# Corpus/Corpora

## Tra materialità e astrazione

*a cura di*
Martina Albertini
Dayron Carrillo Morell
Sara Ferrilli
Alberto Giudici
Seraina Montigel
Stefano Negrinelli
Claudia Tassone

*Contributi di*
Elisa Attanasio, Alessandra Bossone, Valeria Carrieri
Stefano Cristelli, Claudia Dell'Uomo d'Arme, Alessandro Fiorillo
Lise Fuertes, Lorraine Fuhrer, Yelena Kondrashova Sayko
Andrea Menozzi, Mirko Mondillo, Sara Natale
Romina Irene Palacios Espinoza, Noemi Pigini
Andrea Polegato, Giovanni Spalloni, Mario Wild

Aracne editrice

Copyright © MMXXI

ISBN 978-88-255-4095-6

*I diritti di traduzione, di memorizzazione elettronica,*
*di riproduzione e di adattamento anche parziale,*
*con qualsiasi mezzo, sono riservati per tutti i Paesi.*

*Non sono assolutamente consentite le fotocopie*
*senza il permesso scritto dell'Editore.*

I edizione: **Roma**, aprile 2021

*Deditgia*
*agl seminari rumantsch dall'Universitad da Turitg*
*tgi ò festivo igl onn 2019 sies giubileum da 125 onns*

# Indice

11 *Intro:* Corpus et corpora *en corporativitad*
*Dayron Carrillo-Morell, Alberto Giudici*

**Sección I**
***El cuerpo (político) en el Arte***

21 «Corpo fenomenico» e «corpo senza organi». Linee, soglie e ripiegamenti tra filosofia e letteratura
*Elisa Attanasio*

35 Il problema del corpo. *Ragazzi di vita* e *Porcile*
*Alessandro Fiorillo*

49 Representaciones del cuerpo desde la perspectiva del constructivismo social. Un diálogo Foucault-Nochlin
*Yelena Kondrashova Sayko*

63 Il corpo come metafora politica nel primo Machiavelli
*Andrea Polegato*

**Section II**
***Le corps dans la littérature***

83 *Corpora violant*: le *Ritratto del Novecento* de Sanguineti entre Benjamin et Arendt
*Claudia Dell'Uomo d'Arme*

103 Le corps des personnages comme frontière allégorique dans le *Livre d'Artus* (XIII$^{e}$ siècle)
*Lise Fuertes*

115 Reificare e riscattare. Il “corpo” di Walter Siti tra *Troppi paradisi* e *Bruciare tutto*
*Mirko Mondillo*

135 Narraciones de una paridad intraducible. La dinámica *Leib – Körper* en *El último cuerpo de Úrsula* (2009), de la escritora peruana Patricia de Souza
*Romina Irene Palacios Espinoza*

**Sezione III**
**Corpora *linguistici***

149 L’exploitation des *Documents linguistiques galloromans* en graphématique et lexicologie
*Alessandra Bossone, Lorraine Fuhrer*

165 «E cche ccianno in ner corpus?». Il progetto *GSR* e la sua base di dati
*Stefano Cristelli, Mario Wild*

181 Che fare del proprio strano corpus? Verso la prima banca dati dell’italiano in caratteri ebraici (il *corpus* ICE)
*Sara Natale*

**Section IV**
**Corpus *de manuscrits***

195 La presenza dei corpi nel *corpus* dialogico dei trovatori
*Valeria Carrieri*

209 Nuovi testimoni manoscritti del *Fiore di virtù*
*Andrea Menozzi*

219 Fonti e rimandi nel *Dialogo* di Caterina da Siena: il contributo del *corpus* OVI. Prime riflessioni sul lessico cateriniano
*Noemi Pigini*

229 I volgarizzamenti fiorentini delle lettere politiche del Duecento: varianti di struttura
*Giovanni Spalloni*

243 *Abstracts*

249 Schede delle autrici e degli autori

255 Indice dei nomi

*Corpus/Corpora*
ISBN 978-88-255-4095-6
DOI 10.4399/978882554095610
pp. 149-163 (aprile 2021)

# L'exploitation des *Documents linguistiques gallo-romans* en graphématique et lexicologie*

de ALESSANDRA BOSSONE[1], LORRAINE FUHRER[2]

L'analyse des spécificités des langues, anciennes et modernes, se réalise désormais au moyen d'outils numériques donnant accès à des corpus textuels ou oraux de grande ampleur, tout en permettant des interrogations diverses[3]. Dans cette optique, cette contribution est pour nous l'occasion de montrer les possibilités d'exploitation linguistique d'un corpus numérique de textes médiévaux appartenant au genre documentaire : les *Documents linguistiques galloromans* (DocLing). En présentant les opérations concrètes qui peuvent y être effectuées, nous aborderons deux possibilités d'analyses qu'offre ce corpus : l'analyse graphématique et l'analyse lexicologique.

Le corpus des DocLing est à la fois une base de données très riche et un outil performant pour la réalisation d'analyses linguistiques complexes. Il offre des possibilités d'interrogations et d'annotations graphématiques permettant d'aborder des questions centrales de la recherche en linguistique historique : l'élaboration scripturale des langues galloromanes, le dialogue entre les différentes *scriptae* régionales du domaine galloroman ou encore le processus de standardisation du français. Réalisé sur des données philologiquement assurées et d'un accès optimal, il permet aussi des analyses lexicales axées sur plusieurs perspectives de recherche : l'identification des premières attestations,

* Les chapitres 1, 2 et 5 de cette contribution ont été rédigés conjointement par ses deux auteures. Lorraine Fuhrer est responsable de la réalisation et de la rédaction de l'analyse graphématique (chapitre 3) et Alessandra Bossone de l'analyse lexicale (chapitre 4).

[1] Université de Neuchâtel et EPHE.

[2] Universités de Neuchâtel et de Strasbourg.

[3] Voir par exemple C. PUSCH/W. RAIBLE, *Romanistische Korpuslinguistik I : Korpora und gesprochene Sprache*, Gunter Narr, Tübingen 2002; C. PUSCH/J. KABATEK/W. RAIBLE, *Romanistische Korpuslinguistik II : Korpora und diachrone Sprachwissenschaft*, Tübingen, Gunter Narr, Tübingen, 2005.

l'étude approfondie du lexique spécialisé parfois non répertorié dans les dictionnaires et aussi l'étude de la régionalité lexicale rendue possible grâce à l'ample couverture géographique.

Nous verrons que les analyses graphématiques combinées aux analyses lexicales permettent d'observer les interactions entre les trois domaines galloromans, en fonction surtout du *continuum* linguistique qui les caractérise.

## 1. Le corpus électronique des DocLing

Le corpus électronique des DocLing, disponible et interrogeable en ligne[4], réunit à l'heure actuelle *ca* 4 000 textes de type documentaire compris entre le 13e et le 15e siècle et provenant des trois domaines linguistiques de la *Galloromania*[5].

Les documents qu'il contient ont été édités selon les critères philologiques les plus actuels et ont fait l'objet d'analyses diplomatiques liées au contexte de genèse de chaque document (type de document, objet, acteurs concernés, rédacteur, etc.). Le corpus comprend, pour l'essentiel, des chartes de diverses typologies (donation, vente, accensement, etc.) et dans une certaine mesure des documents de gestion interne (comptes, tarifs de péages, etc.) et également des testaments.

Les DocLing des domaines oïlique et francoprovençal sont organisés en plusieurs corpus dont les lieux de conservation sont, pour les documents français, les archives départementales et, pour les documents suisses, les archives cantonales. Pour l'aire oïlique les chartes couvrent toute l'aire orientale comprenant la Picardie, la Champagne, la Lorraine, la Bourgogne et la Franche–Comté, le Jura suisse, ainsi que les chartes royales. Pour l'aire francoprovençale suisse sont intégrés les documents des cantons de Fribourg, Vaud/Genève, Neuchâtel et

[4] Ce projet a été entrepris à l'École nationale des chartes en 1890 par Paul Meyer. L'entreprise a été poursuivie dans les années suivantes par Clovis Brunel d'abord, et Jacques Monfrin ensuite. En dernier lieu, Martin Glessgen a repris la direction du projet en développant sa version électronique. Pour plus de détails, nous renvoyons aux publications de Glessgen sur le sujet (cf. notamment M.–D. GLESSGEN, *Les "documents linguistiques" de la France Histoire, présent et perspectives d'un projet centenaire*, comptes rendus des séances de l'année 2017 juillet–octobre, Académie des inscriptions & Belles–Lettres, Paris 2017) et aux différentes pages du site : http://www.rose.uzh.ch/docling.

[5] L'intégration de l'aire occitane étant dans une phase préparatoire, le domaine occitan ne sera pas pris en compte ici.

Berne ; pour la partie française ceux de l'Ain, du Forez, de Grenoble et du Lyonnais[6]. À cela il faut ajouter la *Somme du Code* francoprovençale provenant de la région de Grenoble.

L'étendue géographique sur deux domaines linguistiques permet, comme nous le verrons par la suite, de réaliser des analyses linguistiques de type contrastif. Ces dernières sont effectuées au moyen de Phoenix2[7], un outil performant qui offre des possibilités d'interrogations linguistiques approfondies. Concrètement, il est possible de réaliser des balisages graphématiques, morphologiques, ainsi qu'une lemmatisation du lexique. En raison du caractère évolutif du projet, chacune de ces opérations est susceptible d'être en cours de réalisation, seule la lemmatisation des corpus oïliques est terminée.

## 2. Les *scriptae* galloromanes

La production vernaculaire galloromane médiévale représente un ensemble de langues écrites que la recherche a appelées *scriptae*[8]. Sous cette dénomination on entend une langue écrite qui appartient, pour reprendre la terminologie de Koch/Oesterreicher[9], au code graphique.

L'élaboration des *scriptae* est due à des professionnels de l'écrit dont l'activité se déroule dans des *scriptoria* (= lieux d'écriture). Cette élaboration se base sur plusieurs modèles préexistants : le modèle oral, la tradition textuelle latine, la tradition pré–textuelle romane et enfin la *scripta* supra–régionale découlant d'un processus de neutralisation

[6] En Valais, le latin est resté la langue de l'administration jusqu'à la fin de l'ancien régime. Quant à la Savoie, on y observe le maintien du latin jusqu'au début du 16e siècle (A. KRISTOL, *Francoprovençal*, in : A. LEDGEWAY/M. MAIDEN (éds.), « The Oxford Guide to the Romance Languages », Oxford University Press, Oxford 2016, p. 350).

[7] Phoenix2 est le logiciel développé dans le cadre de la base des DocLing. Pour une description sommaire de celui–ci, voir le site des DocLing, sous l'onglet 'Descriptif du projet — architecture de la base de données'.

[8] L. Remacle, *Le problème de l'ancien wallon*, Faculté de Philosophie et Lettres, Liège 1948, est à l'origine du terme et a été le premier à aborder cette question en basant sa recherche sur la documentation wallonne.

[9] P. KOCH/W. OESTERREICHER, *Sprache der Nähe — Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte*, in « Romanistisches Jahrbuch » 36 (1985), pp. 15–43; P. KOCH/W. OESTERREICHER, *Gesprochene Sprache und geschriebene Sprache / Langage parlé et langage écrit*, in : G. HOLTUS/M. METZELIN/C. SCHMITT, *Lexikon der Romanistischen Linguistik (LRL)*, volume 1/2, Niemeyer, Tübingen 2001, pp. 584–627.

inhérent à la langue écrite[10]. L'influence conjointe de chacun de ces modèles génère des *scriptae* composites caractérisées par une forte variance interne, qui se manifeste notamment à travers des diatopismes graphiques et lexicaux. Ainsi, les *scriptae* galloromanes sont régionalement définies et présentent un plus ou moins important degré de variation en fonction de plusieurs paramètres, tels que le genre textuel du document, sa date de rédaction ou encore son rédacteur[11].

L'étendue géographique des DocLing permet d'accéder à de nombreuses *scriptae* galloromanes des 13e, 14e et 15e siècles. Chacune d'entre elles se caractérise par des spécificités qui lui sont propres, tout en partageant des éléments en commun avec les *scriptae* des régions voisines. Dans ce sens, le tissu textuel galloroman forme un *continuum* linguistique au sein duquel les frontières scriptologiques restent floues. Or, grâce aux DocLing, nous pouvons décrire les particularismes graphiques et lexicaux de chacune des *scriptae* qu'ils contiennent, tout en atteignant une meilleure compréhension de leurs interactions.

## 3. L'exploitation graphématique

À l'aide des recherches antérieures[12], nous avons établi une liste de paramètres pour lesquels des balisages graphématiques des DocLing ont été réalisés avec Phoenix2. Ces balisages permettent de dégager des

[10] Cf. Y. GREUB, *Sur un mécanisme de la préstandardisation de la langue d'oïl*, in « Bulletin de la Société de linguistique de Paris » 102 (2007), pp. 429–434; M.–D. GLESSGEN, *La genèse d'une norme en français au Moyen Âge : mythe et réalité du 'francien'*, in « RLiR » 81 (2017), pp. 313–397.

[11] Cf. M.–D. GLESSGEN, *Les lieux d'écritures dans les chartes lorraines du XIIIe siècle*, « RLiR » 72 (2008), pp. 413–540.

[12] Par exemple C.T. GOSSEN, *Französische Skriptastudien : Untersuchungen zu den nordfranzösischen Urkundensprachen des Mittelalters*, H. Böhlaus, Wien 1967; H. GOEBL, *"Le Rey est mort, vive le Roy". Nouveaux regards sur la scriptologie*, in « TraLiLi » 13 (1975), pp. 145–210 ; H. GOEBL, *Qu'est–ce que la scriptologie ?*, in « Medioevo Romanzo » 2 (1975), pp. 3–43; A. DEES, *Atlas des formes et des constructions des chartes françaises au 13e siècle,* Niemeyer, Tübingen 1980; A. DEES, *Atlas des formes linguistiques des textes littéraires de l'ancien français*, Niemeyer, Tübingen 1987; M.–D. GLESSGEN, *Les lieux d'écritures dans les chartes lorraines du XIIIe siècle*, cit.; H. VÖLKER, *Skripta und Variation. Untersuchngen zur Negation und zur Substantivflexion in altfranzösischen Urkunden des Grafschaft Luxemburg (1237–1281)*, Niemeyer, Tübingen 2003; K. GRÜBL, *Varietätenkontakt und Standardisierung im mittelalterlichen Französisch. Theorie, Forschungsgeschichte und Untersuchung eines Urkundenkorpus aus Beauvais (1241–1455)*, Narr Verlag, Tübingen 2014; P. VIDESOTT, *Les plus anciens documents en français de la chancellerie royale capétienne (1241–1300). Présentation et édition*, ELiPhi, Strasbourg 2015.

tendances quant aux choix des scripteurs des documents et par là de mieux appréhender le tissu scriptologique galloroman, tout en fournissant des données empiriques inhérentes à une recherche basée sur un corpus textuel.

D'un point de vue méthodologique, chacune de nos analyses comprend trois étapes. Dans un premier temps, il s'agit de définir le changement phonétique qui nous intéresse dans toute la *Galloromania*. Ensuite, il est question d'établir une liste des notations possibles et d'en proposer une hiérarchie. Enfin, nous réalisons le balisage graphématique dans Phoenix2.

C'est par la présentation des graphies des issues du latin –'OCU — présent dans les mots 'FOCU et 'LOCU — que nous souhaitons illustrer ce type d'analyse. Notre discussion sera axée sur les caractéristiques les plus neutres, mais aussi les plus diatopiquement marquées des différentes graphies observées dans les DocLing.

Dans ['ɔku], l'amuïssement de [k] intervocalique, propre aux domaines oïlique et francoprovençal, précède la diphtongaison spontanée de ['ɔ] vers ['uo̯][13]. Ceci génère une triphtongue de coalescence [*'uɔu̯] commune aux deux domaines linguistiques. Généralement, dans l'aire oïlique, [*'uɔu̯] se réduit à [jø] dans les issues de ['lɔku] pour aboutir à [ljø] (*lieu*) et à [ø] dans les issues de ['fɔku] aboutissant à [fø] (*feu*)[14]. Pour ce qui concerne le domaine francoprovençal, la triphtongue [*'uɔu̯] se réduit à ['ue̯] dans les deux exemples : ['lɔku] > [lue̯] (*lue*) et ['fɔku] > [fue̯] (*fue*)[15]. En revanche, en territoire occitan, la consonne intervocalique est maintenue, ce qui occasionne une diphtongaison conditionnée de ['ɔ] vers ['uo̯], attestée sous des notations de type *fuoc* [fuo̯k][16]. Nous verrons qu'il s'agit de changements phonétiques qui ne sont pas forcément représentés par l'entier des trois domaines. En réalité, chaque région est susceptible de correspondre à un changement phonétique différent et c'est justement sur ces divergences que reposent nos analyses.

[13] H. HAFNER, *Grundzüge einer Lautlehre des Altfrankoprovenzalischen*, Francke, Bern 1955, pp. 112–114; J.–M. PIERRET, *Phonétique historique du français et notions de phonétique générale*, Peeters, Louvain–la–Neuve 1994[2], p. 190.

[14] J.–M. PIERRET, *op. cit.*, §418.

[15] H. HAFNER, *op. cit.*, pp. 112–114.

[16] Å. GRAFSTRÖM, *Étude sur la graphie des plus anciennes chartes languedociennes avec un essai d'interprétation phonétique*, Almqvist & Wiksells Boktryckeri AB, Uppsala 1958, pp. 75, 79.

Les changements phonétiques définis, nous pouvons dégager les notations susceptibles d'apparaître dans les DocLing. Pour ce changement, nous avons pu identifier onze notations : ‹ue›, ‹uo›, ‹ua›, ‹eu›, ‹ieu›, ‹ueu›, ‹ou›, ‹iou›, ‹uou›, ‹u›, ‹iu›. Dans Phoenix2, elles ont été hiérarchisées de la manière suivante : n000 = traitement de la voyelle tonique ; 0n00 = absence ou présence de ‹i, y, u› ; 9999 = varia[17]. Pour ce paramètre, notre hiérarchisation, suivie du balisage des formes, a abouti au tableau suivant :

| | Nr. | Variant | Occ. | Texts | Corp. |
|---|---|---|---|---|---|
| | » | all variants | 2155 | ... | ... |
| ✓ | 1100 | ue | 815 | ... | ... |
| ✓ | 2100 | uo | 25 | ... | ... |
| ✓ | 3100 | ua | 42 | ... | ... |
| ✓ | 4100 | eu | 908 | ... | ... |
| ✓ | 4200 | ieu | 228 | ... | ... |
| ✓ | 4300 | ueu | 4 | ... | ... |
| ✓ | 5100 | ou | 3 | ... | ... |
| ✓ | 5200 | iou | 9 | ... | ... |
| ✓ | 5300 | uou | 4 | ... | ... |
| ✓ | 6100 | u | 2 | ... | ... |
| ✓ | 6200 | iu | 97 | ... | ... |
| ✓ | 9999 | varia | 18 | ... | ... |

Ce tableau extrait de Phoenix2 permet de voir la hiérarchisation des notations, mais également la dernière étape du travail : le balisage de celles–ci. Celui–ci nous permet de dégager une première quantification, puis de regarder dans le détail la répartition transversale des notations que nous présentons dans le tableau ci–dessous. Ce dernier expose la répartition des différentes notations en fonction de chaque corpus pour lequel nous avons réalisé le balisage, tout en opérant une distinction entre les corpus du domaine oïlique et ceux du domaine francoprovençal[18] :

[17] Dans chaque paramètre, il a été nécessaire d'ajouter une catégorie 'varia' comprenant des notations isolées qui ne pouvaient pas être intégrées dans la structure hiérarchique définie.

[18] Pour chaque corpus, nous avons rempli en gris foncé les cases qui correspondent aux notations les plus fréquentes et en gris clair les cases qui correspondent aux notations plus rares, mais également bien attestées. Nous avons laissé en blanc les cases représentant les notations très rares.

| *Corpus* | ‹ue› | ‹uo› | ‹ua› | ‹eu› | ‹ieu› | ‹ueu› | ‹ou› | ‹iou› | ‹uou› | ‹u› | ‹iu› | varia | *Tot.* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Corpus du domaine oïlique* | | | | | | | | | | | | | |
| chDo | | | | 3 | 16 | | | | | | 53 | | 72 |
| chRoy | | | | 44 | 57 | | | | | | 2 | | 103 |
| chMa | | | | 68 | 30 | | | | | | 20 | 1 | 119 |
| chMe | 1 | | | 52 | 1 | | | | | | 10 | | 64 |
| chMM | | | | 55 | 3 | | | | | 1 | | 1 | 60 |
| chV | | | | 35 | 6 | | | | | | | 1 | 42 |
| chHM | 1 | | | 77 | 1 | | | | | | 5 | 1 | 85 |
| chN | | | | 18 | 10 | | | | | | | | 28 |
| chCOr | 4 | | | 95 | | | | | | | 1 | 1 | 101 |
| chHS | | | | 47 | | | 3 | | | | 2 | 3 | 55 |
| chJu | | | | 47 | | | | | | 1 | | | 48 |
| chSL | 15 | | 1 | 60 | 5 | 1 | | | | | 2 | 2 | 86 |
| chJuBe | 468 | | | 123 | 1 | | | | | | | 2 | 594 |
| *Corpus du domaine francoprovençal* | | | | | | | | | | | | | |
| chNCh | 55 | | | 135 | 13 | | | 6 | | | | 3 | 212 |
| chFrib | 72 | 25 | | 3 | 28 | 3 | | 3 | 1 | | | | 135 |
| chVauGe | 14 | | | 46 | 15 | | | | 3 | | 2 | | 80 |
| docAin | 64 | | 20 | | | | | | | | | 1 | 85 |
| docLyon | 11 | | 21 | | | | | | | | | | 32 |
| docForez | 64 | | | | 42 | | | | | | | 2 | 108 |
| docGren | 11 | | | | | | | | | | | | 11 |
| Codi | 35 | | | | | | | | | | | | 35 |
| *Tot.* | 815 | 25 | 42 | 908 | 228 | 4 | 3 | 9 | 4 | 2 | 97 | 18 | 2155 |

La notation ‹eu› est la plus fréquente (908 occ., *ca* 42%), mais elle est absente des corpus du domaine francoprovençal français (docAin, docLyon, docForez, docGren, Codi). Ceci prouve que ces corpus répondent à d'autres habitudes scripturales que les corpus oïliques et que les corpus de la Suisse romande francoprovençale. La notation ‹eu› est une notation qui s'est développée en domaine oïlique et qui a été adoptée par le système graphique du français dans le substantif *feu*. Sa présence en domaine francoprovençal suisse romand (chNCh, chFrib, chVauGe) illustre un emprunt au système graphique oïlique.

La notation ‹ue› est celle qui est la plus fréquente après ‹eu› (815 occ., *ca* 38%). Elle se limite surtout aux corpus du domaine oïlique sud–oriental (chSL, chJuBe) et aux corpus du domaine francoprovençal (chNCh, chFrib, chVauGe, docAin, docLyon, docForez, docGren, Codi). Elle doit ainsi être considérée comme une notation diatopiquement marquée pour les deux domaines linguistiques où elle dénote un changement phonétique commun, soit la réduction de la triphtongue vers ['ue̯].

La notation ‹ieu› est aussi une notation fréquente (228 occ., *ca* 10%). Très présente dans les corpus de Douai, des chartes royales et de la Marne, elle est également bien attestée par certains corpus francoprovençaux (chNCh, chFrib, chVauGe, docForez) et dans le corpus de la Nièvre (chN). Alors que dans les corpus oïliques elle est une notation neutralisée représentant un traitement phonétique de 'LOCU propre à une zone plutôt septentrionale de ce domaine, en domaine francoprovençal elle illustre une notation d'emprunt au système graphique du français, s'expliquant par l'influence de la *scripta* exemplaire de Paris à partir de la moitié du 14e siècle[19]. Son absence dans les autres corpus oïliques s'explique par le fait que les documents qui les composent sont anciens et reflètent des habitudes graphiques qui précèdent la constitution de cette *scripta* exemplaire.

Quant aux notations qui sont plus rares (‹iu›, ‹ua›, ‹uo›, ‹ueu›, ‹ou›, ‹iou›, ‹uou› et ‹u›), elles se limitent à des zones restreintes du domaine et correspondent à des traitements phonétiques locaux qui ont bénéficié d'une tentative de mise à l'écrit sans lendemain.

- La notation ‹iu› est septentrionale et picardisante. Sa présence dans d'autres corpus que celui de Douai se réfère probablement à des documents dont la *scripta* est elle aussi picardisante.
- La notation ‹ua› est francoprovençale et uniquement attestée dans les corpus du Lyonnais et de l'Ain. Sa présence dans le corpus de la Haute–Saône signifie probablement que le document qui la contient comporte des diatopismes francoprovençaux et mériterait une analyse approfondie sous le prisme du francoprovençal.

[19] M.–D. GLESSGEN, *La genèse d'une norme en français au Moyen Âge : mythe et réalité du 'francien'*, cit.

- La notation ‹uo› est uniquement attestée dans le corpus de Fribourg et dénote ainsi un particularisme fribourgeois qui mérite d'être analysé en profondeur[20].
- Sans entrer dans le détail, notons que les graphèmes ‹ueu›, ‹ou›, ‹iou›, ‹uou› et ‹u› qui sont très rares sont surtout attestés dans les corpus oïliques sud–orientaux et dans les corpus francoprovençaux.

Avec ce cas précis, nous pouvons voir l'intérêt que revêt ce type d'analyse dans la discussion sur la variation interne des *scriptae* du domaine galloroman, tout en mettant en avant l'influence toujours plus croissante de la *scripta* exemplaire. Notre étude de cas a aussi démontré l'importance de s'attarder non pas sur une seule région, mais bien sur toutes les régions de la *Galloromania*. Le fait de croiser les issues de ces étymons en domaine oïlique et en domaine francoprovençal nous a permis de révéler le *continuum* linguistique que représente le réseau scriptologique galloroman, mais surtout de révéler les liens étroits entre le domaine francoprovençal et le domaine oïlique sud–oriental. Nous verrons que l'exploitation des DocLing sous l'angle de la lexicologie fournit une preuve supplémentaire à cette observation.

## 4. L'exploitation en lexicologie : la régionalité

Ainsi que nous l'avons rapidement mentionné, tous les corpus oïliques disponibles dans les DocLing ont fait l'objet d'une lemmatisation[21]. Ceci, permet d'accéder de façon immédiate à la totalité du lexique transmis par les chartes en langue d'oïl. Or, notre étude entend exploiter ces données à la lumière de la régionalité lexicale[22].

Avant tout, dans le cadre de cette contribution, il est nécessaire de définir ce que nous entendons par régionalité : « un régionalisme est un

[20] Cette analyse est présentée dans le cadre de la thèse de doctorat de Lorraine Fuhrer, *Les scriptae de l'espace francoprovençal au Moyen Âge (XIII^e–XV^e siècles) : l'élaboration d'une norme de chancellerie à Fribourg (Suisse)*, thèse de doctorat, Université de Strasbourg.

[21] Notons par ailleurs, que le vocabulaire des DocLing est désormais intégré dans le DEAFpré et donc accessible en ligne.

[22] Ce travail s'inscrit dans le cadre de la thèse de doctorat d'Alessandra Bossone, *Le Dictionnaire des régionalismes du français médiéval (DRFM). Principes méthodologiques, résultat et perspectives*, soutenue à l'EPHE en cotutelle avec l'université de Neuchâtel. Cette thèse s'inscrit dans un projet plus vaste intitulé "La scriptologie lexicale : le cas du domaine oïlique oriental" qui a vu la participation de plusieurs chercheurs : nous–même (Alessandra Bossone) en tant que doctorante, Marco Robecchi en tant que PostDoc, Martin Glessgen et Hélène Carles comme directeurs.

lexème dont la forme et/ou le sens se caractérise par une diffusion régionale identifiable à l'intérieur de l'espace de la langue en question […]. Il s'agit donc de lexèmes qui sont utilisés exclusivement ou de manière préférentielle dans une seule ou dans plusieurs régions données »[23].

Ainsi, nous considérons comme régionaux les mots qui, sur la base de leur diffusion géographique, illustrent une concentration d'attestations dans une aire définie en combinaison avec l'absence d'attestations dans le reste du territoire. Ajoutons qu'au–delà du critère géographique, nous nous basons, pour notre identification, également sur des paramètres étymologiques et sémantiques dans une optique historique allant de la première occurrence du mot jusqu'aux attestations modernes[24]. Citons à ce propos Henry qui avait déjà esquissé cette méthode :

> Elle [la méthode] est d'ailleurs plutôt une convergence de méthodes, caractérisée peut–être […] par une philologie aussi « abondante » et aussi rigoureuse que possible. Quoi qu'il en soit, il n'est pas inutile d'en rappeler les grandes lignes : études critique des témoignages anciens, à savoir : datation et localisation, au moins régionale, des textes ou des copies ; études phonétique (valeur des rimes) et sémantique des mots en questions ; détermination approximative de l'aire ancienne, sans oublier les enseignement de la géographie linguistique ; l'étude philologique doit absolument tenir compte de l'apparat critique et des caractéristiques régionales des divers manuscrits […][25].

Cela étant dit, notre analyse de la régionalité se place dans une perspective régiolectale ciblée sur l'aire orientale du domaine d'oïl, comprenant les régions suivantes : la Lorraine, la Champagne, la Bourgogne et la Franche–Comté qui sont amplement représentées dans le corpus des DocLing. Dans cette étude, nous nous limitons à la variance diatopique des lexèmes et non aux variantes graphiques qui seront certes signalées mais non analysées de façon primaire[26] ; cet aspect étant réservé à l'exploitation graphématique.

[23] M.–D. GLESSGEN, *La régionalité lexicale au Moyen Âge : hypothèses, constats et perspectives*, in : M.–D. GLESSGEN/D. TROTTER, (ed.), *La régionalité lexicale du français au Moyen Âge*, ELiPhi, Strasbourg 2016, p. 3.

[24] J.–P. CHAUVEAU, *Régionalismes médiévaux et dialectalismes contemporains en Haute–Bretagne*, in : M.–D. GLESSGEN/D. TROTTER, (éds.), *La régionalité lexicale du français au Moyen Âge*, ELiPhi, Strasbourg 2016, pp. 131–166; G. ROQUES, *Aspects régionaux du vocabulaire de l'ancien français*, thèse de doctorat, Strasbourg, 1980.

[25] A. HENRY, *Études de lexicologie française et gallo–romane*, PUF, Paris 1960, p. 5.

[26] Dans cette optique nous nous plaçons en continuité avec la voie ouverte par Y. GREUB, *Les mots régionaux dans les farces françaises*, SLiR, Strasbourg 2003.

C'est sur la base de la documentation lemmatisée des DocLing que nous avons initié le travail d'identification de la nomenclature régionale[27]. Ce travail s'est organisé en plusieurs phases. Dans un premier temps, nous avons entrepris une analyse lettre par lettre de tous les mots lemmatisés dans la base des DocLing, soit un total de 4 600 lemmes[28], en prenant en compte aussi les corpus qui concernaient les régions qui ne sont pas immédiatement concernées par notre étude (à savoir : le corpus de Douai et des chartes royales). Dans un second temps, grâce à la consultation des articles des dictionnaires de référence pour l'ancienne langue (FEW[29], Gdf[30], TL[31], DEAF[32], DMF[33] et DFM[34])[35], nous avons identifié une série de lexèmes régionaux qui ont été traités par la suite au moyen d'articles lexicographiques. Ceci a donné lieu au *Dictionnaire des régionalismes du français médiéval* (DRFM) qui compte à l'heure actuelle *ca* 400 articles lexicographiques[36].

Pour illustrer les possibilités d'exploitations lexicales offertes par le corpus, nous aborderons le cas du régionalisme *fossorer*, pour lequel nous avons rédigé l'article qui paraîtra dans le DRFM. Ce lexème, attesté depuis le 13e siècle, est un terme d'agriculture qui désigne l'action de labourer la terre avec une houe. Dans l'article — dont nous reproduisons ci–dessous le début — nous avons répertorié, par ordre

[27] Cette phase du projet, a été réalisée par nous–même (Alessandra Bossone) et par une autre doctorante : Mathilde Cornu.

[28] Rappelons que la base supporte les expressions régulières, par ex. '^a' permet de sélectionner tous les mots commençant par 'a–'.

[29] W. VON WARTBURG et al., *Französisches Etymologisches Wörterbuch. Eine Darstellung des galloromanischen Sprachschatzes*, 25 vol., Leipzig / Bonn / Bâle 1922.

[30] F. GODEFROY, *Dictionnaire de l'ancienne langue française et de tous ses dialectes*, 10 vol., Paris 1880–1902.

[31] A. TOBLER/E. LOMMATZSCH et al., *Altfranzösisches Wörterbuch*, 11 vol., Berlin 1925–2002.

[32] K. BALDINGER/F. MÖHREN/T. STÄDTLER (dir.), *Dictionnaire étymologique de l'ancien français*, Tübingen, 1975–.

[33] R. MARTIN (dir.), *Dictionnaire du moyen français*, publication en ligne, 2015 (http://www.atilf.fr/dmf/).

[34] T. MATSUMURA, *Dictionnaire du français médiéval*, Les Belles Lettres, Paris 2015.

[35] Nous avons également intégré les matériaux vernaculaires en contexte latin enregistrés dans les dictionnaires tels que Niermeyer et Du Cange, ainsi que ceux de la base de données des *Chartae Galliae* (AA.VV., 2014. publication en ligne http://www.cn–telma.fr//chartae–galliae/index/).

[36] Ce dictionnaire est actuellement en cours de réalisation sous la direction d'Hélène Carles et Martin Glessgen avec la collaboration d'Alessandra Bossone et Marco Robecchi en tant que rédacteurs. Nous profitons de cette communication pour remercier Gilles Roques qui a accepté de relire la totalité des articles rédigés et a apporté des précieux commentaires.

chronologique, les occurrences anciennes du lexème enregistrées dans la lexicographie de référence[37] :

bourg., frcomt. **FOSSORER** v.

bourg. *fessorer* (13e s., 1390/92), *fessourer* 1450
frcomt. *fous(s)erer* 1286, *fossorer* (1290, 1313)

cf. bourb. *fessourer* 1498

- ♦ (t. d'agriculture) "labourer la terre avec une houe (ici : *vigne*, *cortil*, *terre*)"

1286 (AD Doubs, quittance de la chambre des comptes de Dole) : « Pour fouserer la plantee (…) pour fouserer les vignes monseignour (…) pour fousserer dou premier cop » (= Gdf)
1290 (JPrioratR 5985 [frcomt.]) : « La terre fossorent et muevent » (= DFM)
13e s. (CoutChartreux f° 11 r° [Dijon]) : « Li curtilliers fessore le cortil, il seme les herbes et les sacle » (= Gdf)
1313 (ComptMahautArtois [Franche–Comté]) : « Pour fossorer toutes ces vignes premiere fois (…) pour fossorer ycelles vignes seconde fois (…) pour fossorer ycelles vignes tierce fois (…) » (= Gdf)
1390/92 (AD COr B 4786, inventaire) : « Pour les journees de 647 ouvriers de bras qui ont fessoré le dit cloux la premiere foiz » (= Gdf)
1450 (ProstInv 1, 455 [Bourgogne]) : « (…) fessourer 3 quartiers de terre qui sont ou jardin » (= DMF)
1450 (ArchNord B 1684, f° 138 [mentionnant un nom de lieu en Bourgogne]) : « Icelui povre suppliant fust en nostre ville de Dijon avec pluseurs autres laboureurs qui estoient illec venuz pour gaignier et fessourer as vignes » (= DMF)
1498 (ChartBourbonnais 409) : « (…) pour labourer et fessourer ladicte vigne » (= DMF)

*Données supplémentaires* : par ailleurs, en domaine francoprovençale les DocLing enregistrent : frb. *fossorar* (1415, 1420), *fossoreir* (1415) et *fosserar* (1411, 3 occ.). On relève *fossorar* aussi dans Devaux (Dauphiné septentrionale 1401/02 ; = FEW).

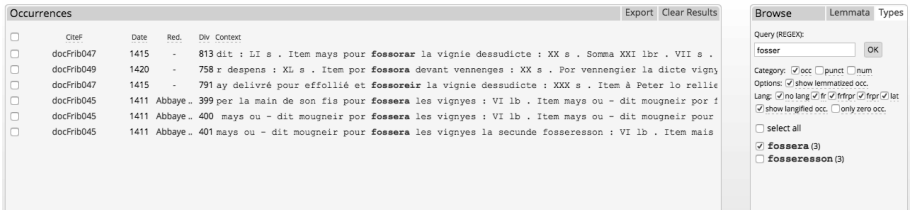
FEW 3, 742b s.v. FOSSORIUM : « abourg., afrcomt. *fossorer* "labourer avec une houe" (depuis 13e s.) », cf. aussi : « afrb. *fosora* ARom 7, 331, adauph. *fossorar* »

[37] Signalons que ce mot n'est pas présent dans les documents oïliques des DocLing mais seulement dans les documents du domaine francoprovençal. Les attestations francoprovençales n'ont pas été intégrées dans le corps de l'article mais dans la section 'données supplémentaires' en vue d'être analysées selon une approche contrastive.

Gdf 4, 106 s.v. *fossorer* v. "labourer avec une houe" (frcomt. 1286, 1550, bourg. 13e s.–1604)
TL 3, 2174 s.v. *fossorer* [renvois au Gdf]
Ø DEAFpré
DMF s.v. *fessorer* v. (bourg. 1375, 1450, bourb. 1498)
DFM s.v. *fossorer* (entré marquée 'sud–est' : frcomt. 1290)
cf. aussi GPSR 7, 737b s.v. *fosserer*, *fòsərā* "labourer avec un outil à main là où on ne peut utiliser la charrue" (afrpr. *fossera*, *–er*, *–orar*, etc. depuis 1411/12, F, Vd, N, B)
v. aussi Vitali[38] s.v. *fossorare* (dep. 1227 en Suisse romande) : « vocables typiques de l'est de la France et de Suisse romande ».

Ainsi que nous pouvons le voir grâce aux étiquettes géolinguistiques qui précèdent l'entrée de l'article, il s'agit d'un verbe propre aux régions de la Bourgogne et de la Franche–Comté, avec une attestation plus tardive dans le Bourbonnais (1498). Si l'on s'attarde sur les dialectes modernes, nous constatons que ce verbe survit avec des sens proches dans les parlers de la Bourgogne et de la Franche–Comté (v. FEW 3, 742b s.v. FOSSORIUM : Beaune, Poligny, Besançon, Vaudioux, etc.).

Par ailleurs, et c'est là l'aspect qui nous intéresse, le FEW mentionne une attestation de *fossorar* dans le Dauphiné (1401/02) qui reflète l'existence de ce verbe en domaine francoprovençal. Ceci est confirmé par les données supplémentaires réunies grâce aux DocLing où l'on relève ce verbe à Fribourg (1411–1420) sous les formes francoprovençales *fossorar* et *fosserar* et sous la forme d'importation oïlique orientale *fossoreir*, dont voici le détail des occurrences extrait de Phoenix2 :




D'un point de vue étymologique, c'est sur la base des cognats galloromans (bourg. *fesso(u)rer*, frcomt. *fossorer*, *fousserer*, frb. *fossorar*, *fosserar*, *fossoreir*, dauph. *fossorar*) que l'on peut raisonnablement supposer qu'il s'agit d'un continuateur de la base protoromane *FOSSORARE qui a été formée sur le latin tardif FOSSORIU "instrument pour

[38] D. VITALI, *Mit dem Latein am Ende ?*, Peter Lang, Bern 2007, p. 475.

creuser" (> afr. *fossoir* s.m. "houe")[39]. Par ailleurs, les DocLing et la lexicographie attestent aussi un dérivé d'agent en *–ier* (< –'ARIU), *fossorier* s.m., avec l'acception "celui qui laboure la terre, laboureur (surtout dans les vignobles)". Celui–ci montre une distribution géolinguistique partiellement semblable à celle du verbe *fossorer*[40].

Ces deux lexèmes remontant à la même famille étymologique illustrent le fait qu'il existe un lexique régional conjoint aux domaines oïlique sud–oriental et francoprovençal. Nos analyses nous ont démontré que ce cas de figure est plutôt fréquent dans la langue des textes documentaires que nous étudions[41]. Or, cette convergence peut s'expliquer à la lumière de la tendance à employer des lexèmes à large diffusion propre à toute langue écrite. Elle suit un *continuum* géographique au–delà des frontières linguistiques, qui — comme nous l'avons vu pour l'analyse graphématique — est perceptible grâce à la prise en compte d'une documentation suffisamment étendue permettant une approche contrastive essentielle[42].

## 5. Conclusion

Au premier abord, nos recherches portent sur des sujets différents et répondent à des méthodologies divergentes. Toutefois, elles offrent l'opportunité d'aborder des questions importantes pour l'histoire linguistique de la *Galloromania*.

Nous avons surtout axé notre communication sur les liens qui unissent le domaine francoprovençal et le Sud–Est du domaine oïlique.

[39] Cf. L. GAUCHAT/J. JEANJAQUET/E. TAPPOLET (dir.), *Glossaire des patois de la Suisse romande*, Neuchâtel/Paris 1924– (https://www.unine.ch/gpsr), vol 7, p. 737b s.v. *fossorer* pour la question étymologique. D'après le GPSR une dérivation romane sur le substantif *fossoir* s.m. "sorte de houe pour labourer la vigne" (< FOSSORIU/*FOSSOTORIU) s'exclut, vu que le verbe est bien attesté dans des cantons suisses où le nom a peu de vitalité.

[40] Le régionalisme est connu par trois attestations anciennes (1209–1338) localisées en Franche–Comté et Bourgogne. Il est également attesté à date ancienne en Suisse romande (cf. DocLing Neuchâtel et Vaud/Genève), dans l'Isère et l'Ain. Il survit dans les dialectes francoprovençaux (Isère, Ain, Valais et Vaud).

[41] Par ex. les régionalismes suivants qui seront traités dans le dictionnaire : la famille d'afr. rég. *afoër* v. "fournir le bois nécessaire au chauffage", *albue* s.f. "terre glaise de couleur grise ou rougeâtre, argileuse", *apaiser* v. "protéger, défendre la possession de quelque chose" et *beveuge* s.m. "lieu où on loge les bestiaux, étable".

[42] Cf. aussi l'article H. CARLES/M.–D. GLESSGEN, *L'élaboration scripturale du francoprovençal au Moyen Âge*, « ZrP » 135 (2019), pp. 68–157 où est soulignée l'importance d'une approche contrastive pour l'étude des langues écrites médiévales.

Bien que ces observations ne soient pas révolutionnaires[43], notre but était ici de démontrer qu'il est possible d'appréhender cette caractéristique sur la base du corpus sur lequel nous travaillons. Ces deux zones reflètent un *continuum* linguistique qui s'observe par la présence de diatopismes graphiques et lexicaux communs, ce qui atteste l'absence d'une rupture nette entre elles.

Bien qu'ici, nous avons concentré notre analyse sur un seul lexème, nous supposons qu'une étude lexicale systématique permettrait d'observer la présence d'un lexique régional partagé entre la zone oïlique sud–orientale et le domaine francoprovençal (de façon partielle ou totale) parfois d'origine protoromane. Cette observation est d'ailleurs confirmée par les analyses lexicologiques réalisées dans le cadre du DRFM. Dans d'autres cas, le DRFM permet aussi l'identification d'emprunts au français dans les documents en *scripta* francoprovençale, ainsi que — plus rarement — de lexèmes d'origine francoprovençale empruntés dans les *scriptae* oïliques.

L'analyse graphématique permet, quant à elle, de mieux cerner le processus d'élaboration scripturale au sein d'un domaine étendu. La *Galloromania* représente un réseau scriptologique où de nombreux modèles se sont développés et où certains d'entre eux se sont répandus, notamment à partir de centres d'écriture importants[44]. Parmi ceux–ci se trouve la chancellerie royale qui a élaboré une *scripta* qui a ensuite eu une incidence dans la totalité du domaine galloroman. L'extension de cette *scripta* exemplaire de Paris qui s'est faite à partir de la première moitié du 14e s.[45] se constate de manière flagrante dans les corpus plus tardifs du domaine francoprovençal.

[43] À ce sujet voir *Ibidem* et Y. GREUB, *La diffusion de la langue littéraire au Moyen Age : français et francoprovençal*, « Langages » 215 (2019), pp. 15–26.

[44] Concernant l'importance des centres d'écritures, nous renvoyons aux publications de M.–D. Glessgen, *Les lieux d'écritures dans les chartes lorraines du XIIIe siècle*, cit. et de P. Videsott, *Les plus anciens documents en français de la chancellerie royale capétienne (1241–1300)*, cit.

[45] M.–D. GLESSGEN, *La genèse d'une norme en français au Moyen Âge*, cit.