *et alii* éd., Trier, 2001, p. 257-294, ici p. 268 et suiv.

les autres genres textuels confondus par un facteur 100. Certaines variétés romanes ne nous sont connues que grâce aux témoignages documentaires, comme le sarde médiéval ou, dans la Galloromania, le gascon, le francoprovençal et l'arverno-limousin.

Malgré leur richesse indéniable, autant dans les renseignements grapho-phonétiques que morphologiques, lexicaux et syntaxiques, les textes documentaires médiévaux ont été infiniment moins étudiés par les philologues romanistes que les textes littéraires, religieux, historiographiques ou scientifiques. Leur potentiel de recherche pour la linguistique historique est à peine exploité et leur rôle dans les pratiques de l'écrit et dans les processus d'élaboration des langues romanes reste mal cerné.

Étant donné par ailleurs la nature sérielle de l'écrit documentaire, il est indispensable de l'appréhender dans une logique globalisante avec une emprise sur l'intégralité d'un territoire et, idéalement, sur toutes les époques concernées. C'est ici que réside toute la valeur du projet entrepris par Paul Meyer, qui englobe à présent toute la Galloromania entre le XII^e et le XV^e siècle.

1. 2. PAUL MEYER (1840-1917)

Le projet des *Documents linguistiques* (DocLing) a connu quatre phases qui sont intimement liées aux différents protagonistes, à leurs choix et aux moyens dont disposait la science à leur époque². Paul Meyer a ouvert la voie, en défrichant un terrain encore vierge. Il décrit lui-même, vers la fin de sa vie, son cheminement, avec toute la clarté et la précision qui le caractérisaient :

« J'étais encore sur les bancs de l'École des chartes, que je copiais toutes les chartes provençales auxquelles je pouvais avoir accès. C'est à l'aide du recueil, bien insuffisant, que je m'étais formé et des textes imprimés, encore peu nombreux il y a trente ou quarante ans, que je rédigeai [en 1874 un] essai sur la langue d'oc et ses dialectes (...)[,] trop incomplet pour mériter de voir le jour. J'avais dû me résigner à mettre en œuvre des éléments trop peu nombreux. (...) J'ai donc reconnu de très bonne heure la nécessité

2. Voir, pour un aperçu plus détaillé de l'histoire des DocLing et pour la documentation bibliographique, M. Glessgen, « L'écrit documentaire et le projet des *Plus anciens documents linguistiques de la France* », dans *Manuel de la philologie de l'édition*, D. Trotter éd., Berlin-Boston, 2015, p. 267-295.

d'une exploration méthodique des archives du midi de la France. Mais (...) [c]'est seulement depuis une dizaine d'années que j'ai pu consacrer à des recherches dans les archives de la France méridionale une partie de mes vacances.

Ces recherches sont souvent pénibles. (...) j'ai eu le plus ordinairement à poursuivre mes recherches dans les archives communales, qui, en bien des villes, ne sont ni classées, ni inventoriées³. »

Le travail de Paul Meyer reflète les possibilités de son époque et l'habitus chartiste d'antan : en parcourant les différents dépôts d'archives, il transcrit sur place les documents qui lui semblent les plus significatifs, en couvrant toute la période de l'écrit occitan, allant parfois jusqu'au XVII^e siècle. Meyer s'est toutefois approché de la matière par étapes :

– en effet, dès 1874, il intègre une série de textes documentaires, notamment des chartes [p. 158-173] et coutumes [p. 173-192], dans la partie « provençale » [p. 23-192] de son *Recueil d'ancien textes bas-latins, provençaux et français* (1^{er} vol.) – alors que la partie française [p. 193-384] réunit comme la « provençale » des textes littéraires religieux et profanes, mais aucun texte documentaire ;

– une quinzaine d'années plus tard, il édite une série de documents dans le cadre de la brève étude « Le langage de Die au XIII^e siècle », *Romania* 20, 1891, p. 70-85 ; c'est la première approche de type systématique dans ce contexte ;

– suit une publication partielle, mais déjà importante, des « Documents linguistiques des Basses-Alpes », *Romania* 27, 1898, p. 337-441 ;

– puis l'édition *Le livre-journal de maître Ugo Teralh : notaire et drapier à Forcalquier (1330-1332)*, Paris, Klincksieck, 1898 [42 p.] et des « Fragments du grand livre d'un drapier de Lyon (1320-1323) », *Romania* 35, 1906, p. 428-444 ;

– enfin, en 1909, il publie l'imposant volume des *Documents linguistiques du Midi de la France, recueillis et publiés avec glossaires et cartes. Ain, Basses-Alpes, Hautes-Alpes, Alpes Maritimes* (le corpus de l'Ain [p. 1-166] ayant été préparé par son élève Édouard Philippon [1851-1926]) [655 p.].

3. P. Meyer, *Documents linguistiques du Midi de la France, recueillis et publiés avec glossaires et cartes. Ain, Basses-Alpes, Hautes-Alpes, Alpes Maritimes*, Paris, Honoré Champion, 1909, p. I-III.

Le caractère systématique de l'entreprise ressort pleinement de ce dernier volume qui couronne les efforts consentis par Paul Meyer pendant plusieurs dizaines d'années. Les grandes qualités de l'ouvrage résident dans la variété des genres documentaires réunis (inventaires divers, registres de recettes et dépenses, délibérations municipales, chartes, coutumes) ainsi que dans la couverture chronologique et géographique qu'il fournit pour la région pré-alpine.

Il s'agit d'un choix représentatif, notamment pour le département des *Basses-Alpes* (les « Alpes-de-Haute-Provence »), qui représente de loin la part la plus volumineuse avec 250 pages [p. 169-418]. Pour les *Hautes-Alpes*, « les textes de langue sont relativement rares » [p. 420], ce qui explique la taille réduite de cette partie comportant à peine 60 pages [p. 419-480]. Dans le département des *Alpes-Maritimes* [p. 481-654], Paul Meyer s'est heurté à une situation archivistique très complexe, ce qui a provoqué certains déséquilibres, notamment concernant la ville de Nice [p. 619-634], pour laquelle tout témoignage vernaculaire fait défaut entre 1115 et 1445.

De cette réalisation impressionnante, il ressort toutefois que Paul Meyer a largement sous-estimé l'importance de son entreprise. Il traite personnellement trois départements avec une tradition de l'écrit plutôt restreinte, mais sans en donner d'analyses linguistiques, à l'exception des notes linguistiques et toponymiques, très soignées mais éclatées, et des glossaires, comportant tous les lexèmes significatifs, mais également éclatés en quatre endroits. L'utilisation de ce volume composite n'est pas aisée et cela a sans doute contribué au fait qu'il n'ait jamais donné lieu à d'importantes études lexicologiques, grapho-phonétiques ou grammaticales. Le FEW (*Französische Etymologische Wörterbuch*) cite très systématiquement « MeyerDoc », mais la plupart du temps sans indiquer la date et le lieu des documents et sans entrer dans la réflexion définitoire des lexèmes souvent épineux. Les études grapho-phonétiques sont rares, les travaux sur la morphologie ou la syntaxe inexistant.

Mais Paul Meyer a eu le grand mérite d'ouvrir la voie, conformément à son propos (« L'important est de bien commencer. », p. I). Ses éditions comportent des renseignements précieux sur la *scripta* de la Provence et l'impact sur la recherche que les *Documents linguis-*

tiques du Midi de la France ont eu par leur modèle méthodologique est entier.

1. 3. CLOVIS BRUNEL (1884-1971)

La réalisation de Clovis Brunel a connu comme celle de Paul Meyer une phase préparatoire, représentée par la publication des documents du Gévaudan et par un large tour d'horizon des plus anciens documents occitans⁴. Ce dernier article profite pleinement des acquis de son maître et prédécesseur à la chaire de philologie romane à l'École des chartes. Aussi, il amène Brunel à la conclusion que l'approche de Meyer n'est pas réaliste pour l'ensemble du territoire d'oc :

« Les facilités offertes par le groupement des premiers départements dont le nom commence par *A* (Hautes-Alpes, Basses-Alpes, Alpes-maritimes), aussi bien que par la pauvreté des archives de ces pays de montagne, sont loin de se retrouver dans la suite de l'entreprise⁵. »

Il se concentre par conséquent sur la première tranche chronologique de l'écrit vernaculaire, en réunissant seulement les *plus anciens* « textes d'ordre judiciaire et administratif ». Ce choix, partiel mais systématique, se justifie notamment en vue de la force d'inertie de tout genre textuel, tendant à reproduire pendant des siècles des modèles établis. Le *terminus ad quem* retenu de *ca* 1200 permettait à Clovis Brunel d'achever de son vivant son projet en deux volumes réunissant 541 actes⁶. Lors de la parution du *Supplément*, il avait toutefois atteint un âge avancé, comme auparavant Paul Meyer.

Par d'autres paramètres encore, Clovis Brunel a augmenté le côté systématique des *Documents linguistiques* :

– la présentation des éditions suivant un ordre chronologique, sans considération du lieu d'origine, ce qui facilite l'utilisation ;

4. Cl. Brunel, « Documents linguistiques du Gévaudan », *Bibliothèque de l'École des chartes* 77, 1916, p. 5-57 et 241-285 ; Id., « Les premiers exemples de l'emploi du provençal dans les chartes », *Romania* 48, 1922, p. 335-364.

5. Cl. Brunel, *Les plus anciennes chartes en langue provençale. Recueil des pièces originales antérieures au XIII^e siècle. Publiées avec une étude morphologique*, Paris, 1926, p. VII ; Supplément, Paris, 1952.

6. Voir n. 5.

- la datation et localisation des documents à partir d'indices paléographiques et linguistiques ;
- la présentation homogène des éditions, avec chaque fois un bref regeste et une partie bibliographique ;
- des critères d'éditions plus diplomatiques (indication des résolutions d'abréviation en italique, indication des lignes de l'original, indication des parties en latin, toujours au moyen de l'italique ; dans le deuxième volume, indication des séparations de mots) ;
- dans chacun des deux volumes, un glossaire synthétique (1926, p. 448-495 ; 1952, p. 228-258) et une table de noms propres (1926, p. 345-448 ; 1952, p. 179-227), plus développés que ceux de Paul Meyer ;
- un relevé systématique de la morphologie nominale, pronomiale et verbale (1926, p. XIII-LV ; 1952, p. XIII-XXXVI).

Les deux volumes couvrent le XII^e siècle réunissant pour cette époque environ deux tiers des actes originaux conservés aujourd'hui. La distribution géographique des documents occitans est toutefois déséquilibrée et correspond à « une région qui peut être circonscrite entre les villes de Toulouse, Moissac, Villefranche-de-Rouergue, Rodez, Millau et Castres » (1926, p. IX). Seulement 41 des 541 actes ne sont pas languedociens ou rouergats (7,5 %) ⁷. Comme c'est le cas pour toutes les langues romanes médiévales, l'écrit (vernaculaire) ne couvre pas de manière homogène le territoire, mais est tributaire du développement des « lieux d'écriture » (cf. *infra* 3. 1.), à leur tour liés aux évolutions économiques, socio-politiques et culturelles.

Les genres textuels sont en revanche plus restreints que chez Paul Meyer, lequel avait pris en considération toute la diversification de l'écrit documentaire des XIII^e-XV^e siècles. Ainsi, chez Brunel les chartes sont largement dominantes ⁸. Il est certain qu'il s'agit là d'un des genres textuels les plus complexes et les plus riches, au moins dans l'*expositio*, mais d'autres genres documentaires fournissent d'autres informations, comme la longue liste de redevances de la

7. Vingt documents proviennent du domaine occitan septentrional (Limousin, Périgord, Auvergne et Vivarais), dix documents de la Provence et du Dauphiné et onze documents de la Gascogne (cf. Cl. Brunel, *op. cit.* [n. 5], 1926, p. LVI-LXIII ; 1952, p. XXXVI-XXXIX).

8. Cette différence ressort en effet des titres des deux éditions qui opposent les *Documents (linguistiques)* aux (*Plus anciennes chartes (provençales)*), sachant que les « chartes » en question sont essentiellement languedociennes – « provençal » étant utilisé comme hypéronyme pour la langue d'oc – et que ce sont les « documents » qui proviennent vraiment de la Provence.

commanderie de Manosque de la fin du XII^e siècle⁹ que Brunel a exclu n'ayant pas vu l'apport capital des noms de lieux et de personnes pour la linguistique historique.

Enfin, Clovis Brunel a exclu, de manière très rigoureuse, des copies, même « transmises par des scribes du pays des actes originaux et à peu près du même temps¹⁰ ». Si son souci de contrôler la quantité des textes est parfaitement compréhensible, la différence entre original et copie est graduelle, surtout dans le domaine de l'écrit documentaire. Un original est certes toujours préférable à une copie, mais renoncer à l'étude des copies rendrait impossible, par exemple, le traitement des genres littéraires ou religieux du Moyen Âge comme de l'Antiquité. Dans le cas de l'écrit occitan antérieur à 1200, la prise en considération des copies contemporaines multiplierait par deux ou trois le volume textuel et fournirait une vision géo-chronologique plus équilibrée de l'émergence de la scripturalité en langue d'oc puisque les textes les plus précoces d'autres régions – comme l'Auvergne – ne sont souvent transmis que sous forme copiée.

Les travaux de Clovis Brunel ont connu une plus grande fortune que ceux de Paul Meyer notamment grâce aux études grapho-phonétiques et morphologiques d'Åke Grafström sur les chartes languedociennes (1958 et 1968) et celles de Kalman pour les chartes rouergates (1974), mais aussi grâce à leur utilisation intense par le FEW et, plus généralement, par la lexicographie de l'ancien occitan (DAO, DOM)¹¹. Étant donné la grande richesse lexicale, onomastique, morphologique et syntaxique des actes, *Les plus anciennes chartes* comportent toujours un potentiel important pour la recherche.

En conclusion, les différences d'approche entre Paul Meyer et Clovis Brunel font bien ressortir les paramètres à prendre en considération dans le travail philologique appliqué aux textes documentaires.

9. Cf. *op. cit.* (n. 4), 1922, p. 338.

10. « Ont été seules retenues les pièces conservées en original, tant par souci de ne produire que des textes à l'abri de tout soupçon d'altération due à un copiste que par désir de ne pas donner à la publication une étendue trop grande », Cl. Brunel, *op. cit.* (n. 5), 1926, p. VII et suiv.

11. Notons toutefois que les études de Grafström ne reposent que sur 141 des 541 chartes de Cl. Brunel et que le FEW reste presque aussi sybillin dans la citation des volumes que pour celui de P. Meyer.

1. 4. JACQUES MONFRIN (1924-1998)

Le projet des *Documents linguistiques* change radicalement d'orientation avec Jacques Monfrin. Bien qu'occitaniste, il décide de ne pas poursuivre le travail sur les documents du Midi, mais d'ouvrir le chantier des documents d'oïl. Il implique de nombreux élèves et collègues dans le projet, ce qui permet un recensement plus systématique des régions concernées. En prenant appui sur le modèle établi par Louis Carolus-Barré pour les documents de l'Oise (DocOis)¹², J. Monfrin établit également des critères d'édition plus rigoureux et cohérents et augmente encore le souci du détail des volumes publiés (édition, notes, glossaire, index onomastiques).

Le résultat le plus visible de « l'époque Monfrin » sont les trois volumes de textes documentaires dirigés avec Lucie Fossier, comportant en tout 514 chartes françaises de cinq départements contigus à l'est de Paris (Seine-et-Marne, Yonne, Aube, Haute-Marne et Vosges : 1974, 1975 et 1988)¹³. S'ajoutent deux volumes pour la Wallonie avec les plus anciens actes du Hainaut et de la Flandre (1984, 1987)¹⁴.

Mais les volumes publiés ne donnent qu'une idée partielle de l'avancement du projet des *Documents linguistiques*. En effet, surtout grâce aux thèses de l'École des chartes, J. Monfrin a suivi entre 1964 et 1994 la réalisation de sept autres séries qui sont restées à l'état manuscrit : Douai, le Pas-de-Calais, la Somme et l'Aisne pour la Picardie, la Meurthe-et-Moselle, la Côte d'Or et le Doubs pour le sud-est du territoire d'oïl¹⁵.

12. DocOis = *Les plus anciennes chartes en langue française*, t. 1. *Problèmes généraux et recueil des pièces originales conservées aux Archives de l'Oise (1241-1286)*, L. Carolus-Barré éd., Paris, 1964.

13. DocHM = *Chartes en langue française antérieures à 1271 conservées dans le département de la Haute-Marne*, J.-G. Gigot éd., Paris, 1974.

DocV = *Chartes en langue française (...) des Vosges*, J. Lanher éd., Paris, 1975.

DocAub = *Chartes en langue française (...) de l'Aube, de la Seine-et-Marne et de l'Yonne*, D. Coq éd., Paris, 1988.

14. DocHain = *Chartes en langue française antérieures à 1271 conservées dans la province de Hainaut*, P. Ruelle éd., Paris, 1984.

DocFl = *Chartes en langue française (...) en Flandre orientale et Flandre occidentale*, R. Mantou éd., Paris, 1987.

15. En dehors des DocCO, toutes ces publications ont été annoncées par J. Monfrin dans son « Introduction » aux DocHM, *op. cit.* (n. 12), p. XI-LXXX, ici p. XLIX :

DocAisne_ms = *Chartes de l'Aisne (1214-1350)*, F. Grégoire-Ollivier éd., thèse de l'École des chartes, 1964.

D'un point de vue chronologique, ces volumes répertorient les débuts de l'écrit documentaire vernaculaire jusqu'en 1270 environ, parfois un peu au-delà. C'est donc un choix identique à celui de Clovis Brunel, à la différence près que les actes oïliques font apparition un siècle après les actes occitans.

Quant à la couverture géographique, les volumes publiés et manuscrits couvrent une grande partie du territoire d'oïl septentrional (Picardie et la Wallonie), oriental (Champagne et Lorraine) et sud-oriental (Bourgogne et Franche-Comté). En revanche, Monfrin a fait l'impasse sur la Normandie, l'Ouest et le Sud-Ouest, de même que sur Paris et ses environs immédiats. Curieusement, le sondage commencé par Monfrin dans les Archives Nationales et à la Bibliothèque nationale n'a pas été poursuivi¹⁶, alors que ces deux lieux de conservation sont les plus facilement consultables par les philologues parisiens.

Enfin, J. Monfrin considère avec plus de prudence que Cl. Brunel les questions de transmission, étant donné que « (...) la différence entre les originaux et les copies est moins de nature que de degré » (*op. cit.* [n. 15], p. 46). Les *Documents linguistiques* incluent ainsi régulièrement des copies contemporaines, même s'ils restent ciblés sur les originaux.

Un apport de poids enrichit cette troisième phase des *Documents linguistiques* : il s'agit de deux volumes complémentaires dédiés au territoire francoprovençal qui ont été projetés par Mgr Pierre Gardette (1974, 1975)¹⁷. Ces éditions couvrent essentiellement le

DocCOR_ms = *Chartes de la Côte d'Or (1239-1270)*, V. Neveu éd., thèse de l'École des chartes, 1988.

DocDo_ms = *Chartes de Douai (1204-1270)*, M. Mestayer éd., ca 1974.

DocDoub_ms = *Chartes du Doubs (1233-1261)*, M. Lefèvre éd., thèse de l'École des chartes, 1975 ; *Chartes du Doubs (1260-1271)*, J. Ducourtieux éd., thèse de l'École des chartes, 1994.

DocMM_ms = *Chartes de la Meurthe-et-Moselle (1232-1265)*, M. Arnod éd., thèse de 3^e cycle, 1974.

DocPC_ms = *Chartes du Pas-de-Calais*, P. Bougard éd., ca 1980.

DocSom_ms = *Chartes de la Somme*, J. Estienne éd., ca 1980.

16. J. Monfrin a relevé dans le Trésor des chartes aux A. N. « à peu près 275 pièces en français antérieures à 1270 », « plus de cent cinquante à la Bibliothèque nationale », sans compter « les collection des Provinces, Picardie, Lorraine, Champagne » : cf. Id. « Le mode de tradition des actes écrits et les études de dialectologie », *Revue de Linguistique Romane* 32, 1968, p. 17-47, ici p. 24.

17. DocLyo = *Documents linguistiques du Lyonnais (1225-1425)*, P. Durdilly éd., Paris, 1975.

DocFor = *Documents linguistiques du Forez (1260-1498)*, M. Gonon éd., Paris, 1974.

XIV^e ainsi qu'une partie du XV^e siècle et donnent un recensement intégral avant d'éditer un choix de documents. Ils suivent donc la logique de Paul Meyer et continuent le travail d'Édouard Philippon sur l'Ain. Quant aux genres textuels, si les séries françaises sont constituées dans l'immense majorité par des chartes, la série francoprovençale en est presque exempte : ici dominent les documents de gestion et diverses listes.

D'un point de vue linguistique, les 62 documents du Forez correspondent pour la plupart, comme ceux de l'Ain, à une *scripta* mixte de type oïlique, intégrant des éléments lexicaux et toponymiques francoprovençaux¹⁸. L'important volume sur le Lyonnais, avec près de 600 pages, présente en revanche une *scripta* essentiellement francoprovençale, ce qui souligne le rôle de Lyon comme épiceutre de cette langue.

Les avancées éditoriales des *Documents linguistiques* sont donc considérables : grâce aux volumes publiés sous la direction de Jacques Monfrin, Lucie Fossier et Pierre Gardette ainsi qu'au volume de Louis Carolus-Barré, on disposait alors de près de 800 chartes oïliques du XIII^e siècle et d'une remarquable série de documents francoprovençaux du XIV^e siècle, tous dans un état éditorial exemplaire. S'ajoutait l'édition manuscrite de plus de 1700 chartes françaises, dans un état encore préliminaire mais très avancé.

Les éléments analytiques en revanche n'ont pas été développés. Les volumes de la série française disposent chacun d'un glossaire soigné selon la conception de Clovis Brunel, mais leur éclatement en six volumes et leur caractère trop peu interprétatif n'en facilitent pas l'utilisation¹⁹. Les volumes francoprovençaux, quant à eux, ont renoncé à des glossaires individuels en attendant un glossaire général dont la rédaction a avorté, suite à la mort prématurée de Mgr Gardette (1906-1973).

18. Les cinq documents véritablement francoprovençaux (n° 4-7, 14) sont toutefois pour certains assez volumineux (le *Terrier de Chazelles*, 1290, et les *Comptes de l'hôtel du comte de Forez*, 1322/1323).

19. Cf. le plaidoyer incontournable de J.-P. Chambon, « Lexicographie et philologie : réflexions sur les glossaires d'éditions de textes (français médiéval et préclassique, ancien occitan) », *Revue de Linguistique Romane* 70, 2006, p. 123-141 (reproduit aussi dans Id., *Méthodes de recherches en linguistique et en philologie romanes*, É. Buchi et alii éd., Strasbourg, 2017, vol. 2, p. 657-675).

1. 5. SYNTHÈSE : L'ÉTAT DES *DOCUMENTS LINGUISTIQUES* VERS 2000

En conclusion, vers l'an 2000, les *Documents linguistiques* sont présents dans les trois territoires de la Galloromania. La qualité et le soin des éditions sont le grand atout de cette collection qui garde comme objectif un recensement systématique de l'écrit documentaire médiéval. Pourtant, les réalisations accessibles de l'époque de Jacques Monfrin sont restées, plus encore que le volume de Paul Meyer, une œuvre inachevée, géographiquement disparate et difficile à exploiter²⁰.

Dans un certain sens, le projet de Monfrin s'inscrivait dans l'esprit de l'époque gaullienne, tout comme le *Trésor de la langue française* (TLF) ou l'entreprise des *Nouveaux atlas linguistiques de la France* (NALF). En ce qui concerne le TLF, grâce à un financement très considérable, il a été possible de saisir les données textuelles à trois reprises²¹ et d'achever sur cette base le « Nouveau Littré » projeté ; son pendant médiéval, le *Dictionnaire du Moyen Français* (DMF), quant à lui, a eu la chance d'avoir été repris, bien après sa première conception en 1980, par Robert Martin. Les *Nouveaux atlas*, en revanche, partagent plutôt le destin des *Documents linguistiques*, puisque de nombreux volumes restent encore aujourd'hui à l'état de manuscrits et que peu de séries ont connu ne serait-ce que des index lexicologiques conséquents²². Le caractère monumental de l'entreprise a en quelque sorte empêché son achèvement.

S'ajoutent deux problèmes intrinsèques : le projet de Meyer, Brunel et Monfrin était pensé, tout comme les *Nouveaux atlas*, dans une logique de répertoire plus que d'utilisation. Il est évident que des éditions de qualité et des cartes d'atlas bien pensées sont des sources

20. Même les volumes publiés de la série française n'ont que rarement servi à des analyses linguistiques, mis à part leur exploitation systématique dans le fichier du DEAF (*Dictionnaire Étymologique de l'Ancien Français*) et, par cette voie, leur présence dans le DEAFpré. L'utilisation peut-être la plus adéquate est l'intégration des trois volumes des DocOis, DocHM et DocV (voir n. 10 et 11) dans l'*Atlas des formes et des constructions des chartes françaises du XIII^e siècle* d'Anthonij Dees, Tübingen, 1980 (cf. la liste de la « provenance des chartes », p. 307-312).

21. Notre élève Ruth Radermacher a retracé dans sa thèse (*Le Trésor de la langue française. Analyse métalxicographique*, Strasbourg-Saarbrücken, 2004) cette triple saisie, d'abord sur des cartes perforées, puis sur des bandes de papier, puis enfin sur un support électronique.

22. Voir par exemple, C. Dondaine, *Trésors étymologiques de la Franche-Comté (d'après l'Atlas linguistique et ethnographique de la Franche-Comté)*, Strasbourg, Éditions de la Société de linguistique romane, 2002. Actuellement, G. Brun-Trigaud poursuit le projet de rendre accessibles les volumes restés manuscrits des NALF.

indispensables pour la linguistique historique et variationnelle. Mais la mise à disposition des matériaux devrait toujours s'accompagner d'analyses interprétatives pour ne pas rester lettre morte et en susciter d'autres. Le deuxième problème de tels projets à dimension nationale réside dans l'immensité de la matière à traiter que Jacques Monfrin a sous-estimée, tout comme ses prédécesseurs et – il faut bien l'avouer – tout comme nous-même, dans la quatrième phase des *Documents linguistiques*.

2. Les corpus électroniques actuels

2. 1. LES TEXTES RÉUNIS ACTUELLEMENT DANS L'ÉDITION ÉLECTRONIQUE DES *DOCUMENTS LINGUISTIQUES*

Après la disparition de Jacques Monfrin en 1998, Françoise Vielliard hérite de ses matériaux des *Documents linguistiques* et met généreusement à ma disposition ces documents, son expérience ainsi que le soutien institutionnel de l'École des chartes. Nous nous sommes donné alors l'objectif, avec Olivier Guyotjeannin, de poursuivre le dessein de Jacques Monfrin sous les auspices de la philologie informatique : informatiser les volumes existants, saisir et corriger les volumes manuscrits, élargir géographiquement les recensements²³. Si cette idée nous semble toujours rationnelle, la saisie et la révision philologique des volumes non publiés ainsi que le développement informatique du projet se sont révélés extrêmement lourds, au-delà de toute attente.

Les efforts conjoints et le soutien de nombreux collègues et amis, élèves et jeunes chercheurs, de même que l'apport financier considérable de plusieurs institutions académiques²⁴ ont toutefois permis de réunir 2416 actes en douze séries sous un format élec-

23. Ce travail engagé a conduit à une première version des DocLing électroniques : *Les plus anciens documents linguistiques de la France. Édition électronique*, collection fondée par J. Monfrin, poursuivie par M. Glessgen, première édition en collaboration avec F. Vielliard et O. Guyotjeannin, 2009. Après la reprogrammation de la base de données et l'intégration des documents du territoire francoprovençal, nous disposons désormais de la troisième version : *Les plus anciens documents linguistiques galloromans (DocLing). Édition électronique*, dirigée par M. Glessgen, en partenariat avec H. Carles, F. Duval et P. Videsott, troisième édition revue et fortement élargie, 2016, <www.rose.uzh.ch/docling>.

24. Cf. les indications détaillées concernant les collaborateurs et les financements sur le site des DocLing-3, *op. cit.* (n. 23).

tronique homogène, accompagnés pour l'essentiel de reproductions photographiques et permettant des interrogations linguistiques poussées :

sigle	département (ou lieu d'écriture)	dates extrêmes	auteur d'une édition papier	auteur ou responsable de l'édition électronique	nombre de doc.
DocCO	Côte-d'Or	1239-1306	Neveu 1988	Glessgen 2016	231
DocDo	Douai	1204-1270	Mestayer <i>ca</i> 1970	Brunner 2014	506
DocJu	Jura	1243-1296		Muller 2014	95
DocHM	Haute-Marne	1232-1275	Gigot 1974	Kihaï 2009	276
DocHS	Haute-Saône	1242-1300		Muller 2014	132
DocMa	Marne	1234-1272		Kihaï 2009	230
DocMe	Meuse	1225-1270		Matthey 2009	237
DocMM	Meurthe-et- Moselle	1232-1265	Arnod 1974	Glessgen 2014	290
DocNi	Nièvre	1289-1330		Alletsgruber 2014	34
DocR	Chancellerie royale	1241-1300		Videsott 2014	140
DocSL	Saône-et- Loire	1227-1331		Alletsgruber 2014	100
DocV	Vosges	1235-1275	Lanher 1975	Trotter 2014	146 ?

Cette collection comprend deux des volumes anciennement publiés (DocHM et DocV), trois des volumes restés manuscrits (DocMM – intégralement revus –, DocDo et DocCO) ainsi que sept nouvelles séries élaborées par quatre de mes élèves ainsi que par Paul Videsott, de l'université de Bolzano.

Pour documenter les débuts de l'écrit dans les différentes régions, nous avons reculé la date butoir jusqu'en 1300 voire, pour la Bourgogne, jusqu'en 1331. Les documents réunis couvrent ainsi une zone relativement continue du Nord et, surtout, du Nord-Est de la France (Flandre, Champagne, Lorraine, Franche-Comté, Bourgogne) en incluant la chancellerie royale.

En collaboration avec Hélène Carles, de l'université de Neuchâtel, et son assistante Lorraine Fuhrer, nous avons par ailleurs numérisés en 2016 les textes documentaires du territoire francoprovençal pour lequel il existe une édition de qualité. Grâce à ces quelque 760 documents, souvent longs, la base textuelle des *Documents linguistiques* a presque doublée :

sigle	région	dates extrêmes	auteur de l'édition papier	nombre de doc.
DocAin	Ain	1265-1466	Philipon 1909 [dans : Meyer 1909]	18
DocLyon	Lyonnais	[1225]- 1425	Durdilly 1975	49
DocForez	Forez	1260-1498	Gonon 1974	63
DocGrenoble	Grenoble	1338/1339	Devaux 1892	[comptes consulaires]
Codi	Dauphiné	ca 1200 ?	Royer 1929-1933 Royer/Thomas 1983#?#	[Somme du code]
DocFrib	Fribourg	1296-1496	Fuhrer en prép.	74
DocBeJu	Jura et Berne	1244-1395	Schüle/Scheurer/ Marzys 2002	299
DocNCh	Neuchâtel	1238-1379	Scheurer/Morerod/ Kristol à paraître	182
DocGeVa	Genève et Vaud	1247-1406	Gavillet 2011	72

Nous avons également entrepris la numérisation des volumes de Paul Meyer et de Clovis Brunel, ce qui ouvrira enfin de nouvelles perspectives d'exploitation pour ces éditions précieuses.

Dans les prochaines années, nous prévoyons par ailleurs diverses intégrations complémentaires :

- les volumes publiés par Louis Carolus-Barré et par Dominique Coq ;
- les deux volumes publiés des séries wallonnes ;
- le volume de la Moselle, préparé par notre regrettée amie d'études Martina Pitz ;

– une série d’actes de la prévôté de Paris (1260-1300, *ca* 120 documents), en cours d’étude par Paul Videsott ;

– enfin, les quatre thèses manuscrites restées en suspens (DocAisne, DocDoub, DocPC, DocSom) ainsi que, pour la série occitane, les documents de la Provence occidentale édités sous un format word par Pierre et Maud Payan, Marie-Rose Bonnet et Pierre Paul.

Il restera alors la grande lacune de l’Ouest oïlique, de la Normandie au Poitou²⁵, sans oublier l’Angleterre avec sa scripturalité précoce. Le XIV^e et le XV^e siècle ne seront présents que pour les domaines francoprovençal et occitan et le gascon tout comme l’occitan septentrional seront absents du corpus. Mais chaque ensemble ajouté demande un effort considérable pour vérifier et homogénéiser les données textuelles et électroniques ce qui oblige à la prudence.

La documentation couvrira toutefois une grande partie de la Galloromania de manière relativement systématique et ceci toujours sur une base philologique fiable. Même si l’étendue de la matière n’aura pas été pleinement dominée, les efforts réunis de quatre générations de chercheurs auront créé une référence solide pour les travaux futurs. Grâce à la gestion informatique des données, il reste par ailleurs possible à tout moment de poursuivre l’enrichissement du corpus, même au-delà de l’écrit documentaire.

2. 2. LES PRINCIPES D’ÉDITION DES *DOCUMENTS LINGUISTIQUES* ÉLECTRONIQUES

Pour l’édition électronique, les différentes séries des *Documents linguistiques* ont été saisies dans un encodage neutre de type XML et intégrées dans une base de données MySQL. Le site qui les héberge actuellement est le serveur du *Romanisches Seminar* de Zurich (<www.rose.uzh.ch/docling>) (fig. 1).

Le site permet de visualiser l’édition de chacun des *ca* 3200 documents, pour la plupart accompagnés d’une photo. En passant par l’onglet « Les corpus textuels » et en choisissant l’un

25. L’édition de M. S. La Du des *Chartes et documents poitevins du XIII^e siècle en langue vulgaire* (DocPoit, 1960-1964) comporte *ca* 100 actes antérieurs à 1300, mais elle intègre de nombreuses copies postérieures et s’avère très peu fiable des points de vue paléographique et philologique.

Accueil

Descriptif du projet

Les corpus textuels

Contact

Les plus anciens documents linguistiques galloromans
Édition électronique
dirigée par Martin Glessgen

en partenariat avec Hélène Carles, Frédéric Duval et Paul Videsott

Troisième édition (2016)

La collection des *Plus anciens documents linguistiques galloromans* (DocLing) fournit une édition fiable d'un nombre important de textes documentaires médiévaux conservés pour l'essentiel sous une forme originale et en grande partie inédits. Le projet prend appui sur les *Plus anciens documents linguistiques de la France* consacrée aux chartes françaises du XIII^e siècle. Cette collection, qui poursuivait pour le domaine d'oïl les éditions occitanes de Paul Meyer (1909 etc.) et de Clovis Brunel (1926/52), fut initiée en 1974 par Jacques Monfrin et poursuivie depuis 1998, sous un format électronique, par Martin Glessgen, en collaboration avec Françoise Vieliard et Olivier Guyotjeannin.

Une première édition électronique, mise en ligne en 2009, comprenait les textes, les tableaux analytiques et les photographies numérisées de 1133 documents oïliques compris entre 1224 et 1275, conservés dans les quatre départements de la Meurthe-et-Moselle, de la Meuse, de la Marne et de la Haute-Marne. La deuxième édition, élargie et révisée, comportait 2182 documents de 1204 à 1331 couvrant une grande partie du Nord-Est de la France (Flandre, Champagne, Lorraine, Franche-Comté, Bourgogne), ainsi que les plus anciens actes de la chancellerie royale (133 documents). Cette version permettait d'interroger la base de données textuelles de différentes manières à des fins linguistiques, philologiques et historiques.

La présente édition ajoute deux autres collections du Sud-Est oïlique (le département de la Côte d'Or et les documents de Berne et du Jura bernois en Suisse romande). Par ailleurs, elle intègre les documents du territoire francoprovençal, autant de France (Lyon, Forez, Ain, Grenoble) que de Suisse (Fribourg, Neuchâtel, Genève, pays de Vaud). Ces documents se placent à nouveau dans l'optique de Paul Meyer et couvrent, de manière représentative, l'époque entre le 13^e et le 15^e siècle (ca 1250 – ca 1450). L'ouverture vers le domaine d'oc, actuellement en chantier, est prévue pour la quatrième édition du site.

Les éditions sont de type diplomatique, mais elles intègrent en même temps des éléments interprétatifs afin de permettre une meilleure compréhension des textes; grâce à leur encodage en xml, chaque acte peut être visualisé, selon les choix des lecteurs, sous une forme purement diplomatique, sous une forme interprétative ou sous la forme d'un encodage 'double' qui allie les deux principes d'édition. Cela vaut pour tous les documents nouvellement édités dans le cadre du projet, pour lesquels le site reproduit également les photographies des originaux. Pour les documents repris d'éditions antérieures, l'édition est de type semi-diplomatique, plus traditionnelle. Tous les actes sont accompagnés de tableaux analytiques détaillés qui les replacent dans le diachronisme historique de leur genèse et fournissent les informations diplomatiques nécessaires.

La base de données textuelles est gérée par le logiciel «phoenix», qui fait appel à différents outils informatiques (surtout MySQL et PHP). Ce logiciel, développé aux fins du présent projet, a comme vocation l'édition et l'analyse linguistique des textes anciens et répond aux exigences de l'open source.

Les formes lexicales contenues dans les actes oïliques sont lemmatisés et intégrés dans la base de données du *Dictionnaire étymologique de l'Ancien Français* (DEAFPré). On aura ainsi accès dès 2017 à l'intégralité du vocabulaire français des DocLing via la nomenclature du DEAF électronique.

L'édition électronique des *Plus anciens documents* est le fruit d'un travail d'équipe réunissant les efforts de nombreux chercheurs européens depuis 1998. Les financements indispensables ont été assurés en premier lieu par le Fonds National Suisse (FNS, dans le cadre du centre national de la recherche «Mediality»), mais aussi par les Universités de Zurich, de Strasbourg (MISHA), de Bolzano, d'Aberystwyth et de Neuchâtel, l'École Nationale des Chartes, le Fonds de la recherche autrichien, l'Association William Pierrehumbert et l'Académie des Sciences de Heidelberg.




FIG. 1. – Site des *Plus anciens documents*.

parmi la vingtaine d'ensembles actuellement disponibles, l'on arrive à travers la « Table des dates » aux éditions individuelles, comme ce document tiré du corpus de la Meurthe-et-Moselle (fig. 2).

Les principes d'édition des *Documents linguistiques* sous la forme que Jacques Monfrin avait retenue ont été respectés pour l'édition électronique. Nous avons toutefois introduit des précisions ultérieures, par une segmentation sémantico-syntaxique des actes et par l'indication des majuscules et de la ponctuation originales. La représentation que nous favorisons est une vue « mixte » qui rend les séparations médiévales transparentes et combine la ponctuation

Accueil

Descriptif du projet

Les corpus textuels

chartes de Meurthe-et-Moselle

Table des dates

Interrogations linguistiques

[vers une autre interface]

Contact

Choix d'édition

édition critique

édition interprétative

édition diplomatique

Montrer la structure rhétorique

non

oui

Téléchargement du document

fichier xml

fichier pdf

Les plus anciens documents linguistiques de la France
Corpus : Array (chMM)

Responsable du corpus : Martin-D. Glessgen
Édition de la charte : Martin-D. Glessgen / Michel Arnold (révision: D. Kihal)

chMM002

1234 (25 mars-31 décembre) ou 1235 (1^{er} janvier-24 mars)

TYPE DE DOCUMENT: charte: accensement de terres
OBJET: L'abbé et le chapitre de Salival accensent à Wirrion et Houillon treize journaux de terre au finage de Juvelize contre un cens de treize deniers et deux hémines de grain.

AUTEUR: non annoncé
DISPOSANT: abbaye de Salival
SCEAU: non annoncé
BÉNÉFICIAIRE: abbaye Notre-Dame de Salival
AUTRES ACTEURS: Wirrion et Houillon, paysans de Juvelize
RÉDACTEUR: scriptorium de l'abbaye Notre-Dame de Salival [pal]

SUPPORT: parchemin jadis scellé sur simple queue; 58x141
LIEU DE CONSERVATION: AD MM H 1244, fonds de l'abbaye de Salival

1 Conue chose soit à-toz 2 *que* li abes *et* li chapitles de Salinvas: at laissié à Wirion 12 *et* Huillon, les dous freres de Geverlise, les anfanz Bertran Bacheler, 3 ·XIII· jor·nas de terre treisse: en la fin de Geverlise: *et* à lor oirs: 4 parmi ·XIII· deniers de cens: *et* 10 ·II· himas de blef: l'un d'avoine: l'autre de froment: 5 *et* s'il ne paievent à jor 15 nomei à la feste sent Remi: à Giverlise, en la maison de Salinvas ^[i]. *que* l'on se taneroit à la terre: *et* ce que sus averoit:

6 Si est ensi devisee: q'au Tramble en 17 at ·III· jor·nas: un par lui ^[ii]. *et* ·III· ensemble: 7 *et* en la voie de Hignicort en at 18 V· jor·nas: II· d'une part *et* ·III· d'autre: 8 *et* en la voie de Marsal ·II· jor·nas: 19 après la terre les Voves ^[iii]. 9 *et* en la voie de Donnerais: as Genoivres: en at 10 ·II· jor·nas:

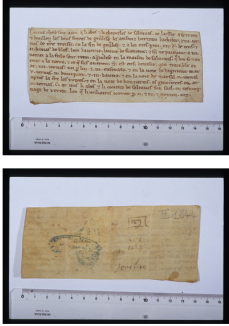


FIG. 2. – Édition d'un acte sur le site des *Documents linguistiques*.

1 Conue chose soit à-toz 2 *que* li abes *et* li chapitles de Salinvas: at laissié à Wirion 12 *et* Huillon, les dous freres de Geverlise, les anfanz Bertran Bacheler, 3 ·XIII· jor·nas de terre treisse: en la fin de Geverlise: *et* à lor oirs: 4 parmi ·XIII· deniers de cens: *et* 10 ·II· himas de blef: l'un d'avoine: l'autre de froment: 5 *et* s'il ne paievent à jor 15 nomei à la feste sent Remi: à Giverlise, en la maison de Salinvas ^[i]. *que* l'on se taneroit à la terre: *et* ce que sus averoit:

6 Si est ensi devisee: q'au Tramble en 17 at ·III· jor·nas: un par lui ^[ii]. *et* ·III· ensemble: 7 *et* en la voie de Hignicort en at 18 V· jor·nas: II· d'une part *et* ·III· d'autre: 8 *et* en la voie de Marsal ·II· jor·nas: 19 après la terre les Voves ^[iii]. 9 *et* en la voie de Donnerais: as Genoivres: en at 10 ·II· jor·nas:

10 Ci at mis li abes *et* li covenz de Salinvas son sael: en tesmoignage de verité:

11 l'an que li miliaires corroit par ·M· *et* CC· *et* XXXIII· anz:

FIG. 3. – Édition « mixte ».

médiévale [au milieu de la ligne] et moderne de même que les majuscules médiévales [en gras] avec les majuscules modernes (fig. 3).

Cette vue mixte peut être exportée pour tous les documents de manière automatique sous un format PDF. L'encodage XML permet de représenter les mêmes textes également sous deux autres vues, l'une diplomatique et l'autre pleinement interprétative :

– la vue diplomatique est alors organisée d’après les lignes du manuscrit et reproduit les séparations des mots, la ponctuation et les majuscules de l’original, les abréviations étant résolues en italiques : le texte de la charte DocMM 2 (1234, Abbaye de Salival) devient ainsi :

- 1 Conue chose soit atoz *que* li abes *et* li chapitles de salinvas · at
laissie a wirion
- 2 *et* huillon les dous freres de *geverlise* les anfanz bertran bachelier
·XIII· jor
- 3 nas de *terre* treisse · en la fin de *geverlise* · *et* a lor oirs · parmi
·XIII· *deniers* de cens · *et*
- 4 ·II· himas de blef · lun davoine · lautre de froment · *et* sil ne
paievent a jor
- 5 nomei a la feste sent remi· a *giverlise* en la maison de salinvas
· *que* lon se tan
- 6 roit a la terre · *et* ce *que* sus averoit· Si est ensi devisee· qau
Tramble en
- 7 at ·III· jornas · un *par* lui · *et* ·III· ensemble · *et* en la voie (...)

Cette vue diplomatique peut être immédiatement comparée avec la reproduction photographique (fig. 4).

– la vue interprétative introduit en revanche des séparations de mots, la ponctuation et les majuscules selon l’usage actuel, ajoute des accents, des numéros de structuration sémantico-formelle et, le cas échéant, des paragraphes ; cela donne à voir un texte en apparence plus moderne :

1 Conue chose soit à toz 2 *que* li abes *et* li chapitles de Salinvas at
laissé à Wirion et Huillon, les dous freres de Geverlise, les anfanz
Bertran Bachelier, 3 XIII jornas de terre treisse, en la fin de Geverlise,
et à lor oirs, 4 parmi XIII deniers de cens et II himas de blef, l’un
d’avoine, l’autre de froment ; 5 *et* s’il ne paievent à jor nomei à la
feste sent Remi, à Giverlise en la maison de Salinvas, que l’on se
tanroit à la terre *et* ce *que* sus averoit (...)

La base textuelle permet enfin des mises en page plus élaborées à travers une interface avec le logiciel d’édition InDesign. Nous avons réalisé par cette voie notamment l’édition des *Plus anciens documents en français de la chancellerie royale (1241-1300)* de Paul Videsott (Strasbourg, 2015) (fig. 5).

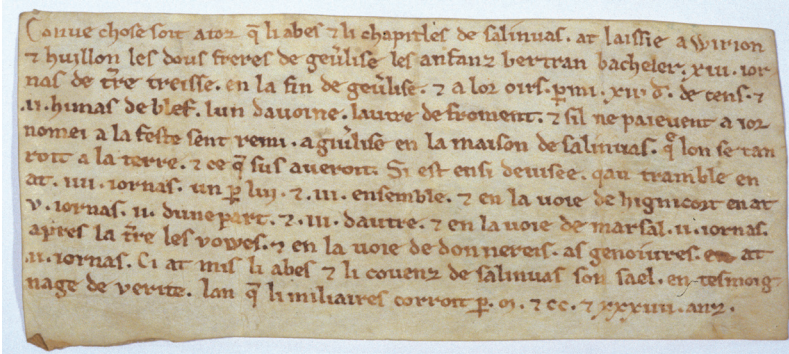


FIG. 4. – Facsimilé du manuscrit.

DOCUMENTS EN FRANÇAIS DE LA CHANCELLERIE ROYALE (1241-1300)

1 Looyz par la grace de Dieu rois de France, **2** à touz ceus à qui ces lettres vendront, saluz .

3 Nos ¹² faisons à savoir **4** que pais a esté porpallee de nostre chier cousin . et nostre fael Thibaut, roi de Navarre, et des Templiers, ¹³ **5** de ce que li Templier disoient . qu'il pueent aquerre en la conté de Champagne et de Brye et en ses fiez, . fors seg¹⁴ norie de chatel . et de cité, **6** que li Templier ne doivent jamés aquerre en la conté de Champagne et ¹⁵ de Brye ne es fiez de celle conté, . se ce n'estoit par le gré le conte de Champagne et de Brye .

7 Et des aquez *que* ¹⁶ li Templier ont fez . puis que li rois de Navarre fist pais à euls, il se sont acordé en tel maniere . *que* li rois de ¹⁷ Navarre, touz ces aquez . et les issues, . nos a bailliez **8** et nos les avons bailliez nostre amé et nostre fael . Johan ¹⁸ de Thorete, . en tel maniere que cil Jahans de Thorete les doit tenir et les yssues en sa main . dés Brandons ¹⁹ qui viennent en un an . **9** Et se la pais est tenue par l'a¹⁰ cort dou grant maistre du Temple d'Outremer et dou ¹⁰ couvent, . li Templier doivent avoir ces aquez . et ces issues . par le gré faisant le roi de Navarre. **10** Et se la pais ¹¹ n'est tenue, cil Jahans de Thorete doit reseisir le roi de Navarre de ces aquez et des issues, . **11** sauf le droit ¹² de chacun . qui demorra en toutes choses . en tel point cum il estoit au jor que la pais fu porpallee .

12 Et est à-savoir ¹³ que li Templier ne porront aquerre . en la conté de Champagne et de Brye ne en ses fiez . dedenz le terme ¹⁴ devant dit, . quar nos lor avons deffendu . et il distrent que non feroient il . **13** Et li aquest *que* li Templier ¹⁵ avoient fait devant la pais qu'il firent au roi de Navarre, . lor doivent demourer à totes les issues . **14** Ne ¹⁶ por sael *que* nos metains en ces lettres, . nulles choses ne lor sont comfermees . se la pais ne se fait . **15** Et ceste ¹⁷ pais devant dite doivent li Templier tenir se li grant maistres du Temple d'Outremer et li couvenz ¹⁸ le louent, . et li rois de Navarre la doit tenir s'il s'i acorde . **16** Et se la pais n'est tenue, . la chose sera ¹⁹ en tel point . comme devant en totes choses . **17** Et se la pais ne tient, ces lettres ne vaudront rien ou²⁰ tre les Brandons devant diz . **18** Et se ceste pais est tenue, . li rois de Navarre . et li Templier . nos ²¹ ont requis que nos la tesmognein ou confermain, . sauve la garde . au seignor de Champagne . ²² et de Brye .

19 Ce fu fait à Pontoise, **20** l'an de l'incarnation nostre Seignor . mil . deus cenz . et ²³ quarante et un, **21** ou mais d'aoust .

FIG. 5. – Édition d'un acte de la chancellerie royale (Videsott, 2015).

La part de programmation dans le projet actuel a occupé une place prépondérante et son financement est difficile puisque les institutions considèrent qu'il s'agit là d'une part préliminaire et non-scientifique du travail. Ainsi, la programmation de base n'a pu être achevée qu'en 2014, sous une architecture à la fois claire et fonctionnelle, après plus d'une décennie et une reprogrammation intégrale²⁶. Le système d'exploitation *Phoenix* a la qualité d'être utilisable par tout projet philologique d'une certaine complexité, même s'il faut prévoir, là encore, un budget certain pour son adaptation à d'autres ensembles textuels. L'informatique est, l'on pourrait dire, à la fois une bénédiction et une malédiction pour le philologue médiéviste.

Malgré ces contraintes, la mise en ligne des documents s'avère le vecteur essentiel pour dépasser l'éclatement de l'information linguistique intrinsèque aux *Documents linguistiques* depuis ses débuts. Dans l'édition de Paul Meyer, tout lexème doit être recherché dans quatre glossaires différents et l'on doit se reporter aux références indiquées pour chaque occurrence afin d'identifier le lieu et la date du texte. La même procédure s'impose pour les deux glossaires de Clovis Brunel ou les six glossaires (publiés) de l'époque Monfrin. Dans l'édition électronique actuelle, la recherche d'une forme apporte immédiatement l'ensemble des occurrences avec l'indication de la date, de la provenance, du contexte et, si elle a été identifiée préalablement, du rédacteur. On peut voir par exemple les formes lemmatisées du mot *mouture* (fig. 6).

Pour les recherches graphématiques ou morphologiques, la gestion parallèle des 3200 documents ouvre les mêmes perspectives permettant enfin de considérer cet ensemble documentaire comme une unité et un corpus de référence avec, actuellement, près de deux millions d'occurrences.

3. Volets de l'analyse linguistique

La part éditoriale représente le fondement de la grande entreprise des *Documents linguistiques* – mais elle n'est aucunement une fin en

26. Cf. notre mise au point « Le long chemin d'une charte vers l'ère multimédiale », dans *Lexikon, Varietät, Philologie, Romanistische Studien. Günter Holtus zum 65. Geburtstag*, A. Overbeck, W. Schweickard et H. Völker éd., Berlin-New York, 2011, p. 649-664.

All texts

Selected texts only (see Restrictions)

Occurrences	Date	Red.	Div	Context
<input type="checkbox"/>	chouai0468A	1270	-	30 ra au moelin et de tout prendre mouture ; et li sergans au signeur doit sir sour le
<input type="checkbox"/>	chouai0468A	1270	-	30 a le mosnerie , doit molre sans mouture . Et chis briés est fais par amendement des
<input type="checkbox"/>	chouai0468B	1270	-	5 au - moelin et de tout prendre mouture ; et li siergans au signeur doit sir sour l
<input type="checkbox"/>	chouai0468B	1270	-	5 - le mosnerie , doit morre sans mouture . Et chis briés est fais par amendement des
<input type="checkbox"/>	chouai0468C	1270	-	4 ra al moelin et de tout prendre mouture ; et li siergans au signeur doit sir sor le
<input type="checkbox"/>	chouai0468C	1270	-	4 a le mosnerie , doit more sans mouture . Et cis briés est fais par amendement des e
<input type="checkbox"/>	chV0144	1271	-	6 oller ausi com autre gens , nos moutures paiens et nostre argent . Et si li dis abbes
<input type="checkbox"/>	chH5058	1271	VigVil.	3 /. une quartre et une demie de - moutures à - la - mesure de - Vesoul que je avoie ou
<input type="checkbox"/>	chN 004	1290	-	12 . cent /. et . neuf . bichez de mouture /. et por /. vint . bichez d ' avoinne /. et
<input type="checkbox"/>	chN 004	1290	-	13 . Des quex /. cent /. bichez de mouture et . neuf . avec /. et . vint /. bichez d '
<input type="checkbox"/>	ChMM001	1232	EpMerzl	122 d ' amande les signors /. et la mouture .//. Et des musures que li signor ont detenu
<input type="checkbox"/>	chMM006	1237	SCommer.	12 ement à - toz Jorz , sanz paier mouture , aprez ceu qu ' il troveront en la tremene
<input type="checkbox"/>	chMM027	1243	EpVerdu.	3 lef , moitiel froment , moitiel mouture , chascun an en la partie l ' evesque de Ver
<input type="checkbox"/>	chMM041	1246	EpMerzl	5 if , moitié froment ,//. moitié mouture ,//. à la parfin en la justice le devant dit
<input type="checkbox"/>	chMM065	1249	AbbSVan.	5 , à Creveichamp , sanz paier /. mouture ,//. à la parfin en la justice le devant c
<input type="checkbox"/>	chMM065	1249	AbbSVan.	9 madamme Aeliz ne paera point de mouture de bleif au molin de Creveichamp le devant c
<input type="checkbox"/>	chMM066	1249	AbbSVan.	9 it priour ,//. et si i /. paera mouture de ferine .//. Et pour ce que ce soit ferme
<input type="checkbox"/>	chMM066	1249	PfFlavi.	5 ur , à Creveichamp , sanz paier mouture ,//. li devant diz priours de Flavignei , de
<input type="checkbox"/>	chMM066	1249	PfFlavi.	6 finage de Creveichamp /. et la mouture dou molin de ferine ,//. an tel mensire ke l
<input type="checkbox"/>	chMM066	1249	PfFlavi.	8 me Aeliz et sei oir paeroent la mouture de ferine de ce que il mouroent au molin le
<input type="checkbox"/>	chMM066	1249	PfFlavi.	8 t et ce ne paeroent /. point de mouture de bleif . Et pour ce que ce fut fait devant
<input type="checkbox"/>	chMM082	1251	EpMerzl	11 t dit /. paicent entierement la mouture .//. Et de toutes ces choses devant dites av
<input type="checkbox"/>	ChMM103	1254	EpVerdu.	5 son de Bouconville , sanz panre mouture ,//. et se la farine venoit à Viez Mostier ,

Browse Lemmata Types

Query/REGEX:

Include: p i c lp cp

d lat

null, **mouture** (63)

OK

FIG. 6. — Relevé d'un lexème lemmatisé dans les DocLing.

soi. Une édition se doit d'être utilisée en vue d'interprétations scientifiques, littéraires, linguistiques ou historiques. Le développement des volets d'analyse, avec toutes les implications qu'il comporte, nous a en effet bien plus occupé que l'édition des documents. Nous présenterons par la suite les trois domaines interprétatifs qui ont connu les avancées les plus importantes depuis 2000, la scriptologie, la lexicologie et l'histoire linguistique, en laissant de côté ici la syntaxe et l'onomastique, même si ces deux domaines sont tout aussi porteurs.

3. 1. INTERROGATIONS SCRIPTOLOGIQUES

La scriptologie est traditionnellement le domaine le plus attendu dans l'analyse de textes documentaires. En raison de leur caractère de témoins datables et même localisables, les chartes originales fournissent des renseignements précis sur la variance grapho-phonétique et morphologique à un moment et en un lieu donnés.

Or, le travail avec les *Documents linguistiques* a apporté non seulement de nouvelles données empiriques par l'importance du corpus et par la couverture systématique du territoire, mais il nous a également amené à développer de nouveaux concepts interprétatifs. Tout d'abord, nous nous sommes rendu compte, lors de l'analyse du corpus de la Meurthe-et-Moselle que la dimension déterminante pour le choix des variantes linguistiques n'est pas en premier lieu l'individu, c'est-à-dire le scribe, ni même un lieu géographique en tant que tel, mais la chancellerie ou le *scriptorium* qui assume la responsabilité de la rédaction de l'acte, soit ce que j'appelle le « rédacteur »²⁷. Dans une même ville comme Nancy par exemple, la nature linguistique des actes varie en fonction de ce « rédacteur », qui peut être le duc de Lorraine, l'église Saint-Epvre ou un scribe indépendant travaillant dans la ville. Parallèlement, une chancellerie comme celle des ducs de Lorraine, des comtes de Bar ou, mieux encore, des rois de France, comporte une dimension de « délocalisation » qui contredit – ou tout au moins complexifie – le rattachement de la langue à un lieu géographique bien défini.

27. Cf. « Les lieux d'écriture dans les chartes lorraines du XIII^e siècle », *Revue de linguistique romane* 72, 2008, p. 413-540.

Nous avons pu prouver que les scribes individuels, reconnaissables par leur calligraphie, se soumettent à la discipline de leur « lieu d'écriture » dans la mise en page des documents tout comme dans leurs choix grapho-phonétiques, morphologiques et lexicaux. Plus encore, un scribe peut intégrer dans la « norme » de son lieu d'écriture des éléments propres à une autre chancellerie, présente dans le même acte, pour mieux impliquer celle-ci dans le processus de rédaction et augmenter ainsi la validité du document. C'est ici, dans la microscopie, que l'on peut identifier le rôle catalyseur des actes dans une configuration médiatisée du pouvoir temporel et symbolique. Les protagonistes laïques et ecclésiastiques font appel à ce « parchemin performatif » pour constituer leur pouvoir et pour se positionner les uns par rapport aux autres. C'est le « *Sprachspiel* » de Wittgenstein qui est ici à l'œuvre de manière très frappante.

Les chancelleries et les *scriptoria* constitués, qui transmettent le savoir pratique de la scripturalité, sont ainsi la dimension réelle de l'élaboration de la langue écrite et de tous les processus évolutifs que cette langue connaîtra. La formation d'une norme linguistique se place précisément dans le cadre des grands lieux d'écriture et résulte des dialogues entre eux.

Nos analyses nous ont également amené à mieux cerner la nature de cette sous-discipline particulière de la philologie médiévale qu'est la « scriptologie »²⁸. Cette méthodologie, qui est particulièrement appréciée dans les pays de langue germanique, des Pays-Bas à l'Autriche, prend appui sur la variation diatopique et régionale de la langue médiévale et exploite ce levier pour identifier d'autres paramètres pertinents : la localisation et la datation de textes littéraires, religieux ou scientifiques, le rattachement des copies à un rédacteur défini, le prestige linguistique d'un texte ou encore des mouvements d'influence ou de dépendance entre les protagonistes de l'écrit. Les *Documents linguistiques* fournissent en cela un repère sûr pour l'époque du XIII^e siècle qui est celle où les constellations de la langue écrite médiévale se mettent en place dans le territoire d'oïl.

Actuellement, nous travaillons au balisage scriptologique de l'ensemble du corpus des *Documents linguistiques*. Cette opération,

28. Cf. « Trajectoires et perspectives en scriptologie galloromane », *Medioevo Romanzo* 36, 2012, p. 5-23.

déjà initiée il y a une décennie, a connu une longue interruption suite à la reprogrammation de notre base de données qui n'a pas affecté la part éditoriale, mais qui a freiné la part interprétative. Jusqu'ici, nous avons balisé, avec différents coéquipiers, quelque 250 000 formes d'un point de vue grapho-phonétique et morphologique. Le paramètre le plus exigeant que nous avons traité jusqu'ici concerne les résultats de 'a[accentué en syllabe libre, après palatale ou non, avant nasale ou non (donc les types lat. [1a] NAVE > fr. mod. *nef* / MATRE > *mère*, [1b] CARU > *chier* > *cher*, [2a] PANE > *pain*, [2b] CANE > *chien*). Même sans prendre en considération les formes devant nasale, ce paramètre concerne non moins de 50 000 formes individuelles structurées dans une quarantaine de configurations précises, selon l'entourage de la voyelle accentuée en question (fig. 7).

Au-delà de la précision descriptive des phénomènes en question, nous avons pu retenir deux observations d'un point de vue macroscopique. La première concerne la phonétique historique du français. L'on a toujours considéré que 'A[passe par une phase de diphtongue *['æ] puis se monophtongue en [e] (ou [ɛ]) en ancien français (MATRE > *mere*, HOSPITALE > *hostel*, NAVE > *nef*, etc.). Or, les *Documents linguistiques* montrent que l'évolution 'A[> [e ~ ɛ] caractérise certes les actes royaux de la deuxième moitié du XIII^e siècle, mais que toutes les autres régions connaissent, dans une mesure variable, le type diphtongué [ej], né dans une évolution parallèle par différenciation de la diphtongue initiale *['æ] (*meire*, *hosteil*, *neif*). Le type [ej] avait été considéré jusqu'ici comme une variante mineure et déviante, alors qu'il a été endémique sur l'essentiel du territoire de l'oïl, sauf – justement – en Île-de-France. Il s'agit par conséquent d'un des rarissimes paramètres phonétiques oraux qui semblent propres à Paris dans la formation d'une norme écrite du français et qui s'est ensuite imposé sur une partie importante du domaine d'oïl.

La deuxième observation concerne la scripturalité vernaculaire en domaine francoprovençal, entre Lyon, Grenoble, Genève, Fribourg et Cluny. Dans ce territoire, relativement important, il n'y a jamais eu d'élaboration systématique et généralisée de la langue autochtone par le biais de textes pleins en francoprovençal. Les scribes sont passés d'une *scripta* latine à une *scripta* française, en continuant d'y intégrer des éléments de leur langue maternelle francoprovençale. L'étude de la coexistence des deux systèmes linguistiques – le

Nr.	Variant	Occ.	Texts Co
»	all variants	49697	...
⊙	1110 non préc pal: ei < -ĀTU	4952	...
⊙	1111 non préc pal: eit < -ĀTU	192	...
⊙	1120 non préc pal: eie < -ĀTA	151	...
⊙	1121 non préc pal: eiey < -ĀTA	1	...
⊙	1130 non préc pal: eir < -ĀRE (inf.)	871	...
⊙	1140 non préc pal: eir < -ĀTRE-	291	...
⊙	1150 non préc pal: eil < -ĀLE	58	...
⊙	1151 non préc pal: queil, teil < -ĀLE	443	...
⊙	1160 non préc pal: ei < varia	203	...
⊙	1210 non préc pal: é < -ĀTU	3371	...
⊙	1211 non préc pal: é < -ĀTU (bis)	7689	...
⊙	1220 non préc pal: ee < -ĀTA	2168	...
⊙	1230 non préc pal: er < -ĀRE	3521	...
⊙	1231 lat. -ĀRE (inf.) > <é> (après phonème non palatal)	6	...
⊙	1240 non préc pal: er < -ĀTRE	2715	...
⊙	1250 non préc pal: el < -ĀLE	1496	...
⊙	1251 non préc pal: quel, tel < -ĀLE	3857	...
⊙	1252 non préc pal: el < rom. -al	30	...
⊙	1253 non préc pal: eus < -ĀLE pl.	10	...
⊙	1254 non préc pal: queus, teus < -ĀLE pl.	44	...
⊙	1260 non préc pal: é < varia	16	...
⊙	1261 non préc. pal: ez < -ĀTIS	1	...
⊙	1310 non préc pal: ié < -ĀTU	1720	...
⊙	1311 non préc pal: iet < -ĀTU	117	...
⊙	1320 non préc pal: iee < -ĀTA	15	...
⊙	1330 non préc pal: ier < -ĀRE	23	...
⊙	1410 non préc pal: iei < -ĀTU	3	...
⊙	1420 non préc pal: iee < -ĀTA	2	...
⊙	1510 non préc pal: a < ĀTU	1430	...
⊙	1520 non préc pal: a < ĀTA	1014	...
⊙	1521 non préc pal: aa < ĀTA	17	...
⊙	1530 non préc pal: ar < -ĀRE	1579	...
⊙	1540 non préc pal: ar < -ĀTRE	227	...
⊙	1541 non préc pal: ai < -ĀTRE	13	...
⊙	1550 non préc pal: al < -ĀLE	1360	...
⊙	1551 non préc pal: qual, tal < -ĀLE	482	...
⊙	1552 non préc pal: au < -ĀLE	18	...
⊙	1553 non préc pal: aul < -ĀLE	142	...
⊙	1554 non préc pal: aus < -ĀLE pl.	222	...
⊙	2110 après pal: ei < -ĀTU	288	...
⊙	2111 après pal: eit < -ĀTU	83	...
⊙	2120 après pal: eie < -ĀTA	8	...
⊙	2130 après pal: eir < -ĀRE	6	...
⊙	2160 après pal: ei < varia	4	...
⊙	2210 après pal: é < -ĀTU	794	...
⊙	2211 après pal: et < -ĀTU	29	...
⊙	2220 après pal: ee < -ĀTA	146	...
⊙	2230 après pal: er < -ĀRE	273	...
⊙	2310 après pal: ié < -ĀTU	1438	...
⊙	2311 après pal: iet < -ĀTU	76	...
⊙	2320 après pal: iee < -ĀTA	75	...
⊙	2330 après pal: ier < -ĀRE	1186	...
⊙	2360 après pal: ié < varia	199	...
⊙	2410 après pal: iei < -ĀTU	31	...
⊙	2411 après pal: iei < -ĀTU	5	...
⊙	2420 après pal: iee < -ĀTA	15	...
⊙	2510 après pal: a < ĀTU	24	...
⊙	2520 après pal: a < ĀTA	29	...
⊙	2530 après pal: ar < ĀRE	0	...
⊙	2610 après pal: ia < ĀTU	117	...
⊙	2620 après pal: ia < ĀTA	387	...
⊙	2630 après pal: iar < -ĀRE	2	...
⊙	3110 après /i j/: ei < -ĀTU	101	...
⊙	3111 après /i j/: eit < -ĀTU	2	...
⊙	3130 après /i j/: eir < -ĀRE	13	...
⊙	3210 après /i j/: iee < -ĀTU	1902	...
⊙	3211 après /i j/: iei < -ĀTU	51	...
⊙	3220 après /i j/: eee < -ĀTA	84	...
⊙	3230 après /i j/: eer < -ĀRE	1262	...
⊙	3231 après /i j/: eir < -ĀRE	3	...
⊙	3310 après /i j/: ié < -ĀTU	134	...
⊙	3311 après /i j/: iet < -ĀTU	22	...
⊙	3330 après /i j/: ier < -ĀRE	89	...
⊙	3410 après /i j/: iei < -ĀTU	1	...
⊙	3510 après /i j/: a < -ĀTU	4	...
⊙	3520 après /i j/: a < -ĀTA	18	...
⊙	3530 après /i j/: ar < -ĀRE	36	...

FIG. 7. – Classification des occurrences relevées dans les DocLing pour ‘A[non suivi de nasale.

français dominant et le francoprovençal sous-jacent – aux XIV^e et XV^e siècles est passionnante et n’a jamais été entreprise sur la base d’un corpus informatique. L’analyse quantitative permet ainsi de tracer avec grande précision la présence des résultats autochtones – le francoprovençal conserve ‘a[non précédé de palatale (*mare, ostal, cla* < CLAVE) – qui est nettement plus forte, à niveau microscopique, de ce que l’on a pu penser jusqu’ici.

Ce grand chantier, entrepris en collaboration avec l’université de Neuchâtel, fournira une nouvelle mesure dans l’analyse de tout texte en langue ancienne et précisera considérablement les évolutions phonétiques et morphologiques du français et du francoprovençal.

3. 2. ÉLABORATIONS LEXICOLOGIQUES

On sait que le vocabulaire médiéval est fortement segmenté selon les genres textuels. Une traduction du Nouveau Testament, une chirurgie, un fabliau, une chronique ou un registre de comptes mettent en relief différents ensembles lexicaux, tout en partageant un certain nombre de mots d'une langue dite « générale ». Les actes de la pratique juridique comportent surtout des termes de droit et de gestion patrimoniale, d'agriculture ainsi que de nombreux toponymes et de noms de personne pertinents du point de vue de l'analyse lexicologique.

Les entrées du *Dictionnaire du Moyen Français* de Robert Martin (DMF) ainsi que les lemmes du *Dictionnaire étymologique de l'Ancien Français* (DEAF) permettent d'évaluer le nombre de mots attestés en français médiéval à environ 80 000 dont une moitié sont des mots savants empruntés au latin ou encore à d'autres langues de contact. Les quelque 3200 documents des *Documents linguistiques* réunissent environ 12 000 lexèmes dont certains sont propres à l'écrit documentaire et rares dans d'autres genres textuels. Ce n'est donc qu'une partie du vocabulaire médiéval, mais c'est un ensemble important et significatif.

Pour rendre compte de ce trésor lexicologique, nous poursuivons deux voies complémentaires : la réalisation d'un répertoire général et de plusieurs analyses thématiques. Durant les quatre dernières années, nous avons lemmatisé, avec différents collaborateurs, l'intégralité du corpus oïlique et nous avons mis en œuvre la programmation d'une interface avec le DEAF électronique. Celle-ci permet désormais d'intégrer les lexèmes des *Documents linguistiques* dans les articles du DEAFpré. Dans un premier temps, ces mots ne sont ni sémantisés ni analysés, mais ils sont placés dans un cadre qui permet assez facilement leur interprétation et la comparaison avec des mots provenant d'autres genres textuels. Il est également possible de passer directement par hyperlien de cet avatar de dictionnaire électronique aux contextes des formes dans les *Documents linguistiques*. Voici un article du DEAFpré comportant en rouge les formes des DocLing (fig. 8).

Dans un deuxième temps, les lexèmes pourront naturellement être traités d'un point de vue sémantique, syntagmatique et variationnel.

The screenshot shows the DEAFpré website interface. At the top left, there are logos for 'DAGÉL' and 'DEAFplus'. A search bar contains the word 'mouture'. Below the search bar, there are radio buttons for 'DEAFplus' and 'DEAFplus + DEAFpré', with the latter selected. On the left side, there is a 'Résultats' section with a list of items: 'mouture' (selected) and 'mouturer'. The main content area displays the article for 'mouture f.' with a list of references and citations, including works by [FEW], [MED], [DCCarp], [TLF], [S], [N], [FEW], [Arnaldi], [DiStefLoc], [ChansArtB], [XXII], [MontRayn], [NoomenFabl], [chdouai], [moiture], [moltiei], [molture], [chMM], [chSL], [CensToulo], [JurésSouend], [DocHMarneG], [DocHMarneG], [ChMe], [ChHM], [CensToulo], [matures], [chV], [NyströmMén], [DocAubeC], [ChMM], [ChMe], [PiérardMons], [1°], [PiérardMons], [MontRayn], [NyströmMén], [Bev], [Drüppel], [Morlet], [sigle], [TLF], [Gdf], [MED], [FEW], and [2°].

FIG. 8. – Article *mouture* du DEAFpré.

Pour ce traitement, nous suivrons, dans le cadre du DEAFpré, le modèle du *Dictionnaire d'Ancien Gascon électronique* (DAGél).

Le DAG, que nous dirigeons depuis 2013, est le premier dictionnaire historique pour le gascon et repose presque exclusivement sur des sources documentaires, qui sont les seules disponibles pour le gascon médiéval. Son architecture électronique est le résultat d'une collaboration avec Sabine Tittel de Heidelberg et son système de rédaction représente une adaptation du DEAFpré²⁹. D'un point de vue informatique, la structure est (presque) identique, mais dans la représentation, l'on passe du dictionnaire synthétique et interprétatif qu'est le DEAF à un dictionnaire historique basé sur un corpus comme, par exemple, le DMF.

Voici une prise de vue du site du DAGél avec le début de la liste de mots actuellement traités à gauche (AB-) et un bref article à droite (agasc. *abilhar* substantivé dans le sens spécialisé de « revêtement, placage », en orfèvrerie) (fig. 9).

29. Cf. M. Glessgen, S. Tittel, « Le *Dictionnaire électronique de l'ancien gascon* (DAGél) », dans *Actes du XXVIII^e Congresso Internazionale di Linguistica e Filologia Romanza*, Rome, 2016, R. Antonelli, M. Glessgen et P. Videsott éd., Strasbourg, SLR-ÉLiPhi [à paraître, 2018].

Une représentation semblable est également possible à partir de la plateforme du DEAFpré. Elle sera alors parallèle au DAG et s'inscrira dans un *Glossaire des textes documentaires galloromans médiévaux* comportant aussi tous les lexèmes des *Documents linguistiques* francoprovençaux et occitans.

En guise de complément à ce traitement de type holistique, nous avons entrepris en 2014 une analyse systématique de la régionalité lexicale présente dans les *Documents linguistiques*. Avec mon regretté ami David Trotter – qui a parlé à l'Académie des Inscriptions et Belles-Lettres, il y a un peu plus de deux ans –, nous avons publié un volume collectif sur la *Régionalité lexicale au Moyen Âge* (Strasbourg, 2016), un aspect largement sous-estimé et sous-exploité dans le passé. Nous supposons que sur les 40 000 lexèmes non savants du français médiéval, près d'un quart devrait légitimement porter une marque régionale et appartenir, dans l'usage oral, à un espace plus ou moins étendu à l'intérieur du diasystème. Simplement, cet aspect n'est que faiblement apparent dans les textes écrits et encore moins dans les genres non-documentaires.

Actuellement, nous poursuivons avec Hélène Carles dans un projet du Fonds national suisse une analyse des quelque 500 lexèmes des *Documents linguistiques* qui sont caractéristiques de l'est et du sud-est du domaine d'oïl. Les résultats demanderont une reconsidération en profondeur des mécanismes de formation et de transformation, de diffusion et d'élaboration du vocabulaire français médiéval.

Dans tous nos travaux, nous défendons l'idée que le lexique forme une partie constitutive du langage et qu'il doit être étudié de manière systémique tout comme la phonologie, la morphologie et la syntaxe³⁰. Le grammaticocentrisme de la linguistique des cinquante dernières années a fortement porté préjudice à la discipline – et il a créé un vide dans l'analyse linguistique du vocabulaire médiéval. Pour cette raison, nous souhaiterions rendre hommage en ce lieu à la réalisation extraordinaire de Robert Martin qu'est le *Dictionnaire du Moyen Français* et qui est l'outil le plus précieux, à côté du *Französisches Etymologisches Wörterbuch* de Walther von Wartburg, de tout lexicologue médiéviste.

30. Cf. « Le statut épistémologique du lexème », *Revue de linguistique romane* 75, 2011, p. 337-416.

The screenshot shows the DAGÉL website interface. On the left is a vertical navigation menu with terms: abadessa, abadia, abaissar, abaluar, abaluat, abaria, abat, abatament, abater, abatut, abenir, aber, abescad. The main content area features a search bar with 'catégorie onomas.' and buttons for 'Exemples' and 'Recherche avancée'. Below the search bar are buttons for 'afficher tout' and 'masquer tout', and the text 'rédaction: Martin Glessgen'. The entry for 'ABILHAR' is displayed as 'inf. substantivé'. It includes several references: '[FEW 1,366b-367b gaul. *BILIA: agasc. *abilhar* seulement dans le sens général "arranger"; cf. aocc. *abilhar* "équiper"; le sens technique de l'orfèvrerie n'est pas répertorié.]', 'DOM *abilhar* "vétir" (= emprunt au fr.)', 'Base étymologique: afr. *abillier*.]', '(agasc. *abilhar* CConsRisclePC 1, p. 275[24])', and '1473, 1473 [Riscle] CConsRisclePC 1, p. 275[24] [..., loqual plom es en la man de Berdot Sobiran, metale, per lo abilhar de mete en la susdita poma;]'. A link '→ catégorie onomas.: tous les lemmes' is also present.

FIG. 9. – Article du DAGÉL.

3. 3. LINGUISTIQUE DIACHRONIQUE

Nous avons laissé de côté pour le moment l’analyse de la syntaxe des *Documents linguistiques* parce que les trois chantiers de l’édition, du balisage grapho-phonétique et de la préparation d’un glossaire nous occupent pleinement. Nous le regrettons toutefois parce que – sans être grammaticocentriste – ce domaine est potentiellement peut-être le plus novateur. La nécessité de décrire un acte juridique complexe de manière au moins partiellement décontextualisée comporte des exigences particulières pour la cohésion textuelle et la subordination. Si le vocabulaire des actes reflète avec un retard souvent de plusieurs siècles des formations lexicales qui s’échelonnent de manière constante entre l’époque antique et le Moyen Âge³¹, leur syntaxe reflète un processus d’élaboration en cours qui prend le relais des récits antérieurs en vers ou en prose.

Les *Documents linguistiques* sont ici significatifs pour le processus d’élaboration de la langue française, tout comme ils reflètent dans leur état grapho-phonétique les aléas de la formation d’une norme langagière. Gardons à l’esprit que les textes français antérieurs à *ca* 1160 proviennent presque exclusivement du contexte

31. Cf. H. Carles, *Trésor galloroman des origines (TGO). Les trajectoires étymologiques et géolinguistiques du lexique galloroman en contexte latin (ca 800-1120)*, Strasbourg, 2017.

anglo-normand, générés par les conditions communicatives particulières de l'Angleterre et inspirés par les modèles anglo-saxon. L'élaboration textuelle en Picardie – séparée de l'Angleterre par le seul canal – commence vers 1170 et implique dès 1204 la scripturalité documentaire qui est donc partie prenante dans la formation de ces normes écrites. Nous avons vu avec quelle attention les *scriptoria* et les chancelleries soignaient les aspects d'une norme « interne » qui avait comme finalité d'être reconnaissable par les professionnels de l'écrit en tout lieu.

Ces réflexions et les travaux de Paul Videsott sur la chancellerie royale m'ont amené à étudier de plus près la situation de la scripturalité vernaculaire à Paris³². J'ai pu voir que celle-ci débute précisément à l'époque de Louis IX avec plusieurs textes et traductions d'envergure :

- les traductions du *Corpus juris civilis* par le réseau universitaire de Paris (1250/1280) ;
- la première traduction intégrale de la bible en Occident avec la *Bible française de Paris* (1250/1260) ;
- la *Chronique de Saint Denis* (1270/1280) ;
- les actes de la prévôté (1260-) et du Parlement de Paris ;
- enfin, les premiers actes royaux rédigés en français (1241, 1254, 1259, 1268)³³.

Nous avons essayé de montrer que c'est à cette époque et dans ce contexte scriptural concret que se réalise la formation d'une norme écrite du français qui sera ensuite diffusée, à partir surtout de 1330, dans tout l'espace oïlique. La chronologie de cette genèse d'une langue française écrite est restée jusqu'ici pratiquement inaperçue et, même si notre argumentation dépasse clairement le contexte des *Documents linguistiques*, elle n'est devenue possible que grâce à ce corpus de référence.

4. Apport du projet à la linguistique (diachronique)

En conclusion, l'utilité scientifique du corpus et de son exploitation linguistique – et historique – nous semble réelle. Les *Documents*

32. Cf. « La genèse d'une norme en français au Moyen Âge : mythe et réalité du "francien" », *Revue de linguistique romane* 81, 2017, p. 313-398.

33. Cf. *op. cit.* (n. 26), p. 379 et suiv.

linguistiques allient en effet la qualité d'une fiabilité philologique et éditoriale dans la meilleure tradition de l'École des chartes à une emprise systématique sur la documentation disponible.

Les possibilités d'interrogation informatique permettent toute forme de quantification et d'interfaces. Le programme *Phoenix* qui gère les bases de données est un outil bien adapté à l'édition et à l'analyse de textes médiévaux comportant une certaine complexité philologique. Comme nous avons pu le montrer, il permet la préparation de glossaires et d'analyses scriptologiques, de même que l'intégration des données dans la lexicographie électronique.

D'un point de vue linguistique, les *Documents linguistiques* fournissent une base nouvelle pour la connaissance de la langue médiévale. Ils permettent notamment d'appréhender la variance de manière précise et quantifiée et non pas de façon ponctuelle et impressionniste. Ils ouvrent le regard sur la diversité géolinguistique du domaine d'oïl, permettant aussi d'identifier les invariants dans l'usage linguistique. L'emprise holistique sur le vocabulaire amène une vue plus systémique sur le lexique et ouvre la voie à des exploitations lexicologiques interprétatives qui font largement défaut dans notre discipline.

De manière plus générale, les analyses linguistiques dans les différents domaines du langage apportent une meilleure compréhension des phénomènes d'élaboration au Moyen Âge. Il est également possible d'établir des liens avec ce que la linguistique appelle l'histoire « externe », donc avec l'histoire tout court.

Les *Documents linguistiques* associent une grande tradition d'études à des perspectives novatrices. C'est également un projet qui « fait l'humilité », comme dirait Jacques Brel, et qui demande un esprit d'équipe. Dans ce sens, nous souhaiterions rendre hommage à nos illustres prédécesseurs et aux nombreux amis, collègues et jeunes chercheurs qui ont contribué et continuent à contribuer à l'avancement de ce grand chantier, portés par une passion commune.

*

* *

Le Vice-Président Jean-Louis FERRARY, MM. Philippe CONTAMINE, Robert MARTIN, Charles DE LAMBERTERIE, ainsi que M. Olivier Soutet, correspondant français de l'Académie, interviennent après cette communication.

SIGLES DES DICTIONNAIRES CITÉS

- DAG = *Dictionnaire onomasiologique de l'ancien gascon*, dirigé par Kurt Baldinger, en collaboration avec Inge Popelar, puis Nicoline Winkler, Tübingen, Niemeyer, fasc. 1 à 10, 1975-2002 ; continué depuis 2003 par N. Winkler.
- DAGél = *Dictionnaire de l'ancien gascon*, dirigé par Martin Glessgen en collaboration avec Sabine Tittel, 2014- [mise en ligne prévue en 2018].
- DAO = *Dictionnaire onomasiologique de l'ancien occitan*, dirigé par Kurt Baldinger, en collaboration avec Inge Popelar, puis Nicoline Winkler, Tübingen, Niemeyer, fasc. 1 à 10, 1975-2007.
- DEAF = *Dictionnaire étymologique de l'ancien français*, fondé par Kurt Baldinger, continué par Frankwalt Möhren, puis par Thomas Städtler, Tübingen-Berlin-Boston, Niemeyer-De Gruyter, 1974- [www.deaf-page.de].
- DEAFBibl = *Dictionnaire étymologique de l'ancien français. Complément bibliographique*, Frankwalt Möhren, 2016³, Berlin-Boston, De Gruyter.
- DEAFpré = *Dictionnaire étymologique de l'ancien français*, Matériaux lemmatisés en ligne, préparés par l'équipe du DEAF [<http://www.deaf-page.de>].
- DMF = *Dictionnaire du Moyen Français (DMF 2015)*, dirigé par Robert Martin, Robert et Sylvie Bazin-Tacchella, Nancy, ATILF/CNRS-université de Lorraine, 2015 [www.atilf.fr/dmf].
- DOM = *Dictionnaire de l'occitan médiéval*, fondé par Helmut Stimm, continué par Wolf-Dieter Stempel, puis par Maria Selig, Tübingen, Niemeyer, fasc. 1 à 8, 1996-2016 ; version électronique : www.dom-en-ligne.de.
- FEW = *Französisches Etymologisches Wörterbuch. Eine Darstellung des galloromanischen Sprachschatzes*, dirigé par Walther von Wartburg *et alii*, Leipzig-Bonn-Bâle, Teubner-Klopp-Zbinden, 25 vol., 1922-2002, [<https://apps.atilf.fr/lecteurFEW>].
- FEWCompl = *Französisches Etymologisches Wörterbuch. Eine Darstellung des galloromanischen Sprachschatzes. Complément*, 3^e édition, publiée par Jean-Paul Chauveau, Yan Greub et Christian Seidl, Strasbourg, ÉliPhi-SLR, 2010.
- TLF = *Trésor de la langue française, Dictionnaire de la langue du XIX^e et du XX^e siècle (1789-1960)*, dirigé par Paul Imbs et Bernard Quémada, Paris, Éditions du CNRS-Gallimard, 16 vol., 1971-1994.
- TLFi = CNRS/ATILF, *Trésor de la langue française informatisé* [<http://atilf.atilf.fr>].