



**HAL**  
open science

## **Perspectives d'analyse computationnelle de la langue du ms. Turin, BNU, L.II.14**

Jean-Baptiste Camps, Marco Robecchi

### ► **To cite this version:**

Jean-Baptiste Camps, Marco Robecchi. Perspectives d'analyse computationnelle de la langue du ms. Turin, BNU, L.II.14. Studi francesi, A paraître, 206 (LXIX | II), pp.310-321. <10.4000/15gi3>. <hal-05184125>

**HAL Id: hal-05184125**

**<https://enc.hal.science/hal-05184125v1>**

Submitted on 24 Jul 2025

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

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



HAL Authorization

## **Perspectives d'analyse computationnelle de la langue du ms. Turin, BNU, L.II.14**

### *Abstract*

The aim of this article is to analyse the language of the texts in the first volume of Manuscript Turin, BU, L.II.14. Using transcriptions obtained as part of the *Rescapé* project, we will subject the texts to stylometric, scriptometric, and scriptological analysis. The objective is to determine whether the *unica* were written by a single author-copist or if the texts originate from different sources and have been arranged in a way that establishes connections between the texts copied within the same manuscript.

### *Mots-clés*

Linguistics analysis; stylometry; scriptometry; scriptology; Digital Humanities

Le manuscrit Turin, Biblioteca Nazionale Universitaria, L.II.14 (*T*) fait partie du lot de manuscrits touchés par l'incendie de la bibliothèque en 1904<sup>1</sup>. Probablement conçu comme une compilation unitaire par un remanieur bien expérimenté, et composé à l'origine d'environ 600 feuillets, ce recueil « colossal [est] de grand format, richement enluminé et transcrit avec soin par différents scribes »<sup>2</sup>. Les quatre volumes qui composent le recueil original contiennent au total 22 «œuvres», bien que la séparation entre certaines unités textuelles fasse l'objet de discussions<sup>3</sup>. À côté de pièces transmises par plusieurs témoins manuscrits, comme la *Bible* d'Herman de Valenciennes, *Hervis de Mes*, le *Roman de Saint Fanuel*..., on y retrouve 13 *unica* – mais ce décompte est également discutable – transmis par le seul manuscrit de Turin.

Cette contribution vise à analyser la langue des textes contenus dans le premier tome du recueil, en essayant de déceler quelle cohérence linguistique se dessine entre les *unica* par rapport aux textes à tradition. Cette question n'est pas anodine, si l'on se tient entre autres au constat de Giannini, qui affirme (nous soulignons) : « Ce projet audacieux est l'œuvre d'un clerc expérimenté qui n'hésite pas à remanier ou à écrire lui-même les textes dont il a besoin et qui ne sont conservés, à présent, que dans notre recueil »<sup>4</sup>. Pour y apporter une réponse, nous avons choisi de tester les méthodes stylométrique, scriptométrique et scriptologique à l'aide de l'informatique. Certains des *unica* posent des difficultés très fortes à toutes ces approches, à

---

1 Sur l'épisode, cf. entre autres la notice de P. Meyer dans les "Comptes rendus des séances de l'Académie des Inscriptions et Belles-Lettres" 48, 1904, pp. 31-32, sa *Chronique*, "Romania" 33, 1904, pp. 306-320 e celle de G. Bourguin, *L'incendie de la bibliothèque nationale et universitaire de Turin*, "Bibliothèque de l'École des Chartes" 65, 1904, pp. 132-140.

2 G. Giannini, *Poser les fondements: lieu, date et contexte (essai sur le recueil L.II.14 de Turin)*, "Études françaises" (n. spécial "Lire en contexte: enquête sur les manuscrits de fabliaux") 48, 3, 2012, pp. 11-31, p. 12.

3 Sur cet aspect, nous renvoyons aux contributions de Marco Maulu et d'Ilaria Molteni dans ce volume.

4 Giannini, *Poser les fondements* cit., p. 13. À propos des *unica* de *T*, parmi lesquels on trouve la *Chanson de Godin*, l'éditeur Meunier affirme: «La partie rimée et la partie assonancée ne sont pas du même auteur ... nous nous trouvons donc en présence de deux œuvres différentes» (*La chanson de Godin*, éd. F. Meunier, Louvain, Bibliothèque de l'Université, 1958, p. LXIX). Toujours d'après Meunier, la deuxième partie assonancée serait de provenance picarde, sensiblement influencée par le «francien» et plus ancienne que la partie rimée.

cause de leur brièveté ou, notamment au début du vol. I, en raison de l'état de conservation du manuscrit.

La *scripta* littéraire de *T* sera mise en comparaison – dans la mesure du possible – avec la *scripta* documentaire du *corpus* des *Documents linguistiques galloromans*<sup>5</sup> et des *Atlas* linguistiques de Dees<sup>6</sup>. Le but est d'essayer de s'affranchir de la circularité qui affecterait une étude linguistique basée sur des textes appartenant à un seul genre, bien conscients du niveau de caractérisation scriptologique appartenant à chaque typologie de texte médiéval. On aura ici la possibilité de tester la portée effective et l'interopérabilité de ces deux typologies scripturaires à l'aide des nouvelles méthodes informatiques.

### 1. *Méthodologie : constitution du corpus*

Notre analyse porte plus particulièrement sur les neuf premiers textes du recueil, pour lesquels nous disposons de transcriptions effectuées dans le cadre du projet *Rescapé* (nous indiquons le numéro de tokens, y compris les noms de personne, en excluant la ponctuation et les chiffres)<sup>7</sup>:

1. HermVal: Herman de Valenciennes, *Bible*, extraits, f. 1ra-1vb, f. 12ra-24vb; non transcrit 47rb-49rb; pic. 4<sup>e</sup> quart du XIII<sup>e</sup> s.; environ 37 mss.; transcription normalisée sauf les ff. 24ra-24vb, 6436 tokens<sup>8</sup>;
2. GenealTur: *Généalogie*, f. 10ra-va, fin du XIII<sup>e</sup> s., dont l'auteur pourrait être le copiste lui-même, cf. infra ; 740 tokens<sup>9</sup>;
3. NoironsTur: *Noirons li arabis*, ff. 10va-b, ff. 2ra-12ra, pic./hain., fin du XIII<sup>e</sup> s.?; 2439 tokens<sup>10</sup>;

---

5 Le projet des *Documents linguistiques galloromans* (dorénavant DocLing) commence à l'École nationale des chartes, au XIX<sup>e</sup> siècle, sous l'impulsion de Paul Meyer et Clovis Brunel ensuite, se stabilise grâce à Jacques Monfrin à partir de 1974, et a enfin été repris par Martin Glessgen en 1998, qui a produit une version informatique actuellement stockée dans la base de données *Documents et analyses de la Galloromania médiévale* (désormais GallRom, <https://gallrom.linguistik.uzh.ch/#/>) de l'Université de Zurich, auquel Marco Robecchi collabore depuis plusieurs années. Le projet *RESCAPÉ*, quant à lui, est né de l'initiative conjointe de Marco Maulu (Université de Sassari) et de Jean-Baptiste Camps (École nationale des chartes, PSL), et a été financé par Biblissima+ en 2021. L'École nationale des chartes représentant le point d'intersection de ces deux projets, il nous a paru convenable de les faire interagir dans le présent article.

6 A. Dees, *Atlas des formes et des constructions des chartes françaises du 13<sup>e</sup> siècle*, Tübingen, Niemeyer, 1980, «Beihefte zur Zeitschrift für romanische Philologie» 178 (dorénavant DeesAtlas); A. Dees, *Atlas des formes linguistiques des textes littéraires de l'ancien français*, Tübingen, Niemeyer, 1987 «Beihefte zur Zeitschrift für romanische Philologie» 212, pp. XIX-XX (dorénavant DeesAtlas2).

7 Les sigles des éditions et des textes, les informations concernant la datation et la localisation des textes sont tirées de la bibliographie en ligne du *Dictionnaire Étymologique de l'Ancien Français* (dorénavant DEAF, <https://alma.hadw-bw.de/deafbibl/>), désormais intégrée à la base de données GallRom.

8 Nous avons également eu recours à l'édition *Li romanz de Dieu et de sa Mere d'Herman de Valenciennes*, éd. I. Spiele, Leyde, Presse Universitaire, 1975.

9 Édité dans *Hervis de Mes*, éd. J.-Ch. Herbin, Genève, Droz, 1992, pp. 520-525, qui édite l'intégralité du texte *Hervis de Mes* (cf. infra, n. 9 de notre liste de textes).

10 Le texte, jusqu'à présent inédit, a récemment été publié par M. Maulu, *Virgile et Néron dans Noirons li arabis: Un unicum du manuscrit L.II.14 de la Bibliothèque nationale universitaire de Turin, "Trojanalexandrina"* (n. spécial « Regards médiévaux sur l'écrivain antique », 23, 2023 (sous presse). Le même auteur prépare une édition de la tirade de Virgile sur la Création dans le cadre du projet Prin Pnrr 2022 *All flames (AFAF): from Philology to artificial Intelligence and Back. Recovering the MS L.II.14 of the "Biblioteca Nazionale Universitaria di Torino" damaged by the Fire of 1904.*

4. SFanuel: *Roman de saint Fanuel*, ff. 25ra-70ra, ca 1200; 17 mss.; transcription diplomatique, 8502 tokens<sup>11</sup>;
5. SVou: *Miracle du Saint Vou de Lucques*, ff. 79rb-82ra, hain., fin XIII<sup>e</sup> s., environ 520 vv.; éd. SVouF, 3690 tokens<sup>12</sup>;
6. MahomTur: *Vie de Mahomet* en décasyllabes, ff. 82ra-83ra, environ 149 vv.; 1423 tokens<sup>13</sup>;
7. VenjNS: *Venjanse nostre Seigneur*, ff. 83rb-102vb, ca 1200?; 10 mss.; transcription normalisée jusqu’au f. 88va, diplomatique à partir du f. 89r, 29897 tokens<sup>14</sup>;
8. ProlLoh: Prologue au cycle des *Lohereins*, ff. 103r-104v, 304 vv.; 2192 tokens<sup>15</sup>;
9. Hervis: *Hervis de Mes*, ff. 105ra-170rb, lorr., 1<sup>er</sup> quart du XIII<sup>e</sup> s.; 5 mss.; environ 11500 vv.; 83122 tokens.

Sur ces 9 textes, 4 connaissent une tradition manuscrite assez importante (de 5 à 37 mss.); les 5 autres textes sont des *unica* du ms. de Turin, notamment GenealTur, NoironsTur, SVou, MahomTur, ProlLoh, dont l’insertion se justifie par l’agencement de l’histoire biblique à l’histoire des Lorrains. Quant au premier des *unica*, selon Jean-Charles Herbin l’auteur de la GenealTur était « vraisemblablement contemporain du copiste, sinon le copiste lui-même »<sup>16</sup>, tout en exprimant la même incertitude quant à l’auteur du ProlLoh<sup>17</sup>.

Entre SFanuel, qui raconte la généalogie de la Vierge Marie et de sa mère sainte Anne, et la VenjNS, qui raconte la destruction de Jérusalem par Titus, le responsable de la composition de notre manuscrit a intercalé la légende du SVou et une brève légende sur la vie de Mahomet dans une version par ailleurs inconnue. Pour le premier texte, qui serait localisé dans le Hainaut à la fin du XIII<sup>e</sup> siècle, la date de composition est très proche de la date de confection du manuscrit. Il convient de remarquer qu’entre ces deux textes il n’y aucune solution de continuité matérielle (aucune lettrine, aucune rubrique) qui justifie leur séparation en deux entités distinctes. Les deux textes sont introduits par un seul titre au f. 79rb, au début du SVou: «Chi faut li roumans d[e Nostre Dame] et la soufranch[e Jhesu Christ], si coumench[e s]a venganche». De plus, la laisse finale du Saint Volt et la première laisse du Mahomet se poursuivent sur la même assonance masculine en -é, ce qui ne rend pas le changement de sujet évident. Le passage d’un texte à l’autre ne serait signalé que par la formule introductive «Or vous dirai»:

Et ceste chose se vous ne m’en creés  
li cors en gist a Roume la cité  
en cler argent et en or esmeré.  
Or vous dirai se il vous vient en gré,  
em Babilone la deserte c[it]é,  
la fu Mahons li fis Esgarinés...

11 Nous avons également utilisé l’édition de C. Chabaneau, *Le Romanz de saint Fanuel et de sainte Anne et de Nostre Dame et de Nostre Segnor et de ses apostres*, “Revue des Langues Romanes” 28, 1885, pp. 118-123 et pp. 157-258.

12 Nous avons également utilisé l’édition de W. Foerster, *Le saint Vou de Luques*, “Romanische Forschungen” 23, 1907, pp. 1-56.

13 Le texte a été édité par éd. A. Graf, *Spigolature sulla leggenda di Maometto*, “Giornale Storico della letteratura Italiana” 14, 1889, pp. 204-211 et, plus récemment, par F. Ninitte, « La figure de Muhammad dans le ms. L.II.14 de Turin, entre prophète et *hons faés* », “Romania” 142, 2024, pp. 35-65.

14 Le texte de *T* a été partiellement édité par W. Suchier, *Ueber das altfranzösische Gedicht von der Zerstörung Jerusalems*, “Zeitschrift für romanische Philologie” 24, 1900, pp. 161-198 et 25, 1901, pp. 94-109.

15 Ce texte, comme GenealTur, a été édité dans Herbin, *Hervis* cit., pp. 525-535.

16 Voir *Ibidem*, p. 521.

17 *Ibidem*, p. 526.

On possède donc tous les éléments pour croire à une conception unitaire du texte qui est, par ailleurs, l'idée de Foerster: «j'ai encore ajouté l'épisode de Mahomet, qui est certainement du même auteur»<sup>18</sup>.

Comme cela a été souligné à plusieurs reprises, les analyses menées jusqu'à présent ont été principalement consacrées à un seul texte à la fois, sans considérer le contenu du manuscrit dans son ensemble, afin de dégager la cohérence de sa constitution. L'urgence d'un changement de perspective radical quant à cette approche obsolète a été le moteur de deux projets de recherche récents : *Rescapé* et *AFAF*, financés respectivement par Biblissima+ en France (2021) et par le Ministère de la Culture et de la Recherche en Italie (2022), Programme Prin PNRR 2022.

Nous connaissons l'origine et la datation du manuscrit de Turin avec une certaine précision : le manuscrit est daté de 1311 dans le colophon, transcrit au f. 583vb: «Cis livres fu escrits en l'an de l'incarnation MCCC et XI ou mois de joing»; quant à la localisation, Gabriele Giannini propose, avec des arguments bien fondés, que le manuscrit a été copié dans le Vermandois, peut-être autour d'Origny-Sainte-Benoite, dans l'actuel département de l'Aisne<sup>19</sup>. Notre enquête confirmera cette hypothèse, mais elle permettra surtout de proposer une méthode d'analyse de la langue d'un texte et d'un manuscrit-recueil médiéval à travers la modélisation informatique.

## 2. Constitution du corpus

Dans le cadre du projet *Rescapé*, les quatre volumes de *T* ont été soumis à une transcription par reconnaissance automatique des écritures (*Handwritten Text Recognition*, ou HTR) qui a produit un fichier texte brut en format xml-ALTO (*Analyzed Layout and Text Object*): il s'agit d'une transcription diplomatique pour l'instant dépourvue de ponctuation, diacritiques, accentuation, et qui n'offre qu'une segmentation des mots partielle. Les textes ont été ensuite révisés et transformés en version interprétative par Marco Maulu. Enfin, les transcriptions ont été tokenisées et lemmatisées à l'aide du modèle Deucalion de l'École nationale des chartes, sans que la lemmatisation ait encore fait l'objet d'une correction manuelle<sup>20</sup>.

Projeté sur la représentation schématique de la chaîne de traitement dessinée par Jean-Baptiste Camps, Lucence Ing et Elena Spadini<sup>21</sup>, nous avons ainsi obtenu, dans cette phase, un «texte enrichi» qui nous permet de passer à la phase de l'analyse statistique (Fig. 1).

---

18 Foerster, *Le saint Vou* cit., p. 30: «Ich habe die Mohammedepisode (M.) noch mitherangezogen, die sicher demselben Verf[asser] angehört».

19 Giannini, *Poser les fondements* cit., pp. 18-19.

20 Cf. J.-B. Camps, Th. Clérice, F. Duval et al., *Corpus and models for lemmatisation and POS-tagging of old French*, preprint, disponible en ligne: <http://arxiv.org/abs/2109.11442>.

21 J.-B. Camps, L. Ing, E. Spadini, *Collating medieval vernacular texts. Aligning witnesses, classifying variants*, in *DH2019 Digital Humanities Conference 2019, Jul 2019*, Utrecht, disponible en ligne: <https://hal.archives-ouvertes.fr/hal-02268348/>.

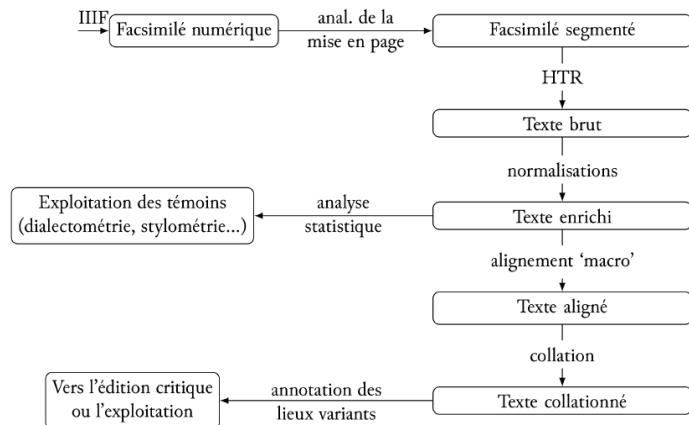


Fig. 1. Chaîne de traitement informatique des textes médiévaux

Une telle procédure de reconnaissance automatique est censée mettre à notre disposition, en peu de temps, une masse textuelle assez importante pour effectuer des analyses quantitatives fiables sur l'ensemble du manuscrit. Cette première phase du travail nous a ainsi permis de disposer d'un corpus textuel d'environ 138.440 tokens.

Grâce à un calcul très approximatif, au net des pertes, nous avons à disposition 580 feuillets recto/verso, soit 1160 pages de 2 colonnes chacune, chaque colonne étant composée de 45 lignes – soit 90 vers par page, pour un total de 104.400 vers environ, moins les décorations, les enluminures, les parties laissées en blanc et les parties illisibles à cause de l'incendie). Si l'on considère une moyenne de 7 mots par vers (mots grammaticaux et lexicaux confondus), nous disposons d'un corpus textuel d'environ 700.000 formes ou tokens. Il s'agit d'un corpus aux dimensions intéressantes pour une analyse quantitative<sup>22</sup>.

Pour étudier la question de l'attribution de ces *unica* et de leur éventuelle attribution à un même auteur, la méthode d'analyse quantitative la plus efficace est la stylométrie<sup>23</sup>. Toutefois, cette approche pose des difficultés particulières en raison de la brièveté des *unica* pris ici en considération et de l'absence d'auteurs candidats auxquels on pourrait comparer les textes à attribuer. Pour cette raison, l'analyse stylométrique (§ 3) sera redoublée d'une analyse scriptométrique (§ 4) visant à permettre de mieux profiler la langue de ces *unica*.

### 3. L'analyse stylométrique

Pour quantifier le style des textes et tenter d'attribuer ceux-ci à un auteur donné, la stylométrie se fonde sur des régularités et des motifs linguistiques d'ordre peu conscient et difficiles à contrefaire. C'est ainsi que, depuis le XIX<sup>e</sup> siècle, les mots-outils, et plus largement les morphèmes grammaticaux, ont formé un type de caractères parmi les plus fiables pour l'attribution des œuvres, y compris médiévales<sup>24</sup>. Pour ce qui est des textes médiévaux, toutefois,

22 Cf. Camps, Clérice, Duval, Ing, Kanaoka, Pinche, *Corpus and Models* cit., p. 4.

23 Sur la stylométrie de l'ancien français, voir notamment par J.-B. Camps, Th. Clérice, A. Pinche, *Noisy medieval data, from digitized manuscript to stylometric analysis: Evaluating Paul Meyer's hagiographic hypothesis*, "Digital Scholarship in the Humanities" 36, supplement 2, oct. 2021, pp. 49-71, disponible en ligne: <https://doi.org/10.1093/lle/fqab033>.

24 M. Kestemont, S. Moens, J. Deploige, *Collaborative authorship in the twelfth century: a stylometric study of Hildegard of Bingen and Guibert of Gembloux*, "Digital Scholarship in the Humanities" 30, 2, 2015, pp. 199-224,

une phase de lemmatisation est généralement nécessaire, pour éviter de prendre en compte les variantes graphiques liées aux scribes. Dans notre cas, nous avons soumis les textes de notre corpus à une lemmatisation, un calcul des fréquences relatives des lemmes de mots-outils, et une réplication de la méthodologie décrite par Florian Cafiero et Jean-Baptiste Camps<sup>25</sup>. La matrice de fréquence relative des lemmes a ainsi été normalisée (z-scores des fréquences pour chaque lemme, normalisation L2 pour chaque texte) et soumise à un algorithme de classification ascendante hiérarchique, avec comme critère de distance le delta de Burrows et le critère d'agglomération de Ward. Le dendrogramme résultant de cette analyse est présenté en Figure 2.

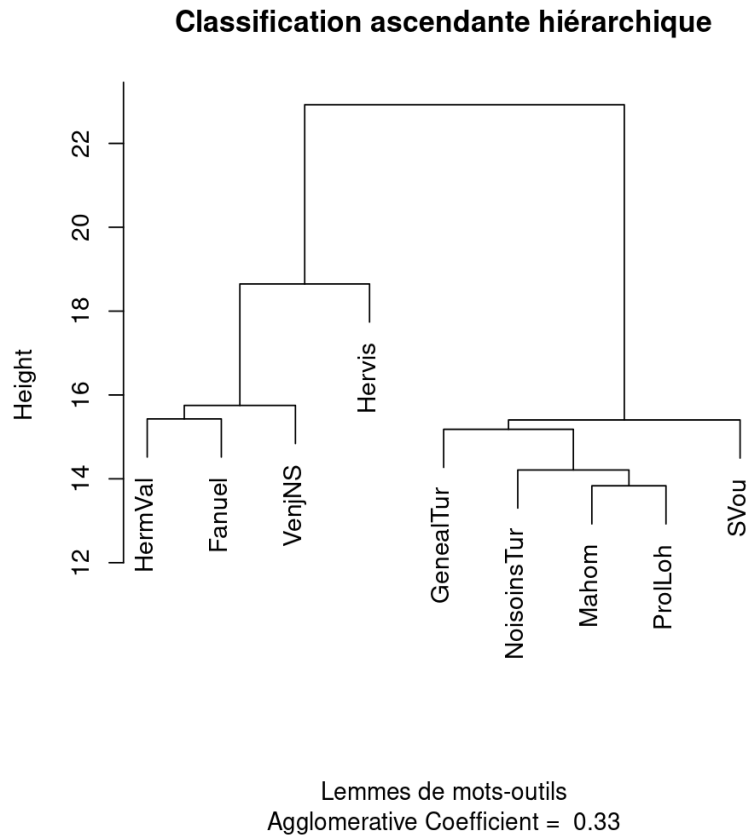


Fig. 2. Classification ascendante hiérarchique des textes du corpus, fondée sur les fréquences de lemmes de mots-outils.

L'analyse montre l'existence de trois groupes de textes: les textes sacrés de sujet religieux HermVal, Fanuel et VenjNS, d'une part; Hervis, texte épique isolé, et enfin le groupe des *unica* du manuscrit de Turin. Toutefois, celle-ci ne permet de conclure qu'à une plus grande proximité des *unica* entre eux plutôt qu'avec les autres textes du corpus, même si l'on note une affinité particulièrement forte entre MahomTur, NoironsTur et le ProlLoh. Cette proximité peut s'interpréter de plusieurs manières: on y retrouve probablement la trace des genres littéraires et traditions discursives, séparant romans à caractère religieux, chansons de geste et des pièces courtes plus hétéroclites. En outre, un examen des variables les plus corrélées, positivement ou

disponible en ligne: <https://lib.ugent.be/en/catalog/pug01:4099638>.

<sup>25</sup> Cf. F. Cafiero et J.-B. Camps, *Why Molière most likely did write his plays*, "Science Advance" 5, 11, 2019, disponible en ligne: <https://www.science.org/doi/10.1126/sciadv.aax5489>.

négativement, aux trois groupes, permet de mieux comprendre la classification (Table 1), et de constater que le groupe des *unica* est surtout défini négativement (et moins significativement) par la plus grande pauvreté du lexique de mots-outils employés, notamment liée à la grande brièveté de ces textes<sup>26</sup>.

Groupe 1 (HermVal, Fanuel, VenjNS)			Groupe 2 ( <i>Unica</i> )			Groupe 3 ( <i>Hervis</i> )		
	v-test	p-value		v-test	p-value		v-test	p-value
<i>gaire</i>	2.76	0.006	<i>gaire</i>	-2.41	0.016	<i>hu</i>	2.83	0.005
<i>mout</i>	2.63	0.008	<i>ez</i>	-2.49	0.013	<i>entretant</i>	2.83	0.005
<i>ambedos</i>	2.58	0.010	<i>desor</i>	-2.56	0.010	<i>aluec</i>	2.83	0.005
<i>chascun</i>	2.49	0.013	<i>mout</i>	-2.57	0.010	<i>entredeus</i>	2.83	0.005
<i>lez</i>	2.44	0.014	<i>suen</i>	-2.59	0.009	<i>ambes</i>	2.83	0.005
<i>suen</i>	2.42	0.015	<i>demain</i>	-2.75	0.005	<i>buer</i>	2.83	0.005

Table 1. Description de chaque groupe par les 6 variables obtenant le plus haut score, positif ou négatif, pour chacune d'entre elles.

#### 4. L'analyse scriptométrique

Vu le caractère peu conclusif de l'analyse stylométrique, nous avons procédé à une analyse scriptométrique (pour un texte médiéval, analyse "scriptométrique" nous paraît plus approprié que le "dialectométrique" du schéma ci-dessus, cf. Fig. 1). Ce procédé est «finalisé à la modélisation et à la classification des données de la scriptologie sur la base d'une grande quantité de données et selon les méthodes de la taxonomie numérique»<sup>27</sup>. Il s'agit donc d'une «approche quantitative et statistique, permettant une prise en compte globale et simultanée d'un nombre élevé de traits»<sup>28</sup>, qui permettrait de «découvrir, par le biais de méthodes quantitatives (mathématiques ou numériques), des structures de profondeur qui échappent à l'introspection directe du chercheur»<sup>29</sup>.

L'analyse scriptométrique a été appliquée par Paul Videsott dans sa *Padania scrittologica*, sur un corpus de 2064 documents médiévaux de l'Italie septentrionale jusqu'à 1525<sup>30</sup>. Pour les textes littéraires français, nous pouvons compter sur l'*Atlas* de Dees de 1987, suivi par les analyses de Goebel et Smečka: grâce à son analyse, Dees a pu proposer ou préciser la localisation d'environ deux cent textes médiévaux<sup>31</sup>. Récemment, la *Chanson d'Otinel* a été soumise à un traitement scriptométrique : les témoins de la chanson ont ainsi été intégrés à un corpus de témoins épiques

26 Une raison supplémentaire tient au fait qu'en termes de «tradition discursive» (*Diskurstradition*), *Hervis* est le seul à être rédigé dans un langage "épique", alors que les autres sont essentiellement des pièces "romanesques" à caractère religieux, dont la tradition discursive présente des caractéristiques propres qui s'écartent du style épique.

27 P. Videsott, *Padania scrittologica. Analisi scrittologica e scrittometriche di testi in italiano settentrionale antico dalle origini al 1525*, Tübingen, Niemeyer, 2009, «Beihefte zur Zeitschrift für romanische Philologie» 343, p. 10 (notre traduction).

28 J.-B. Camps, *La chanson d'Otinel. Édition complète du corpus manuscrit et prolégomènes à l'édition critique*, thèse Paris-Sorbonne 2016, p. CCCLXXIX, disponible en ligne: <https://shs.hal.science/tel-01664932>.

29 H. Goebel, *Regards dialectométriques sur les données de l'"Atlas linguistique de la France" (ALF): relations quantitatives et structures de profondeur*, "Estudis romànics" 25, 2003, pp. 59-120, p. 61; cf. également Videsott, *Padania scrittologica* cit., pp. 9-10.

30 *Ibidem*, p. 1.

31 Une version numérique de l'*Atlas* de Dees de 1987 a été préparée par une équipe de l'Université de Nice coordonnée par T. Scheer et disponible à l'adresse <http://atlasdees.unice.fr/wordpress/>. Cf. aussi H. Goebel et P. Smečka, *L'interprétation dialectométrique des atlas 'scripturaux' d'Anthonij Dees*, "Revue de Linguistique Romane" 80, 2016, pp. 321-368.

copiés dans différentes *scriptae*, et l'ensemble soumis à une analyse non-supervisée (classification ascendante hiérarchique) pour observer à la fois les principaux groupes de témoins (et les *scriptae* auxquels ils correspondent) ainsi que la place des témoins d'*Otinél* en leur sein<sup>32</sup>. Notre analyse scriptométrique vise à créer une routine de comparaisons avec un corpus contrastif pour vérifier la distribution de fréquence des occurrences des mots et placer les 9 textes enrichis de *T* à l'intérieur des traditions textuelles médiévales.

Pour ce faire, nous avons construit un corpus de référence pour les différentes *scriptae* médiévales, en ne retenant, au sein du *Nouveau corpus d'Amsterdam* (NCA), que les témoins pour lesquels la provenance du manuscrit était homogène avec celle du texte (pour limiter la stratification linguistique). Nous avons donc sur cette base constitué un corpus de 73 témoins, en retirant les textes apparaissant en doublon, pour éviter des attractions entre copies d'une même œuvre, qui seraient rapprochées non pas en raison de particularité de *scripta*, mais par leur partage de textes similaires (cf. Table annexe 1). De ce corpus, nous avons extrait des n-grammes de caractères (de longueur 2, 3 et 4) pour tenter de capturer les principaux marqueurs graphiques des différentes *scriptae* (en ne retenant que ceux qui, à l'échelle du corpus, avaient une fréquence totale supérieure à 5 occurrences), et nous avons entraîné des modèles d'apprentissage machine à reconnaître automatiquement la *scripta* d'un échantillon. Pour ce faire, nous avons utilisé des machines à vecteur de support linéaires (Linear SVC, avec une validation croisée de type «k-fold» de groupe, et une stratégie de gestion des déséquilibres des échantillons d'entraînement de type SMOTE et Tomek<sup>33</sup>).

Les scores d'entraînement ainsi que les traits les plus discriminants des modèles sont donnés en Annexe 2. Les scores sont globalement satisfaisants (exactitude de 82%), à l'exception du bourguignon (souvent confondu avec du champenois méridional ou du lorrain) et, dans une moindre mesure, du champenois septentrional (parfois confondu avec du picard<sup>34</sup>).

Une fois ces modèles entraînés, il est possible de les appliquer à une fenêtre roulante parcourant les textes de *T*, pour y identifier d'éventuelles variations de *scripta*. Les résultats sont présentés en Fig. 3. Celle-ci confirme la très nette prédominance de la *scripta* picarde, avec quelques légères fluctuations. En particulier, la *scripta* de SVou, MahomTur et ProlLoh semble plus marquée comme picarde encore, là où certains textes à traditions comme VenjNS ou Hervis présentent quelques variations (même si le picard y reste dominant). On notera en outre que NoironsTur et ProlLoh (peut-être aussi Fanuel) présentent des traits qui semblent faire monter la courbe du champenois septentrional : il est légitime de se demander s'il s'agit d'une trace de la présence de traits picards méridionaux plus marqués. Dans l'ensemble, les fluctuations demeurent malgré tout assez minces, quoiqu'elles ne contredisent pas l'hypothèse d'une proximité un peu plus grande entre MahomTur, NoironsTur et le ProlLoh (voire SVou), déjà observée sur le plan stylométrique.

---

32 Camps, *La chanson d'Otinél* cit.

33 Le code utilisé pour effectuer ces analyses est celui du logiciel SuperStyl: J.-B. Camps, *SUPERvised STYLometry (SuperStyl) (Version 0.9.0)*, 2021, à l'adresse <https://github.com/SupervisedStylometry/SuperStyl/>.

34 D'un point de vue plus largement géolinguistique, ces rapprochements faits automatiquement par la machine poussent à reconsidérer les frontières dialectales (ou du moins scriptologiques) parfois trop étanches et prédéterminées sur la base des limites actuelles qui affectent en partie la recherche en médiévistique.

### Analyse scriptométrique roulante

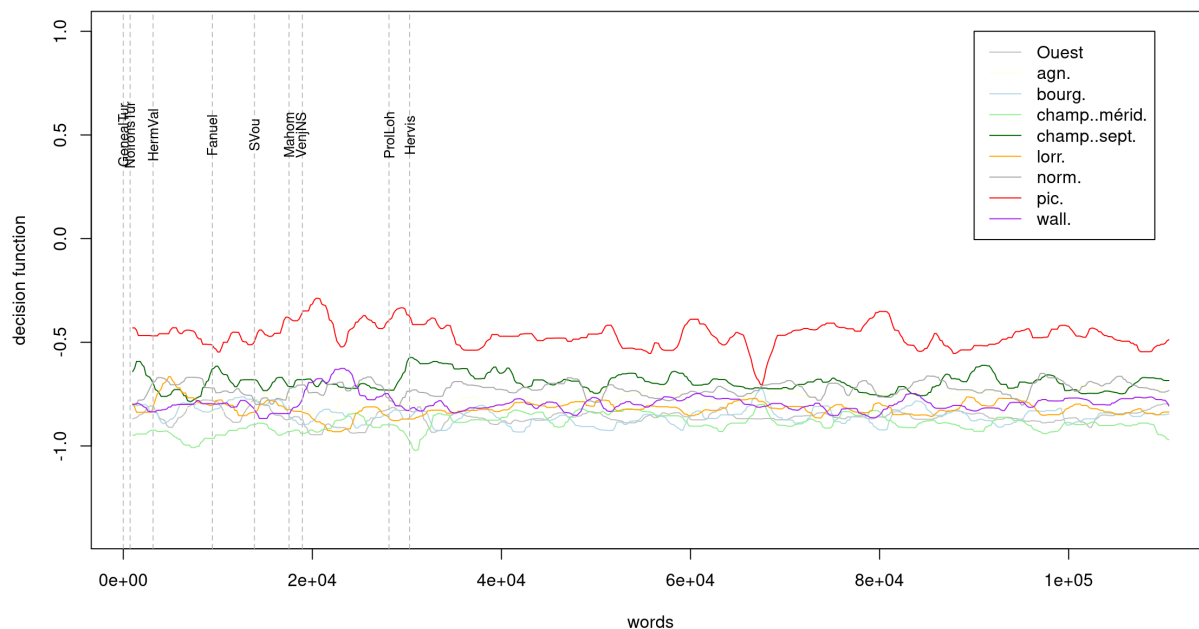


Fig. 3. Analyse scriptométrique roulante (fenêtre de 2000 mots, pas de 200 mots) sur la succession des œuvres étudiées du ms. de Turin.

### 5. Analyse scriptologique

L'analyse scriptologique se focalise en revanche sur les aspects "qualitatifs" de la *scripta*, en s'opposant à l'analyse scriptométrique de caractère "quantitatif" et en même temps en la complétant. En reprenant la définition de Paul Videsott, la scriptologie «s'occupe de la variation linguistique diasystémique des textes produits dans des conditions de normalisation et de standardisation linguistiques peu marquées»<sup>35</sup>; son but est de «présenter synoptiquement l'enchaînement des signes graphiques individuels et leur polymorphisme inter- et intra-textuel»<sup>36</sup>. Si, en général, la scriptologie s'est surtout intéressée à l'aspect graphématique des textes<sup>37</sup>, notre expérience dans la rédaction du *Dictionnaire des régionalismes médiévaux* (DRM)

35 Videsott, *Padania scrittologica* cit., p. 10 (notre traduction), qui reprend H. Völker, *Die Skriptaforschung als eine Philologie der Varietäten. Zur Negation mit (ne) ... nient in den altfranzösischen Urkunden der Grafen von Luxemburg (1237–1281)*, in: *Skripta, Schreiblandschaften und Standardisierungstendenzen. Urkundensprachen im Grenzbereich von Germania und Romania im 13. und 14. Jahrhundert. Beiträge zum Kolloquium (Trier, 16.-18.9.1998)*, dir. K. Gärtner, Trier, Kliomedica, 2001, pp. 75-104, à la p. 96.

36 H. Goebel, 79. *Methoden der Datensammlung und -verarbeitung / Traitement méthodique des données*, in *Lexikon der Romanistischen Linguistik*, dir. G. Holtus, M. Metzeltin, C. Schmitt, Tübingen, Niemeyer, 1988-2005, 8 voll., vol. I, t. 2, *Methodologie (Sprache in der Gesellschaft / Sprache und Klassifikation / Datensammlung und -verarbeitung)*, 2001, pp. 884-893, p. 888 (notre traduction).

37 Les aspects morphologiques en scriptologie rentrent essentiellement dans une approche de type graphématique : étudier les formes de la désinence *-ons*, par exemple (dans ses variantes *-on*, *-om*, *-um*...), se borne au niveau superficiel de la graphie; en revanche, les aspects morpho-syntaxique et syntaxiques se détachent du niveau graphématique (par exemple, présence de la déclinaison bicasuelle ou omission du pronom sujet).

nous a convaincu de l'existence d'une "scriptologie lexicale" qui peut s'appliquer également au niveau lexématique des textes<sup>38</sup>.

La scriptologie des textes médiévaux, qui a jusqu'à présent donné ses meilleurs résultats dans l'étude des textes documentaires, essentiellement transmis comme originaux, intègre désormais l'approche stratigraphique dans l'étude des textes non-documentaires (ou "textes à tradition"), transmis donc en copie par un ou plusieurs manuscrits. La stratigraphie est un aspect de plus en plus approfondi en philologie romane et permet de dégager les différentes couches qui se superposent dans la copie d'un texte<sup>39</sup>. Avec un recueil organique comme *T*, à partir d'une transcription intégrale de l'ensemble de ses textes à la conclusion du projet *AFAF*, il sera possible de mener une analyse stratigraphique verticale similaire à celle utilisée pour l'étude des chansonniers, pour répertorier les variations de traits scriptologiques à l'intérieur de chaque pièce dans l'ensemble de la copie. Si l'on dégager des variations significatives à l'intérieur des *unica*, on peut supposer qu'ils proviennent de sources autoriales différentes (quitte à y intégrer une enquête stylométrique, cf. *supra*). En même temps, une approche de type horizontal devrait permettre de vérifier la présence de certaines formes (graphématiques et lexicales) dans le reste de la tradition manuscrite d'un texte, pourvu que les éditeurs n'aient pas 'normalisé' la *scripta* de leurs textes, qu'ils aient fourni une description linguistique de chaque manuscrit assez fiable et qu'ils aient enregistré assez de variantes, ou du moins un choix raisonné des variantes, dans leurs apparats; dans le cas contraire, il conviendrait de disposer de la transcription de chaque témoin, soit elle intégrale (grâce à l'HTR et à sa fiabilité croissante le rêve n'est pas inatteignable) ou partielle (pour une analyse scriptologique fiable, un milliers de vers pourrait être suffisant). La Fig. 5 offre une représentation schématique des deux approches scriptologiques combinées.

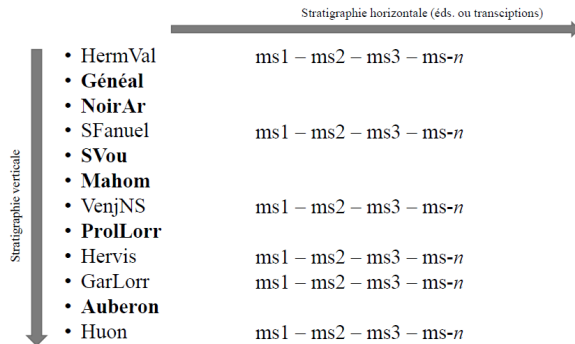


Fig. 5. Approches scriptologiques horizontale et verticale.

Pour éviter un risque de circularité trop élevé, nous avons décidé de recourir aux DocLing comme corpus de comparaison principal. L'expérience de travail sur les chartes nous alerte néanmoins sur la nécessité de considérer les aspects diastratiques et pragmatiques d'un texte, en plus du paramètre diatopique, en recourant au concept de "lieu d'écriture"<sup>40</sup>, qui constitue

38 Cf. H. Carles, M. Glessgen, M. Robecchi et al., *Dictionnaire des régionalismes médiévaux: la Galloromania orientale (DRM). Une analyse des documents linguistiques galloromans (XIII<sup>e</sup>-XV<sup>e</sup> s.)*, Strasbourg, ELiPhi (BiBliRo), 2024, notamment pp. 7-8; M. Robecchi, *La localisation de témoins oiliques médiévaux. Le DRFM comme levier méthodologique*, "Zeitschrift für romanische Philologie" 137, 3, 2021, pp. 744-796; Id., *Régionalité lexicale du français médiéval: analyse lexicologique et enjeux philologiques (Jean le Long, "Liber peregrinationis")*, "Medioevo Romanzo" 45, 2, 2021, pp. 284-320.

39 Sur ce point, cf. surtout Y. Greub, *La stratigraphie linguistique des manuscrits médiévaux et la variation linguistique*, in "Medioevo Romanzo" 42, 1, 2018, pp. 6-30, notamment p. 7.

40 Sur ces aspects, cf. M. Glessgen, *Les lieux d'écriture dans les chartes lorraines du XIII<sup>e</sup> siècle*, "Revue de

désormais la structure portante des DocLing. Le concept de lieu de rédaction, à la fois philologique, historico-culturel et sociolinguistique, permet de dépasser le lien contraignant qui, intuitivement, s'établit entre le lieu de rédaction (normalement un *scriptorium*, laïque ou ecclésiastique) d'un acte écrit et le lieu géographique d'où cet acte provient.

La nouvelle base GallRom, en plus des DocLing, comprend toutes les données des dictionnaires de l'ancien français (le TEAF), de l'ancien occitan (DAO) et du gascon (DAG), avec les bibliographies critiques relatives. Les données lexicales des Docling sont lemmatisées et mises en lien avec ces dictionnaires; les données scriptologiques font l'objet d'un balisage à part, appelé HCC (*Head Category Configuration*) qu'on a développé dans le cadre du projet NewScriptDesc (*Variation linguistique galloromane comme système*, unibz-uzh)<sup>41</sup>.

Le balisage grapho-phonétique que nous avons développé nous permet de dépasser la contrainte de la forme lexicale pour nous focaliser sur les séquences *n*-grammes isolés ou en fonction d'un mot donné<sup>42</sup>. Le projet d'alignement des lemmes sur la base d'un référentiel commun permettrait de faire le lien direct entre le lemmatiseur *Deucalion*, auquel nous avons soumis nos transcriptions, et les données du TEAF, lesquelles peuvent à leur tour être interrogées de manière plus approfondie et dynamique (cf. *infra*). Nous avons ensuite fait un choix de paramètres pondérés sur des formes suffisamment fréquentes dans tous les textes du corpus, qui font aussi l'objet de notre balisage scriptologique pour *NewScriptDesc* et qui avaient également été analysés pour la plupart par Dees dans ses *Atlas*. Toutes les fiches d'analyse détaillées se lisent dans l'Annexe 2, auquel nous renvoyons, nous bornant ici à résumer les données dans les conclusions (§6). Nous avons ainsi distingué entre des faits à «forte pertinence» (Annexe 2.1.), qui se sont révélés particulièrement efficaces pour une première phase d'analyse linguistique, des traits à «faible pertinence» (4.2.), qui ne sont pas distribués d'une manière assez homogène dans l'ensemble des textes examinés ici.

## 5. L'apport du lexique régional

Du point de vue lexical, nous avons jusqu'à présent pu repérer uniquement deux lexèmes du SVou qui se configurent comme nettement caractérisés du point de vue régional.

### 5.1. Le verbe "hoer"

On lit au f. 81ra:

---

Linguistique Romane" 72, 2008, pp. 413-540 et, plus récemment, P. Videsott, M. Robecchi, J. Schaber, (*Re*)cartographier la Galloromania médiévale: enjeux et perspectives quarante ans après l'"Atlas" de Dees, "Zeitschrift für romanische Philologie" 139, 4, 2023, pp. 1003-1047, notamment p. 1036.

41 Pour une description de la nouvelle base, cf. M. Glessgen, *Les documents et analyses linguistiques de la Galloromania médiévale (GallRom): structure et potentiel interprétatif*, in *Perspectives en linguistique et philologie romanes*, dir. D. Corbella, J. Dorta, R. Padrón, Strasbourg, ELiPhi, vol. 2, 2023, pp. 1025-1044; pour un premier essai de recherche scriptologique grâce à la nouvelle base GallRom, cf. M. Robecchi, *Scriptologie médiévale et dialectologie diachronique: le cas du é[ dans la Galloromania et en poitevin*, in Corbella, Dorta, Padrón *Perspectives en linguistique* cit., pp. 863-874. Une des innovation de relief concerne les données textuelles: les documents ne nécessiteront plus d'être conformes au schéma .xml assez particulier des DocLing, mais ils pourront désormais être téléchargés à partir d'un format .xml simple.

42 Cf. *Ibidem* pour plus de détails. Nous avons appelé ces formules de balisage HCC «*Head Category Configuration*», car elles renferment synthétiquement les paramètres pertinents pour une étude scriptologique: l'environnement phonétique, la trajectoire (héréditaire vs savant), le cas échéant les lexèmes individuels, les résultats graphématiques et phonétiques (hypothétiques). Nous disposons actuellement (avril 2024) de 1827 HCC pour 1.384.231 formes balisées.

Tant ont partout et quis et roirwardé  
 et en la terre et fouy et hoé,  
 la crois trueuent si ont le cuer leué.

Le verbe *hoer* “travailler avec la houe” est un dérivé de *hoe* “instrument aratoire” (< afrq. \**hauwa*, FEW 16,185a). Le lemmatiseur *Deucalion* n’a pas été capable de reconnaître le lemme auquel il se rattache (Fig. 4):

Id	Form	Lemma	POS	Morph	Context	Similar	Save	+
2965	hoé	-	VERppe	NOMB.=s CAS=r	ɿ fouy ɿ hoé , la crois		Save	

Fig. 4: difficulté de reconnaissance du lemme *hoer* dans *Deucalion*.

Un alignement des lemmes entre les différents dictionnaires permettrait d’avoir le résultat suivant (Table 4):

Dictionnaire	Lemme
DMF	<i>houer</i>
DEAF (H 526)	<i>hoer</i>
Gdf (4,508a)	<i>houer</i>
TL (4,1126)	<i>höer</i>

Table 4. Exemple de référentiel partagé d’alignement des lemmes des différents dictionnaires

Le nouveau TEAF permet de consulter les 43 formes classées sous le lemme *houer*: les données du DEAF, ordonnées par forme et puis par datation, sont restructurées de manière à pouvoir choisir de visualiser la datation et la localisation du texte et/ou du manuscrit de manière synoptique (figure 6).

FORM	DATE TEXT	DATE MS	LOC. TEXT	LOC. MS	SIGLE	POSITION
houer	1285	fin 13 <sup>s</sup> .	(pic.)	(pic.)	MaccabeS	2716
houer	fin 12es.	3 <sup>et</sup> . 13 <sup>s</sup> .	(pic.)	(pic.)	MonRaineB	6597
houer	fin 12es.	3 <sup>et</sup> . 13 <sup>s</sup> .	(pic.)	(pic.)	MonRaineB	6860
hoer	2em.13es.	1294	(pic.)	(pic.)	SEloiP	76a
houer	mil. 13es.	2 <sup>m</sup> . 14 <sup>s</sup> .	(pic.)	(N.-E.)	DoonMayP	10068
houer	1ert.13es.	fin 13 <sup>s</sup> .	(pic.)	(It. N.-E.)	OgDanB	6250
hoer	3et.12es.	3 <sup>q</sup> . 13 <sup>s</sup> .	(pic.)		AliseRé	591
heuer	doc. 1272		(pic.)		CPont	617,26

Fig. 6: extrait de la visualisation des données lexicales de *hoer* dans le TEAF

Il apparaît avec évidence que le lexème est attesté dans 9 fois dans 8 textes picards (dont 6 transmis par un manuscrit également picard), 7 fois (partout) en Lorraine (4 textes dont 3 avec

ms. lorr.), 5 fois en Flandre (3 textes), 3 fois en Champagne (3 textes) et ensuite dans des textes non localisés.

La correcte lemmatisation permet en outre un accès direct aux données du FEW (figure 7):

Fr. *houer* v. a. n. „travailler avec la houe“ (seit 12. jh.), *hoer* Guill Al, ahain. *hauer* Vox 11, 216, apik. *heuzer* (1499), lütt. *hawer*, nam. *awer*, Giv. *awè*, ard. *awè* „gazonner; recouvrir le seigle de terre“ (Brun 764, 1345), boul. *heuer* „houer“, Gondc. *oēi*, norm. *hower* Dm, Valognes *houwa*, sair. *hovō*, QChamps *hower*, Rethel *hoer*, VilleT. *huē* „piocher“, Buzancy *hōē* Babin 639, Carignan *hawē* „sarcler“ Brun 1413, Florent *haouer* „houer“, argonn. *hawè*, Brillon *haouer*, Dombras *haway*, Cum. *hawer*, Vouth. *houer* [*huē*], gaum. *howèye*, Metz *hawey*, Isle *hōwē*.

Fig. 7 : données du FEW 16,185a, s.v. \**hawwa*.

Wartburg n’a indiqué aucune marque diasystématique précise à *houer*, mais il s’est borné à signaler la *scripta* hennuyère et picarde des formes (pas des lexèmes!) *hauer* et *heuzer*; en revanche, l’identification de la répartition des données dialectales montre clairement que le lexème est attesté, à l’époque moderne, dans la Wallonie, la Flandre, les Ardennes, la Meuse et la Moselle.

#### 4.3.2. Le substantif “terroi”

Sans répéter la procédure adoptée pour *hoer*, nous nous contentons de signaler ici la présence d’un deuxième régionalisme lexical, le substantif *terroi* “territoire” (< TERRA + -ĒTUM). Le lexème n’a pas encore été traité par le DEAF, et dans le TEAF sont répertoriées 5 occurrences, notamment dans TristNantS (hain. mil. XIV<sup>e</sup> s., ms. hain. ca 1475), GodBouillBruxR (hain. ca 1356, ms. pic. mil. 15<sup>e</sup> s.), 2x CartHain (hain. 1274, ms. fin 13<sup>e</sup>s.) et AimeriD (attribué prob. à tort à Bertran de Bar(-sur-Aube) déb. XIII<sup>e</sup> s., var. du ms. *B*, traits pic., prob. Paris, ca 1335)<sup>43</sup>. Le FEW (13/1,244b, s.v. TERRA), mentionne quant à lui la forme *tieroit* «hain. flandr. pic. 13<sup>e</sup> s.», *terroit* (XIII<sup>e</sup> s.-1450), *terroi* (PrécSang, norm. ca 1320; jusqu’à 1520) et *terroy* (JLemaire; 1577). Les DocLing, enfin, enregistrent une forme *tieroit* dans chHain 133 (hain. 1255).

Le lexème est donc nettement caractéristique du Nord-Est, plus précisément de l’aire picarde et hennuyère, en nous livrant une information cohérente avec la distribution de *hoer*.

## 6. Conclusions

L’inspiration de notre contribution est surtout méthodologique. Nous avons voulu fournir un exemple concret de la manière dont l’exploitation des nouveaux outils informatiques (GallRom, avec son accès aux DocLing, au TEAF et aux analyses scriptologiques) associée aux méthodes d’intelligence artificielle (le programme HTR, le lemmatiseur *Deucalion*, le modèle d’interrogation stylométrique et scriptométrique) permet d’appréhender l’étude de la langue des textes littéraires transmis par des volumineux manuscrits-recueil. Il se dégagent alors des perspectives de travail qui dépassent le cadre de cet article mais qui permettront de mener des analyses plus fines également applicables à d’autres textes médiévaux. En particulier:

<sup>43</sup> Cf. AimeriD 624 (= AimeriG 620): «A moi s’en viengne ci clamer demenois» (ms. *A* traits bourg. ca 1270), «s’an va clamer en cest terrois» (ms. *B*), «s’an va clamer tout demenois» (ms. *C* lorr. mérid. 3<sup>e</sup> quart du XIV<sup>e</sup> s.).

- 1) l'amélioration de la qualité des transcriptions HTR permettra de disposer de masses textuelles importantes autant pour les analyses de type quantitatif que qualitatif. Non seulement l'intégralité du manuscrit de Turin pourra être prise en compte, mais aussi les autres copies des textes à tradition pour une analyse de stratigraphie horizontale;
- 2) pour augmenter la qualité des analyses, il est nécessaire de disposer du texte enrichi grâce au lemmatiseur *Deucalion*; dans une étape ultérieure, il serait envisageable d'ajouter le balisage grapho-phonétique structuré dans GallRom afin de pouvoir ainsi soumettre les textes à des analyses grapho-phonétiques, morpho-syntaxiques et lexicologiques fiables;
- 3) toujours dans le cadre d'une synergie entre les différents outils à disposition, il est souhaitable d'avoir à disposition un référentiel partagé d'alignement des lemmes, à son tour placé sous le chapeau du FEW;
- 4) l'élargissement et l'affinement du corpus textuel de comparaison représentent un atout indispensable. Le corpus documentaire des DocLing est presque complet et il sera bientôt possible de proposer des comparaisons entre les textes littéraires et documentaires pour tester l'apport effectif que ces derniers peuvent fournir à l'étude de la langue des premiers, ce qui représente un enjeu de première importance<sup>44</sup>. En revanche, le *Nouveau Corpus d'Amsterdam*, déjà soumis à des révisions et des critiques<sup>45</sup>, demande à être encore amélioré, par exemple sur la base des indications fournies par F. Möhren dans sa bibliographie du DEAF. Comme on l'a testé au cours de cette enquête, les modèles d'analyse basés sur l'IA permettent déjà d'affiner certaines localisations de textes ou de manuscrits;
- 5) en ce qui concerne l'analyse scriptologique, il conviendra de définir une grille de paramètres pertinents, surtout mais non seulement dans la dimension diatopique, comme il a récemment été souhaité par Lino Leonardi<sup>46</sup>.

Nous avons donc essayé de montrer dans quelle mesure l'étude linguistique d'un texte peut être exploitée non seulement dans la tentative de localisation de ce dernier, mais aussi dans l'optique d'une étude littéraire et d'attribution auctoriale.

Revenons donc sur la question que nous nous étions posés au début de cet article : Il est trop tôt et on manque de matériaux pour donner une réponse sur la composition des *unica* du premier tome par rapport au reste du manuscrit. Une réponse plus solide ne pourra venir qu'une fois que la transcription intégrale du manuscrit de Turin sera terminée. Néanmoins, nos analyses nous permettent d'entrevoir quelques pistes:

- 1) on remarque une *tendance* des *unica* de la section «sacrée» à être solidaires contre les autres textes à tradition;

---

44 Les premiers résultats d'une étude scriptométrique comparée se sont relevés être assez décevants. À part le fait que la *scripta* des documents administratifs pourrait, par sa nature, différer de manière assez importante de la *scripta* des textes à tradition, il se peut que l'analyse demande à être affinée avec un choix pondéré d'éléments significatifs sur lesquels le modèle pourra exercer ses analyses, sur l'exemple du choix des paramètres significatifs retenus dans notre analyse scriptologique.

45 Cf. M. Glessgen, X. Gouvert, La base textuelle du *Nouveau Corpus d'Amsterdam: ancrage diasystématique et évaluation philologique*, in *Le "Nouveau Corpus d'Amsterdam": actes de l'atelier de Lauterbad, 23-26 février 2006*, dir. P. Kunstmann, A. Stein, Stuttgart, Steiner, 2007, pp. 51-84, et M. Glessgen, C. Vachon, *L'étude philologique et scriptologique du "Nouveau Corpus d'Amsterdam"*, in *Actes du XXVI<sup>e</sup> Congrès International de Linguistique et Philologie Romanes (València, 6-11 sept. 2010)*, dir. E. Casanova Herrero, C. Calvo Rigual, Berlin, De Gruyter, 2013, pp. 191-204.

46 Nous nous référons à sa conférence *Paradigmi di filologia digitale e testi italiani antichi* donnée dans le cadre du séminaire de *Medioevo Romanzo 2024* "Filologia digitale e medioevo romanzo. Costi, benefici, prospettive" (Siena, 23-24 octobre 2023); le texte est à paraître dans le premier fascicule de *Medioevo Romanzo 2024*.

2) néanmoins, NoironsTur et SVou (les plus longs et les mieux édités), semblent parfois s'écarter légèrement des autres *unica*;

3) les rares données lexicales de SVou confirmeraient la provenance picarde des textes: si *hoer* est également attesté dans les Ardennes, *terroi* renverrait plutôt à la Picardie septentrionale.

Il nous paraît vraisemblable de penser que ce manuscrit soit le résultat d'une mise en recueil commencée avant et peut-être dans une région située plus au nord d'Origny dans l'Aisne. La plupart des textes à tradition transmis par ce manuscrit proviendraient en effet de la Picardie septentrionale. Le copiste aurait ainsi assemblé des matériaux composés en amont, du moins pour ce premier tome. Seule une analyse complète de l'ensemble du manuscrit et des autres *unica* qu'il contient – afin de déterminer s'ils sont cohérents avec ceux du premier tome – pourra apporter une réponse plus précise quant à sa conception globale unitaire, qui ne fait plus de doute à la lumière des études récentes sur T.

Marco Robecchi  
Université de Genève, [marco.robecchi@unige.ch](mailto:marco.robecchi@unige.ch)

Jean-Baptiste Camps,  
École nationale des chartes, Université PSL, [jean-baptiste.camps@chartes.psl.eu](mailto:jean-baptiste.camps@chartes.psl.eu)

## Annexes

### Annexe 1.

*Table des témoins utilisés pour l'analyse scriptométrique.*

Le label retenu est celui que nous avons établi sur la base des métadonnées du NCA (données dans les colonnes suivantes) et de la bibliographie du DÉAF

label retenu	deaf	vers	lieuComposition	dateComposition	lieuManuscrit	dateManuscrit
agn.	AngVieGregM	oui	agn.	1215pm1	agn.	1215pm1
agn.	EdmK*	oui	agn.	1195pm5	agn.	1325pm25
agn.	GuischartG_H	oui	agn.	1190pm10	agn.	1250pm10
agn.	IpH	oui	agn.	1185ca	agn.	1250pm10
agn.	ModvB2	oui	agn.	1230ca	agn.	1250ca
agn.	PlainteAmV	oui	agn.		1312 agn.	1325pm25
agn.	PsCambrM	non	agn.	1125pm25	Canterbury	1150pm10
agn.	Rois	non	agn.	1175pm25	agn.	1190pm10
agn.	SEdmCantB	oui	agn.	1255ca	agn.	1280pm20
agn.	SGillesP	oui	agn.	1183pm17	agn.	1225pm25
bourg.	BestGervM	oui	bourg.	1210pm10	bourg.	1275pm25
bourg.	FloovA	oui	Sud-Est	1190pm10	bourg.	1325pm25
bourg.	MerlinM	non	Sud-Est	1213pm13	bourg.	1250pm50
bourg.	MuleJ	oui	bourg.	1200ca	bourg. sept.	1310pm10
bourg.	SThibAlM	oui	bourg.	1290pm10	bourg.	1290pm10
champ. mérid.	CalendreS	oui	champ.	1215pm2	champ.	1235ca
champ. mérid.	PercLo_A	oui	champ. mérid.	1180ca	champ.	1235ca
champ. mérid.	YvainL_H	oui	champ. mérid.	1177ca	champ.	1235ca
champ. sept.	CoincyI*	oui	nord	1226pm2	pic. mérid.	1310pm10
champ. sept.	DrouartB	oui	champ.		1290 Aisne	1300ca
champ. sept.	RCambr2M	oui	Nord-Est	1190pm10	Nord	1275pm25
champ. sept.	RCambrM	oui	nord-est	1190pm10	pic.	1225pm25
champ. sept.	VolucrK	oui	champ.		1265 champ.	1288pm13
lorr.	BibleMalkS	oui	lorr.	1283pm17	lorr.	1283pm17
lorr.	DialAmeB	non	lorr.	1200ca	lorr.	1200ca

lorr.	PriseCordD	oui	lorr.	1200ca	Meuse	1262pm13
lorr.	PsLorrA	non	lorr.		1365 lorr.	1375pm25
lorr.	SBernAn1F	non	lorr.	1190pm10	Metz	1190pm10
lorr.	SBernAn2S	non	lorr.	1190pm10	lorr.	1190pm10
norm.	ChastPereAM	oui	norm.	1210pm10	norm.	1275pm25
norm.	ChastVergiS_ B	oui	norm.	1240ca	norm.	1300ca
norm.	ChronSMichel M	oui	Nord-Ouest	1163pm13	norm.	1275pm25
norm.	NarcisusP_D	oui	norm.	1165ca	norm.	1300ca
norm.	RouH_CA	oui	norm.	1167pm7	frc.	1300ca
norm.	SEust2P	oui	norm.	1210pm10	norm.	1250ca
Ouest	AdamA	oui	traits occ. et agn.	1175pm25	tour.	1238pm13
Ouest	BenDucF_Harl	oui	poit.	1174ca	agn.	1217pm17
Ouest	BenTroieC	oui	poit.	1170ca	agn.	1210pm10
Ouest	EstFougL	oui	Ouest	1176ca	Ouest	1217pm17
Ouest	SermMaurPB	non	La Rochelle	1250pm10	poit.	1250pm10
Ouest	SSilvCroixP	oui	Ouest	1200ca	tour.	1210pm10
Ouest	ThebesR_D	oui	poit.	1160ca	poit.	1200ca
Ouest	Turpin1A	non	saint.	1217pm17	saint.	1238pm13
pic.	AdHaleFeuillL	oui	art.		1276 pic.	1295
pic.	AdHaleRobS*	oui	art.	1285ca	Arras	1295
pic.	Aiol1/2N	oui	pic.	1160ca	pic.	1275pm25
pic.	AucR3	non	pic.	1225pm25	pic.	1290pm10
pic.	BeaumCoutB_ Harl	non	Oise		1283 frc.	1250pm50
pic.	BodelNicH2	oui	art.	1197pm6	Arras	1295ca
pic.	ChevBarAnL	oui	pic.	1210pm10	pic.	1310
pic.	ChevIIEspF	oui	pic.	1238pm13	Pas-de-Calais	1300ca
pic.	ComtePoitM	oui	pic.	1225pm25	pic.	1310pm10
pic.	ContPerc4Lo_ A	oui	pic.	1250pm17	pic. sept.	1275pm25

pic.	ElieB*	oui	pic.	1190pm10	pic.	1275pm25
pic.	EustMoineC	oui	pic.	1238pm13	pic.	1285
pic.	GarçAvR2	oui	pic.	1275ca	pic.	1275pm25
pic.	JacAmArtK	oui	flandr.	1250pm17	pic. ?	1300ca
pic.	MaccabES	oui	pic.		1285 Nord	1290pm10
pic.	MousketR	oui	hain.		1243 hain.	1275pm25
pic.	PerNeslesTabJ	oui	pic.		1289 pic.	1289
pic.	PhMézPel_A	non	pic.		1389 nil	1400ca
pic.	PriereTheopS	oui	pic.	1290pm10	pic.	1300ca
pic.	RenNouvR	oui	Lille	1290ca	Arras	1295
pic.	RobClariL	non	pic.	1217pm1	pic.	1300ca
pic.	RobOmMirL	oui	pic.		1266 art.	1277
pic.	SJeanBoucheD	oui	pic.	1200ca	pic.	1370
wall.	DialGregF	non	liég.	1190pm10	Est	1210pm10
wall.	GIWallW*	non	wall.	1290pm10	wall.	1290pm10
wall.	JacBaisT	oui	wall.	1300ca	wall.	1300ca
wall.	JuiseR	oui	liég.	1137pm13	wall.	1210pm10
wall.	MédLiégH	non	liég.	1275pm25	wall.	1275pm25
wall.	PoèmeMorB_A	oui	wall.	1200ca	wall.	1210pm10
wall.	SermCarP	non	liég.	1217pm17	liég.	1225pm25

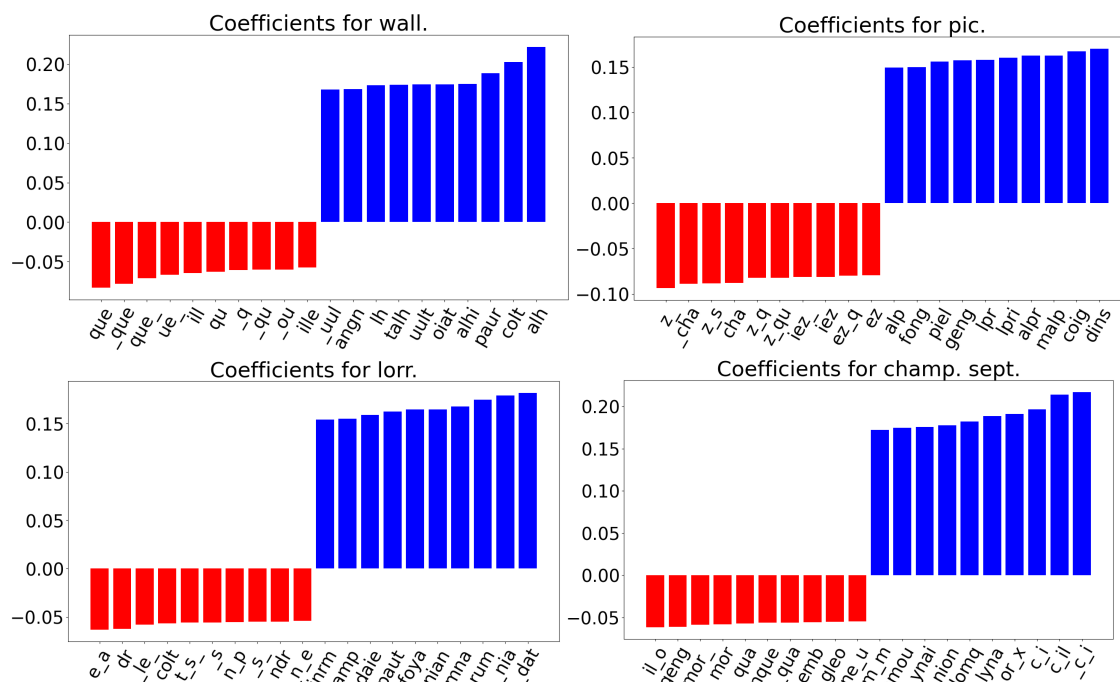
## Annexe 2.

### *Entraînement des modèles scriptométriques*

	<b>précision</b>	<b>rappel</b>	<b>F1-score</b>	<b>support</b>
Ouest	0.90	0.60	0.72	47
agn.	0.86	0.99	0.92	135
bourg.	0.56	0.21	0.30	24
champ. mérid.	0.73	1.00	0.85	11
champ. sept.	0.64	0.57	0.60	28
lorr.	0.91	0.66	0.76	32
norm.	0.56	0.90	0.69	20
pic.	0.83	0.88	0.86	109
wall.	1.00	1.00	1.00	28
<i>exactitude</i>			<i>0.82</i>	<i>434</i>

<i>moy. macro</i>	0.78	0.76	0.74	434
<i>moy. pond.</i>	0.82	0.82	0.81	434

Score obtenus à l'entraînement, sur la base des textes sélectionnés dans le NCA, avec une stratégie de validation croisée de type «k-fold» de groupe (à chaque itération, un texte complet, avec tous ses échantillons, est retiré du corpus d'entraînement pour servir de test, et la stratégie est répétée durant autant d'itérations qu'il y a de textes différents en entraînement).



Principaux coefficients (négatifs et positifs) utilisés par les modèles pour distinguer les “scriptae”, ici pour le wallon, le picard, le lorrain et le champenois septentrional. Les traits en question sont des n-grammes de caractère de différentes longueurs extraits des textes.

### Annexe 3.

#### Fiches d'analyse scriptologique

Dans l'analyse, nous indiquons les pourcentages des régions de Dees pertinentes pour notre recherche, à savoir la Picardie et la Champagne septentrionale. Par la suite, les atlas de Dees sont respectivement cités comme DeesAtlas (la publication de 1980) et DeesAtlas<sup>2</sup> (la publication de 1987), conformément aux sigles de la bibliographie du DEAF; les cartes des Atlas sont citées comme «c.». Nous mentionnons en outre les analyses scriptologiques effectuées dans GallRom (en signalant le nom du Head sous lequel les formes sont analysées) ou, à défaut d'une analyse plus précise, les données des DocLing que nous avons décomptées manuellement. Les sigles des corpus documentaires mentionnés sont celles de la base GallRom, à laquelle nous renvoyons pour plus de détails bibliographiques; nous indiquons ici: chDo = chartes de la ville de Douai (1204-1270); chHain = chartes de la province de Hainaut (1236-1270); chFland = chartes de la province de la Flandre romane (1224-1270); chMa = chartes du département de la Marne (1234-1272); chMe = chartes du département de la Meuse (1225-1270); chMM = chartes du département de la Meurthe-et-Moselle (1232-1265); chMo = chartes du département de la Moselle (1215-1271); chV = chartes du département des Vosges (1235-1275); chOise = chartes

du département de l’Oise (1241-1286); docNorm = chartes de la Normandie (1250-1300). Dans les Atlas et dans GallRom, les oppositions scriptologiques sont signalées par le symbole «~», et le chiffre suivi d’un «x» indique le nombre d’occurrences décomptées. Nous mentionnons en outre les données lexicographiques du *Dictionnaire Étymologique de l’Ancien Français* (DEAF) actuellement intégrées à la base GallRom sous le titre *Trésor Étymologique de l’Ancien Français* (TEAF). Nous proposons un renvoi systématique aux données étymologiques de W. von Wartburg, *Französisches Etymologisches Wörterbuch. Eine darstellung des galloromanischen sprachschatzes*, Bonn/Heidelberg/Leipzig-Berlin/Basel, Schroeder/Klopp/Winter/Teubner/Zbinden, 1922-2002, 25 voll. (FEW); pour les sigles des éditions utilisées par FEW, cf. J.-P. Chauveau, Y. Greub, C. Seidl, *Französisches Etymologisches Wörterbuch. Complément*, Strasbourg, ÉLiPhi, 2010. Nous renvoyons également à F. Godefroy, *Dictionnaire de l’ancienne langue française et de tous ses dialectes du IX<sup>e</sup> au XV<sup>e</sup> siècle*, Paris, 1880-1902, 10 voll. (Gdf). Les abréviations utilisées pour les œuvres littéraires citées sont celles de la bibliographie du DEAF.

Dans les tableaux suivants, nous séparons les *unica* de *T* des textes à tradition copiés dans le même manuscrit<sup>47</sup>.

#### 4.1. Traits à forte pertinence

##### 4.1.1. Pronom “je”<sup>48</sup>

	tot.	je <sup>49</sup>	Ge	jou <sup>50</sup>	jo
<b>GenealTur</b>	1	0	0	1 (100%)	0
<b>NoironsTur</b>	22	12 (54,55%)	4 (18,18%)	6 (27,27%)	0
<b>SVou</b>	27	23 (85,18%)	0	4 (14,82%)	0
<b>MahomTur</b>	4	3 (75%)	0	1 (25%)	0
<b>Prolloh</b>	12	11 (91,67%)	0	1 (8,33%)	0
<b>HermVal</b>	44	41 (93,18%)	1 (2,27%)	2 (4,55%)	0
<b>SFmanuel</b>	28	24 (85,71%)	0	4 (14,29%)	0
<b>VenjNS</b>	176	123 (69,89%)	1 (0,57%)	51 (28,97)	1 (0,57%)
<b>Hervis<sup>51</sup></b>	620	456 (73,55%)	0	164 (26,45%)	0

FEW 3,207a, s.v. EGO «pic. *jou* Esc»

DeesAtlas, c. 1 *g- ~ i-, j-* du pronom “je”: Somme, Pas-de-Calais 2,47%; Oise 4,13%; Aisne 0,40%; Nord 0%; Hainaut 0,74%; Wallonie 6,85%; Ardennes 3,26%; Marne 9,03%

DeesAtlas, c. 2 *-o, -ou, -u ~ -e, -ei* du pronom “je”: Somme, Pas-de-Calais 40,82%; Oise 6,03%; Aisne 37,14%; Nord 76,56%; Hainaut 85,27%; Wallonie 17,27%; Ardennes 0%; Marne 0%

DeesAtlas, c. 3 *jou ~ jo* du pronom “je”: Somme, Pas-de-Calais 96,49%; Oise 100%; Aisne 100%; Nord 82,15%; Hainaut 83,21%; Wallonie 0; Ardennes 0; Marne 0

47 Nous avons décidé de ne pas analyser la forme courte du possessif *no, vo* contre *nostre, vostre*; cette forme, typique du picard, connaît un certain succès dans la tradition littéraire du XIII<sup>e</sup> sec. et vu sa forme monosyllabique est très utile pour des raisons métriques, si bien qu’il est difficile d’évaluer sa valeur au niveau *scriptologique*.

48 Cf. Y. Greub, O. Collet, *La variation régionale de l’ancien français. Manuel pratique*, Strasbourg, ÉLiPhi, 2023, § 4.1: *jo, jou* typiques de l’agn., Picardie, part. Wallonie («dans l’ouest, le centre et l’est de la France, le type *jo, jou* est absent [...] La forme *jou* est particulière, dans les chartes, au nord de la Picardie»); les formes avec *g-* sont attestées surtout dans l’Ouest et le Sud-Ouest.

49 Dans le cas de *j’*, nous le comptons sous *je*.

50 Cf. GodinM, p. LXI, qui indique la forme *jou* tonique ou atone et *je* (et *ge*) uniquement atone dans la partie rimée; en revanche, *jou* est absent dans la partie assonancée (cf. sa liste des occurrences dans la même page).

51 HervisH, p. LII relève «*jeu* (5 exemples), *jou* (9 exemples), *ju* (3 exemples)», ms. de base BN fr. 19160 [lorr. 2<sup>e</sup> t. XIII<sup>e</sup> s.].

DeesAtlas<sup>2</sup>, c. 1 *jo, jou ~ je, ge, gie*: Somme, Pas-de-Calais 24,04%; Oise 0,86%; Aisne 0,97%; Nord 23,40%; Hainaut 48,34%; Wallonie 2,44%; Ardennes 1,87%; Marne 0,94%  
 GallRom: “j\_EGO”, en dehors de l’Ouest, la forme avec consonne *g-* est ponctuellement attestée dans les chOise (*g-* 29x ~ *j-* 1070x) et chMa (*g-* 70x ~ *j-* 799); “eo EGO”, les formes *ou* (tot. 407x) et *o* (tot. 127x) sont attestées dans chOise (*-ou* 88x ~ *-o* 1x ~ *-e* 935x), chDo (*-ou* 52x ~ *-o* 8x ~ *-e* 27x), chHain (*-ou* 146x ~ *-o* 12x ~ *-e* 219x), chMa (*-ou* 9x ~ *-e* 817x).

Dans notre *corpus*, les *unica* et les textes à tradition emploient la forme *je* essentiellement en égale mesure ; dans le ms. de Turin, les rares formes du pronom de P1 commençant par *g-* sont utilisées devant des mots commençant par voyelle, surtout *i-* (ces formes *ge* sont plus fréquentes à l’Ouest du domaine oïlique).

#### 4.1.2. Démonstratifs “*cest*”, “*cel*”, “*celui*”

	tot.	<i>ce(l/st)</i>	<i>che(l/st)</i>
<b>GenealTur</b>	0	0	0
<b>NoironsTur</b>	2	2 (100%)	0
<b>SVou</b>	15	15 (100%)	0
<b>MahomTur</b>	5	5 (100%)	0
<b>Prolloh</b>	8	7 (87,50%)	1 (12,50%)
<b>HermVal</b>	39	1 (24,56%)	38 (97,44%)
<b>SFmanuel</b>	67	63 (94,03%)	4 (5,97%)
<b>VenjNS</b>	148	148 (100%)	0
<b>Hervis</b>	65	64 (98,46%)	1 (1,54%)

FEW 4,552a, s.v. ÎLLE [ECCE+ÎLLE]; FEW 4,820a, s.v. ÎSTE [ECCE+ÎSTE]

DeesAtlas, c. 59 (*ich-* ~ (*ic-*, (*is-* adj. pron. dém.: Somme, Pas-de-Calais 67,38%; Oise 52,69%; Aisne 11,99%; Nord 5,74%; Hainaut 16,87%; Wallonie 15,04%; Ardennes 0,2%; Marne 0%  
 GallRom: “k ECCE” (tot. 24667x): les formes avec le résultat <ch-> (tot. 410x) sont ainsi réparties: 76x docNorm, 209x chOise, 23x chDo, 34x chFland, 17x chHain, 2x chMa

Les formes de type picard commençant par *ch-* sont fortement majoritaires uniquement chez HermVal; les autres textes préfèrent largement les formes non-palatales.

#### 4.1.3. Démonstratif “*ce*”

– *vocalisme (e / o ; ou / eu)*

	tot.	<i>c(h)e</i>	<i>ç(h)ou</i>	<i>ceu</i>	<i>ço</i>
<b>GenealTur</b>	1	1 (100%)	0	0	0
<b>NoironsTur</b>	9	7 (77,78%)	2 (22,22)	0	0
<b>SVou</b>	11	6 (54,55%)	5 (45,45%)	0	0
<b>MahomTur</b>	10	2 (20%)	8 (80%)	0	0
<b>Prolloh</b>	2	2 (100%)	0	0	0
<b>HermVal</b>	42	36 (85,71%)	6 <sup>52</sup> (14,29%)	0	0
<b>SFmanuel</b>	77	28 (36,36%)	47 (61,04%)	0	2 (2,60%)
<b>VenjNS</b>	271	91 (33,58%)	180 (66,42%)	0	0
<b>Hervis</b>	333	289 (86,79%)	41 (12,31%)	3 (0,9%)	0

52 Dont 1 *chu*.

DeesAtlas, c. 29 -o, -ou ~ -e, ei pronom “ce”: Somme, Pas-de-Calais 36,62%; Oise 3,52%; Aisne 15,46%; Nord 42,44%; Hainaut 55,91%; Wallonie 25,36%; Ardennes 2,56%; Marne 0%

DeesAtlas, c. 30 -o, -ou, -u ~ -eo, eou pronom “ce”: Somme, Pas-de-Calais 97,73%; Oise 85,71%; Aisne 100%; Nord 100%; Hainaut 100%; Wallonie 96,41%; Ardennes 0; Marne 0

DeesAtlas, c. 32 -ou ~ -o pronom “ce”: Somme, Pas-de-Calais 91,89%; Oise 100%; Aisne 92,86%; Nord 91,78%; Hainaut 80,25%; Wallonie 81,25%; Ardennes 0; Marne 0

DeesAtlas<sup>2</sup>, c. 6 co, cou, cho, chou ~ ce, che, ceo, ceu pronom “ce”: Somme, Pas-de-Calais 36,86%; Oise 5,97%; Aisne 5,90%; Nord 69,70%; Hainaut 62,97%; Wallonie 0,42%; Ardennes 11,37%; Marne 1,89%

DeesAtlas<sup>2</sup>, c. 7 c(h)ou ~ c(h)o pronom “ce”: Somme, Pas-de-Calais 98,22%; Oise 100%; Aisne 100%; Nord 100%; Hainaut 93,33%; Wallonie 50%; Ardennes 100%; Marne 75%

GallRom: “ce-voc” (tot. 14 385x): les formes avec <ou> (983x), <o> (796x) et <eu> (1258x) sont ainsi réparties: docNorm (eu 11x), chOise (ou 27x ~ o 1x ~ eu 2x), chDo (ou 429x ~ o 26x), chFland (ou 92x ~ o 14x ~ eu 1x), chHain (ou 205x ~ o 18x), chMa (ou 4x ~ o 1x ~ eu 6x)

On remarque une légère prédominance de la forme -ou dans MahomTur, SFanuel et VenjNS, mais la distribution penche plutôt pour -e dans tous les textes, en cohérence avec la distribution de -ou contre -e relevée par Dees pour l’Aisne, ce qui pourrait indiquer une tendance à la neutralisation du scribe de T.

#### “Ce”: consonantisme

	tot.	ce/ou	che/ou
<b>GenealTur</b>	1	0	1 (100%)
<b>NoironsTur</b>	9	5 (55,56%)	4 (44,44%)
<b>SVou</b>	11	2 (18,18%)	9 (81,82%)
<b>MahomTur</b>	10	1 (10%)	9 (90%)
<b>Prolloh</b>	2	1 (50%)	1 (50%)
<b>HermVal</b>	42	2 (4,76%)	40 (95,24%)
<b>SFanuel</b>	77	63 (81,82%)	14 (18,18%)
<b>VenjNS</b>	271	255 (94,10%)	16 (5,90%)
<b>Hervis</b>	333	285 (85,59%)	48 (14,41%)

DeesAtlas, c. 27 ch- ~ c-, s- pronom “ce”: Somme, Pas-de-Calais 75,62%; Oise 53,57%; Aisne 20,56%; Nord 21,46%; Hainaut 38,92%; Wallonie 43,33%; Ardennes 0%; Marne 0%

DeesAtlas<sup>2</sup>, c. 4 ch-, ich- ~ c-, ic-, s-, is- pronom “ce”: Somme, Pas-de-Calais 49,09%; Oise 3,20%; Aisne 2,51%; Nord 0%; Hainaut 18,49%; Wallonie 10,26%; Ardennes 0%; Marne 0%

GallRom: “k\_ECCE” (tot. 15 713x): les formes avec <ch> (1230x) sont ainsi réparties: 144x docNorm, 519x chOise, 231x chDo, 72x chFland, 139x chHain, 10x chMa

Les formes picardes commençant par ch- sont nettement majoritaires dans HermVal (comme *supra*, 4.1.2.) et plutôt fréquentes dans les *unica*. Les pourcentages sont bien majeurs que ceux proposés par Dees pour l’Aisne, autant pour les chartes que pour les textes littéraires : ces données peuvent indiquer une origine peut-être plus septentrionale des textes transcrits dans T, ou bien le copiste a adopté un trait picard plus marqué.

#### “Ce”: vision d’ensemble<sup>53</sup>

<sup>53</sup> Greub, Collet, *Variation* cit., § 6.5: «Les formes çou/chou sont particulièrement picardes (normalement en concurrence avec c(h)e) [...] les formes en <ch> sont picardes et secondairement, sous l’influence picarde, normandes et wallonnes».

	tot.	ce	che	ceu	çou	chou	ço
<b>GenealTur</b>	1	0	1 (100%)	0	0	0	0
<b>NoironsTur</b>	9	5 (55,56%)	2 (22,22%)	0	0	2 (22,22%)	0
<b>SVou</b>	11	1 (9,09%)	5 (45,45%)	0	1 (9,09%)	4 (36,36%)	0
<b>MahomTur</b>	10	1 (10%)	1 (10%)	0	0	8 (80%)	0
<b>Prolloh</b>	2	1 (50%)	1 (50%)	0	0	0	0
<b>HermVal</b>	42	2 (4,76%)	34 (80,95%)	0	0	6 <sup>54</sup> (14,29%)	0
<b>SFmanuel</b>	77	23 (29,87%)	5 (6,49%)	0	38 (49,35%)	9 (11,69%)	2 (2,58%)
<b>VenjNS</b>	271	88 (32,47%)	3 (1,11%)	0	167 (61,62%)	13 (4,80%)	0
<b>Hervis</b>	333	282 (84,68%)	7 (2,10%)	3 (0,90%)	0	41 (12,32%)	0

FEW 4,442a, s.v. HÖC: «aflandr. *chou* Roisin, apic. anorm. *cho*» (< ECCE+HÖC)

GallRom, “k ECCE”: la forme *çou* (636x) ne se profile pas spécifiquement comme picarde (elle apparaît également dans les autres corpus orientaux ainsi qu’en Poitou et Anjou-Touraine), quoiqu’on décompte 223x chDo, 43x chFland, 120x chHain. La forme *che* est fréquente en Picardie mais surtout dans l’Oise (422x) contre 28x chDo, 19x chFland, 41x chHain.

Les distributions désignent un cadre où Hervis privilégie la forme ‘neutre’ *ce*, HermVal la forme de l’Oise (soit picarde mérid.) *che*, VenjNS la forme picarde sept. *çou*. Les autres textes montrent des distributions moins cohérentes.

#### 4.1.4. Démonstratif “cel” (formes du pluriel)<sup>55</sup>

	tot.	c(h)eus, che(l)s	ciaus, ciax	ciax	chiaus, chiaux	chil
<b>GenealTur</b>	0	0	0	0	0	0
<b>NoironsTur</b>	0	0	0	0	0	0
<b>SVou</b>	3	0	0	0	3 (100%)	0
<b>MahomTur</b>	2	0	0	0	0	2 (100%)
<b>Prolloh</b>	3	0	0	0	1 (33,33%)	2 (66,67%)
<b>HermVal</b>	4	2 (50%)	0	0	1 (25%)	1 (25%)
<b>SFmanuel</b>	8	4 (50%)	?	1 (12,50%)	3 (37,50%)	0
<b>VenjNS</b>	26	18 (69,23%)	6 (23,08%)	0	2 (7,69%)	0
<b>Hervis</b>	24	0	0	4 (16,67%)	20 (83,33%)	0

FEW 4,552a, s.v. ÎLLE [ECCE+ÎLLOS]

DeesAtlas, c. 68 *-el-*, *-eu-* ~ *-(e,i)a-*, *-(e)o-* pronom “ceux”: Somme, Pas-de-Calais 16,80%; Oise 31,78%; Aisne 19,44%; Nord 17,36%; Hainaut 0%; Wallonie 3,60%; Ardennes 38,57%; Marne 48,96%

DeesAtlas, c. 71 *c(h)ia-* ~ *c(h)ea-* pronom “ceux”: Somme, Pas-de-Calais 85,48%; Oise 89,23%; Aisne 92%; Nord 70,69%; Hainaut 91,80%; Wallonie 20,45%; Ardennes 45,74%; Marne 53,33%

DeesAtlas, c. 73 *-l-* ~ *-u-*, *-ul-*, pronom “ceux”: Somme, Pas-de-Calais 0%; Oise 0,68%; Aisne 8,33%; Nord 15,28%; Hainaut 15,62%; Wallonie 4,29%; Ardennes 5,42%; Marne 15,62%

GallRom: “ceus-voc” (tot. 6410x): les formes avec <ia> (325x) et <i> (1231x) se distribuent ainsi: chOise (*ia* 121x ~ *i* 39x), chDo (*ia* 17x ~ *i* 500x), chFland (*ia* 23x ~ *i* 11x), chHain (*ia* 84x ~ *i* 72x), chMa (*ia* 29x ~ *i* 70x)

<sup>54</sup> Dont 1 *chu*.

<sup>55</sup> Dans HermVal et SFmanuel on trouve *cheus*; dans VenjNS, deux occurrences de *ceus* se placent au tout début du texte, tandis que par la suite on ne trouve que *ciaus*. Cf. Greub, Collet, *Variation* cit., § 6.1: «*cels* est beaucoup plus fréquent dans l’Est que dans l’Ouest»; «le type *c(h)iaus* avec *-i-*, est presque exclusivement picard (quelques prolongements en Champagne et Wallonie)».

Les *unica* privilégient largement les formes picardes *chiaus* et *chil*, en cohérence avec le relevé de Dees pour l’Aisne; les autres montrent une bonne présence de la forme neutre *ceus*.

#### 4.1.5. Article contracté “*du*”<sup>56</sup>

	tot.	<i>du</i>	<i>Dou</i>	<i>del</i> (suivi cons.)
<b>GenealTur</b>	9	0	9 (100%)	0
<b>NoironsTur</b>	4	1 (25%)	3 (75%)	0
<b>SVou</b>	9	2 (22,22%)	7 (77,78%)	0
<b>MahomTur</b>	2	1 (50%)	1 (50%)	0
<b>Prolloh</b>	9	4 (44,44%)	4 (44,44%)	1 (11,12%)
<b>HermVal</b>	28	27 (96,43%)	0	1 (3,57%)
<b>SFmanuel</b>	29	26 (89,66%)	1 (?)	2 (6,90%)
<b>VenjNS</b>	67	47 (79,15%)	20 (29,85%)	0
<b>Hervis</b>	294	99 (33,67%)	90 (30,61%)	105 (35,72%)

DeesAtlas, c. 42 *del, deu, du ~ do, dou, dau, dor, dour* art. “*du*”: Somme, Pas-de-Calais 93,92%; Oise 51,73%; Aisne 84,38%; Nord 35,48%; Hainaut 4,60%; Wallonie 57,43%; Ardennes 2,60%; Marne 9,65%

DeesAtlas, c. 43 *deu, du ~ del* art. “*du*”: Somme, Pas-de-Calais 79,95%; Oise 98,44%; Aisne 80,15%; Nord 6,39%; Hainaut 0%; Wallonie 2,50%; Ardennes 0; Marne 0

DeesAtlas, c. 44 *du ~ deu* art. “*du*”: Somme, Pas-de-Calais 93,13%; Oise 91,40%; Aisne 25%; Nord 100%; Hainaut 0; Wallonie 0; Ardennes 0; Marne 0

DeesAtlas<sup>2</sup>, c. 84 *dau, do, dor, dour ~ del, deu, du* de l’art. “*du*”: Somme, Pas-de-Calais 13,67%; Oise 44,23%; Aisne 44,25%; Nord 31,25%; Hainaut 4,60%; Wallonie 14,98%; Ardennes 47,18%; Marne 69,11%

DeesAtlas<sup>2</sup>, c. 85 *deu, du ~ del* de l’art. “*du*”: Somme, Pas-de-Calais 63,22%; Oise 6,41%; Aisne 35,45%; Nord 50%; Hainaut 98,52%; Wallonie 100%; Ardennes 75%; Marne 21,56%

GallRom: “*du*” (tot. 37478x), les formes *dou* (10798x) sont nettement majoritaires par rapport à *del* (6130x) et *du* (4060x); voir chOise (*dou* 317x ~ *del* 5x ~ *du* 256x), chDo (*dou* 280x ~ *del* 185x ~ *du* 7x), chFland (*dou* 106x ~ *del* 44x ~ *du* 4x), chHain (*dou* 87x ~ *del* 128x ~ *du* 2x) et chMa (*dou* 356x ~ *del* 40x ~ *du* 25x)

On remarque ici une certaine bipartition dans les résultats: sauf Hervis, qui adopte la forme *del* (plutôt typique de la Picardie septentrionale), les *unica* préfèrent la forme *dou* (apparemment plus neutre au XIII<sup>e</sup> siècle et pas majoritaire dans l’Aisne, selon Dees) tandis que les textes à tradition préfèrent la forme *du*.

#### 4.1.6. Pronom indéfini “*chascun*”<sup>57</sup>

	tot.	<i>chascun</i>	<i>Cascun</i>	<i>chescun</i>	<i>cescun</i>
<b>GenealTur</b>	0	0	0	0	0
<b>NoironsTur</b>	0	0	0	0	0
<b>SVou</b>	1	0	0	0	1 (100%)
<b>MahomTur</b>	1	0	0	1 (100%)	0
<b>Prolloh</b>	3	0	0	1 (33,33%)	2 (66,67%)
<b>HermVal</b>	8	5 (62,50%)	3 (37,5%)	0	0

<sup>56</sup> Greub, Collet, *Variation* cit., § 3.2: «*deu* et *du* sont presque complètement absents en Wallonie, Hainaut, Nord [...] *del* est presque complètement absent dans l’Ouest et dans le Centre [...], et serait rare en Champagne méridionale, en Lorraine et en Bourgogne [...] l’opposition entre *del* et *du* passe entre Picardie méridionale (*du*) et Picardie septentrionale; *dou* est presque entièrement absent en Normandie et en Picardie».

<sup>57</sup> Cf. Greub, Collet, *Variation* cit., § 1.47 consacré à l’aboutissements de A initial atone: «Les formes en -e- de *chascun* semblent très largement répandues (tout l’Est et le Nord-Est, ainsi que l’anglo-normand, FEW 2,482a)».

<b>SFmanuel</b>	11	1 (9,09%)	10 (90,91%)	0	0
<b>VenjNS</b>	64	0	63 (98,44%)	1 (1,56%)	0
<b>Hervis</b>	48	3 (6,25%)	44 (91,67%)	1 (2,08%)	0

FEW 2,482a, s.v. CATA: «aflandr. *cascun* Roisin, *chescun* (Tournai 1280) ... achamp. *chesquen* (XIV<sup>e</sup> s.)»  
 DeesAtlas<sup>2</sup>, c. 47 *c(h)e-* ~ *c(h)a-* pronom indéf. “chacun”: Somme, Pas-de-Calais 0,48%; Oise 0%; Aisne 0%; Nord 0%; Hainaut 0%; Wallonie 6,6%; Ardennes 0%; Marne 0% (pourcentages plus élevés que 0 Vendée, Deux-Sèvres 21,03%, Indre-et-Loire 80%, Normandie 37,45%, Aube 0,56%, Haute-Marne 0,63%, Meuse 33,33%, Franche-Comté 25%, Nièvre, Allier 4,64%, Angleterre 55,16%)  
 DeesAtlas<sup>2</sup>, c. 48 *c-*, *k-*, *qu-* ~ *ch-* pronom indéf. “chacun”: Somme, Pas-de-Calais 71,44%; Oise 33,33%; Aisne 22,22%; Nord 100%; Hainaut 100%; Wallonie 34,80%; Ardennes 20%; Marne 14,29%  
 DocLing: forme *cascun* (tot. 862x): 259x chDo, 32x chFland, 38x chHain, 2x chMa, 5x chOise (et 4x dans 2 chartes des docNorm), le reste des attestations proviennent de l’aire occitane; forme *chascun* (tot. 3160x) attestée dans tous les corpus; forme *cescun* (tot. 42x): 8x chFland, 23x chHain (et 11x docNorm); forme *chescun* (tot. 728x): attestée partout sauf dans nos corpus de référence (1x chHain, 1x chMa).

La bipartition est nette entre le choix des *unica* (forme en *-e-*, externe au Nord-Est d’après Dees) et celui des textes à tradition qui montrent la forme plus “neutre” en *-a-*, avec une nette préférence pour le consonantisme de type picard *c-*.

#### 4.1.7. Évolution de *K* initial suivi de voyelle palatale<sup>58</sup>

Formes répertoriées: *cherché* (< CIRCELLUS)<sup>59</sup>, *cheler* (< CELARE), *chelier* (< CELLARIUM), *chiterne* (< CISTERNA), *chité* (< CIVITAS), *cens* (< CENTU), *celestial* (< CELESTIALIS), *cesser* (< CESSARE), *chierge* (< CEREUS), *cire*, *ciree* (< CERA), *citole* (< CITARA), *cerkier* (< CIRCARE), *celebrer* (< CELEBRARE), *cent*, *centoisme* (< CENTU), *cedre* (< CINIS)  
 Formes individuelles: *certes* (< CERTUS), *ciel* (< CAELUM).

	tot.	<i>ce - cie</i>	<i>certes</i>	<i>ciel</i>	<i>che - chi</i>	<i>chertes</i>	<i>chiel</i>
<b>GenealTur</b>	0	0	0	0	0	0	0
<b>NoironsTur</b>	7	2 (28,57%)	0	0	0	0	5 (71,43%)
<b>SVou</b>	2	0	0	0	0	0	2 (100%)
<b>MahomTur</b>	3	0	0	0	0	0	3 (100%)
<b>Prolloh</b>	0	0	0	0	0	0	0
<b>HermVal</b>	19	1 (5,26%)	0	0	6 + 2 (42,11%)	10 (52,63%)	0
<b>SFmanuel</b>	32	7 (21,87%)	3 (9,38%)	17 (53,12%)	3	0	2 <sup>60</sup>

58 Cf. Greub, Collet, *Variation* cit., § 2.1: *ch-* est typique de Normandie et Picardie; «dans l’Est de la Picardie, le digraphe <ch> est relativement moins fréquent». Remarque importante: «il n’est pas exclu que la répartition entre les graphies <ch> et <c> varie aussi en fonction de la position du phonème dans le mot, et plus généralement selon les unités lexicales» (C. T. Gossen, *Grammaire de l’ancien picard*, Paris, Klincksieck, 1970 «Bibliothèque française et romane» 19, pp. 92-93). Les démonstratifs et les pronoms sont exclus (cf. *supra* 4.1.2.-4.1.4.). Cf. GodinM, p. XLIX, qui indique une attestation de la forme pl. *chius*, en cohérence avec la section des *unica*; en revanche, GodinM, p. LV, signale uniquement le résultat *ch-* sans donner de contrexemples, quoiqu’il enregistre *cerkier* (< CIRCARE) dans la section consacrée à *k+a*. Les formes avec *ch-* caractérisent également la partie assonancée, aussi pour les démonstratifs 1x *chiaus*, 1x *chou* et 1x *che* (GodinM, p. LXX). En revanche, AubS, p. XV signale uniquement «*c+e* ou *i = ch*», sans plus.

59 Cf. 1195 (ms. lorr. fin XIII<sup>e</sup> s.): «*et* contreval sa barbe menu *recercelee*», var. ms. de Turin: «*et* en aval se barbe menue *cherchelee*»; l’adj. *cercelé* ‘frisé, bouclé’ (FEW 2,699b s.v. CIRCĒLLUS) est plus rare du préfixé *recerché* et attesté uniquement dans l’*Escoufle* (pic. ca 1201, ms. pic. fin XIII<sup>e</sup> s.), dans l’*Entrée d’Espagne* (françoit. 1<sup>e</sup> moitié du XIV<sup>e</sup> s.) (cf. FEW, *ibid.*), dans une légende de Saint George du ms. BnF fr. 20330 (début du XIV<sup>e</sup> s.) (cf. Gdf 18,2c s.v. *cerceler*) et le *Roman de toute chevalerie* de Thomas de Kent (agn. 4<sup>e</sup> quart du XII<sup>e</sup> s., ms. agn. 3<sup>e</sup> quart du XIV<sup>e</sup> s.).

					(9,38%)		(6,25%)
<b>VenjNS</b>	62	21 + 28 (79,03%)	9 (14,52%)	4 <sup>61</sup> (6,45%)	0	0	0
<b>Hervis</b>	58	0	55 (94,83%)	3 (5,17%)	0	0	0

FEW 2,34a, s.v. CAELUM; 2,610b, s.v. CERTUS

DeesAtlas, c. 117 *ch-* ~ *c-* subst. “cent(s)”: Somme, Pas-de-Calais 37,65%; Oise 56,01%; Aisne 0%; Nord 0%; Hainaut 1,69%; Wallonie 2,41%; Ardennes 0; Marne 0%

DeesAtlas, c. 133 *ch-* ~ *c-* subst. “cens”: Somme, Pas-de-Calais 65%; Oise 59,57%; Aisne 0%; Nord 0%; Hainaut 0,67%; Wallonie 7,68%; Ardennes 0; Marne 0%

DeesAtlas<sup>2</sup>, c. 165 *ch-* ~ *c-*, *s-* subst. “ciel”: Somme, Pas-de-Calais 27,91%; Oise 0%; Aisne 0%; Nord 0%; Hainaut 8,33%; Wallonie 18,33%; Ardennes 0%; Marne 14,29%

DeesAtlas<sup>2</sup>, c. 480 *chertes* ~ *certes*, *sertes* adv. “certes”: Somme, Pas-de-Calais 26,21%; Oise 0%; Aisne 0%; Nord 0%; Hainaut 0%; Wallonie 0%; Ardennes 0%; Marne 0%

DocLing: cf. 4x *chertain* chFland et 3x *chertain* chOise

GallRom: “k E/I” (tot. 17 826x): les formes avec le résultat <ch> initial (tot. 1095) sont ainsi réparties: 150x docNorm, 658x chOise, 147x chDo, 50x chFland, 61x chHain, 4x chMa.

On remarque ici une certaine préférence des *unica* pour la forme picarde *chiel*; les autres textes, sauf HermVal, choisissent les formes plus ‘neutres’ en *c-*. Cette distribution semble s’opposer à celle relevée par Dees pour l’Aisne, où le pourcentage de *ch-* est 0%.

#### 4.1.8. Substantif “*dieu*”<sup>62</sup>

	tot.	<i>diu</i>	<i>dix</i>	<i>dieu</i>	<i>diex</i>	<i>dex</i>	<i>deu</i>
<b>GenealTur</b>	0	0	0	0	0	0	0
<b>NoironsTur</b>	20	0	0	4 (20%)	16 (80%)	0	0
<b>SVou</b>	15	0	0	7 (46,67%)	8 (53,34%)	0	0
<b>MahomTur</b>	5	0	0	4 (80%)	1 (20%)	0	0
<b>Prolloh</b>	3	0	0	3 (100%)	0	0	0
<b>HermVal</b>	29	4 (13,79%)	8 (27,59%)	8 (27,59%)	9 (31,03%)	0	0
<b>SFaniel</b>	93	22 (23,66%)	41 (44,09%)	13 (13,98%)	15 (16,13%)	0	2 (2,15%)
<b>VenjNS</b>	115	53 (46,09%)	53 (46,09%)	5 (4,35%)	4 (3,48%)	0	0
<b>Hervis</b> <sup>63</sup>	383	277 (72,32%)	90 (23,50%)	2 (0,52%)	0	3 (0,78%)	11 (2,87%)

FEW 3,57a, s.v. DEUS: «nam. *dié*, *diu*, Tourc. *di*, St-Omer *diu*, pic. *dyü*, *guiü*»

DeesAtlas, c. 144 *-ieu-*, *-iu-* ~ *-e-*, *-ei-*, *-eu-* subst. “dieu”: Somme, Pas-de-Calais 97,67%; Oise 97,22%; Aisne 88,24%; Nord 83,33%; Hainaut 64,29%; Wallonie 44,85%; Ardennes 66,67%; Marne 40,91%

DeesAtlas, c. 144 *-iu-* ~ *-ieu-* subst. “dieu”: Somme, Pas-de-Calais 23,81%; Oise 2,86%; Aisne 6,67%; Nord 66,67%; Hainaut 14,81%; Wallonie 13,64%; Ardennes 0%; Marne -

60 Les deux occurrences de *chius* au pluriel se suivent dans le même passage au f. 35va; la première occurrence est à la rime en fin de vers: quoique le dernier mot du deuxième vers soit difficilement lisible, le résultat de la transcription donne: «il les emporterent es chius | a la terre remest li drus».

61 Dont une occ. de *cils*.

62 Cf. GodinM, p. L: «*dëus* aboutit le plus souvent à *diu(s)* [...] parfois on trouve *dieu(s)*» dans la partie rimée, tandis qu’on trouve 2x *dius* contre 2x *dé(s)* dans la partie assonancée. Cf. Greub, Collet, *Variation* cit., § 1.12 «Le groupe [jø] passe dans le Nord-Est (picard, wallon) à [jy] dans plusieurs positions [...] l’usage de la graphie <iu> correspond à des préférences graphiques, qui ne sont cependant attestées qu’à l’intérieur de la Picardie et de la Wallonie».

63 HervisH, p. LVIII, relève 440x *deus* (dans la partie en vers) et 37x de *Dé*, *Dex*, *Dez* à l’assonance, chiffres qui se distinguent assez nettement de notre relevé pour le ms. de Turin.

DeesAtlas<sup>2</sup>, c. 177 *diu* (*dix*, *dis*) ~ *dieu* (*diex*) subst “dieu”: Somme, Pas-de-Calais 29,17%; Oise 22,06%; Aisne 12,18%; Nord 80%; Hainaut 5,37%; Wallonie 0%; Ardennes 0%; Marne 0%

DeesAtlas<sup>2</sup>, c. 178 *de*, *deu* ~ *die*, *dieu*, *diu* subst “dieu”: Somme, Pas-de-Calais 9,41%; Oise 1,39%; Aisne 6,16%; Nord 26,56%; Hainaut 20,50%; Wallonie 75%; Ardennes 10,94%; Marne 31,93%

DeesAtlas<sup>2</sup>, c. 179 *de*, *dei* ~ *deu* subst “dieu” Somme, Pas-de-Calais 33,33%; Oise -; Aisne 100%; Nord 0%; Hainaut 0%; Wallonie 0%; Ardennes 0%; Marne 5,83%

GallRom: “ê[” (tot. 16’133x), formes de DĒUM (2563x): la forme *diu* (66x) est ainsi distribuée: 2x chOise, 39x chDo, 5x chFland, 2x chHain, 2x chMa; les formes *deu* (1707x) et *dieu* (718) sont réparties sur tous les corpus.

On remarque une bipartition nette entre les formes en *-ieu* typiques de *unica* et celles en *-iu* typiques des autres textes est très fréquentes dans l’aire picarde.

#### 4.1.9. Voyelle nasale dans “feme, temps”<sup>64</sup>

	tot. (% de a)	FEMINA			TEMPUS		
		tot. f.	<i>fame</i>	<i>feme</i>	tot. t.	<i>tans</i>	<i>tens</i>
<b>GenealTur</b>	0	0	0	0	0	0	0
<b>NoironsTur</b>	3 (100%)	2	2 (100%)	0	1	1 (100%)	0
<b>SVou</b>	1 (100%)	1	1 (100%)	0	0	0	0
<b>MahomTur</b>	2 (100%)	0	0	0	2	2 (100%)	0
<b>Prolloh</b>	6 (100%)	6	6 (100%)	0	0	0	0
<b>HermVal</b>	11 (36,36%)	7	0	7 (100%)	4	4 (100%)	0
<b>SFmanuel</b>	29 (65,52%)	11	1 (9,09%)	10 (90,90%)	18	18 (100%)	0
<b>VenjNS</b>	40 (52,50%)	19	0	19 (100%)	21	21 (100%)	0
<b>Hervis</b>	75 (30,67%)	53	1 (1,89%)	52 (98,11%)	22	22 (100%)	0

FEW 3,449a, s.v. FEMINA; 13/1,185a s.v. TĒMPUS

DeesAtlas, c. 148 *-a-* ~ *-e-* subst. “femme(s)”: Somme, Pas-de-Calais 19,12%; Oise 75,96%; Aisne 2,33%; Nord 1,53%; Hainaut 1,94%; Wallonie 7,84%; Ardennes 18,75%; Marne 75%

DeesAtlas, c. 195 *-an-*, *-am-* ~ *-en-*, *-em* subst “temps”: Somme, Pas-de-Calais 100%; Oise 90%; Aisne 90,91%; Nord 97,06%; Hainaut 100%; Wallonie 11,11%; Ardennes 40%; Marne 16,67%

DeesAtlas<sup>2</sup>, c. 188 *fa-* ~ *fe-* subst “femme”: Somme, Pas-de-Calais 6,01%; Oise 48,33%; Aisne 59,23%; Nord 0%; Hainaut 20,83%; Wallonie 2,22%; Ardennes 20%; Marne 70,75% (*fame* essentiellement absent du Nord-Est)

DeesAtlas<sup>2</sup>, c. 215 *tans* ~ *tens* subst “temps”: Somme, Pas-de-Calais 97,81%; Oise 74,18%; Aisne 67,67%; Nord 92,31%; Hainaut 89,41%; Wallonie 1,11%; Ardennes 59,65%; Marne 56,67%

GallRom: “en-an” (tot. 46’855x), formes de TEMPUS (1730x): la forme *tamps* est ainsi distribuée: 16x chDo, 7x chFland, 18x chHain, 43x chOise, mais elle est présente dans tous les corpus orientaux et dans la prévôté de Paris (jusqu’à 1292). Head “ê[” (tot. 11 511x), formes de *fame* (2355x) est minoritaire, voire absente dans les corpus du Nord-Est, tandis que *feme* (3012) y est fortement représentée.

Si la forme *tans* est la seule attestée dans l’ensemble du manuscrit, les formes de *feme* montrent une bipartition assez claire entre les *unica*, qui préfèrent la forme plus ‘neutre’ et non picarde *fame*, des autres textes.

#### 4.1.10. Voyelle et groupe consonantique dans “sembler”, “ensemble”<sup>65</sup>

64 Cf. GodinM, p. XLIX, qui relève de «très nombreuses graphies *an* pour *en* étymologique [...] *fame* 17719 (mais *femme* 8564, 19092)» (assez cohérent avec nos données) et de *tans* (qu’on imagine être une forme exclusive). Cf. Greub, Collet, *Variation* cit., § 1.15: «Il semble que devant [m], l’ouverture de [ê] en [ā] se soit imposée plus largement que devant [n] [*fame*, *tens*]».

	tot.	<i>sanler</i>	<i>ensanle</i>	<i>sambler</i>	<i>ensamble</i>	<i>sembler</i>
<b>GenealTur</b>	0	0	0	0	0	0
<b>NoironsTur</b>	1	0	0	1 (100%)	0	0
<b>SVou</b>	8	0	0	8 (100%)	0	0
<b>MahomTur</b>	4	0	0	3 (50%)	1 (50%)	0
<b>ProLoh</b>	1	0	0	1 (100%)	0	0
<b>HermVal</b>	11	3 (27,27%)	0	6 (54,55%)	2 (18,18%)	0
<b>SFaniel</b>	26	1 (3,85%)	0	9 (34,61%)	16 (61,54%)	0
<b>VenjNS</b>	41	29 (70,73%)	2 (4,88%)	6 (14,63%)	4 (9,76%)	0
<b>Hervis</b>	66	0	0	61 (92,42%)	4 (6,06%)	1 (1,52%)

FEW 11,623b, s.v. SĪMĪLARE; 4,716a s.v. ĪNSĪMUL; 25,541a s.v. ASSĪMŪLARE

DeesAtlas<sup>2</sup>, c. 382 *sam*, *san-* ~ *sem-*, *sen-* vb “semble” et composés: Somme, Pas-de-Calais 96,63%; Oise 64,41%; Aisne 69,38%; Nord 100%; Hainaut 100%; Wallonie 4,94%; Ardennes 95%; Marne 69,38%

DeesAtlas<sup>2</sup>, c. 382 *sent-*, *sanl-* ~ *sembl-*, *sambl-* vb “semble” et composés: Somme, Pas-de-Calais 25,89%; Oise 20%; Aisne 2,27%; Nord 64,29%; Hainaut 13,89%; Wallonie 5,56%; Ardennes 0%; Marne 2,50%

DeesAtlas<sup>2</sup>, c. 488 *ensemble*, *ensenle* ~ *ensemble*, *ensenle* adv “ensemble”: Somme, Pas-de-Calais 22,46%; Oise 25%; Aisne 0%; Nord 62,50%; Hainaut 0%; Wallonie 1,33%; Ardennes 0%; Marne 2,50%

DocLing: 23x de *-sanl-* dans 10x chDo, 2x chFland, 11x chHain; les formes *-sambl-* sont attestées dans presque tous les corpus mais très rare dans les corpus picards (sauf chOise)

La bipartition est assez claire, puisque les formes ‘picardes’ en *-anl-* sont absentes des *unica*, qui montrent la forme plus ‘neutre’ avec insertion de la consonne épenthétique *-sambl-*.

#### 4.1.11. Préposition “jusque”, “dusque”, “tresque”

	tot.	<i>jusque</i>	<i>dusque</i>	<i>tresque</i>
<b>GenealTur</b>	0	0	0	0
<b>NoironsTur</b>	4	0	4 (100%)	0
<b>SVou</b>	7	6 (85,71%)	0	1 (14,29%)
<b>MahomTur</b>	0	0	0	0
<b>ProLoh</b>	4	1 (25%)	3 (75%)	0
<b>HermVal</b>	1	0	1 (100%)	0
<b>SFaniel</b>	5	0	5 (100%)	0
<b>VenjNS</b>	22	0	22 (100%)	0
<b>Hervis</b>	133	2 (1,50%)	128 (96,24%)	3 (2,25%)

FEW 14,73a-b, s.v. USQUE: «Afr. mfr. *dusque* (Île-de-France, wall. pic. norm. agn. champ. XII<sup>e</sup>-XIII<sup>e</sup> s.)»; FEW 13/2,198a, s.v. TRANS

DeesAtlas<sup>2</sup>, c. 491 *dusque* ~ *jusque*, *jesque*, *trosque*, *tresque* “jusque”: Somme, Pas-de-Calais 57,57%; Oise 31,13%; Aisne 34,75%; Nord 50%; Hainaut 33,33%; Wallonie 0%; Ardennes 30,71%; Marne 51,56%

DocLing: forme *dusque* (tot. 397x): 319x chDo, 24x chFland, 5x chHain, 25x chOise, 5x docNorm; *tresque* 9x chMe, 2x chOise

La bipartition est ici moins nette et le seul texte qui privilégie la forme neutre *jusque* est le SVou.

65 Cf. également GodinM, p. XLIX, qui indique uniquement des formes en *-ambl-*, *-anbl-* pour la partie rimée, en cohérence avec la section des *unica*; en revanche, GodinM, p. LXXI, signale la forme *asanlés* dans la partie assonancée. Cf. Greub, Collet, *Variation* cit., § 2.32: «Au Nord et à l’Est, l’épenthèse est limitée aux groupes [m’r], [s’r] et [λ’r], et ne se rencontre donc pas dans les groupes [m’l], [l’r], [n’r]». Ajoutons que *sanler* et *ensale* se montrent comme typiques de Picardie et Wallonie.

#### 4.1.12. Préposition et adverbe “avec”

	tot.	avec	avecques	avoec	avoecques	avuec	avuecques
<b>GenealTur</b>	0	0	0	0	0	0	0
<b>NoironsTur</b>	5	0	0	4 (80%)	1 (100%)	0	0
<b>SVou</b>	3	0	0	1 (33,33%)	2 (66,67%)	0	0
<b>MahomTur</b>	0	0	0	0	0	0	0
<b>ProLoh</b>	1	0	0	1 (100%)	0	0	0
<b>HermVal</b>	0	0	0	0	0	0	0
<b>SFaniel</b>	7	3 (42,86%)	1 (14,29%)	1 (14,29%)	0	2 (28,57%)	0
<b>VenjNS</b>	14	1 (7,14%)	1 (7,14%)	0	0	11 (78,58%)	1 (7,14%)
<b>Hervis<sup>66</sup></b>	62	8 (12,90%)	0	50 (80,65%)	4 (6,45%)	0	0

FEW 24,30a, s.v. AB HOC

DeesAtlas<sup>2</sup>, c. 472 *aveque, aveques, oveque, oveques* ~ *avec, avec* prép. “avec”: Somme, Pas-de-Calais 8,35%; Oise 29,17%; Aisne 14,88%; Nord 16,67%; Hainaut 0%; Wallonie 0%; Ardennes 0,56%; Marne 14,97% (les pourcentages sont assez bas dans tout le domaine oïlique et ils ne vont jamais au-dessous de 36%: ce paramètre semble peu significatif)

DeesAtlas<sup>2</sup>, c. 475 *avec, avec* ~ *avuec, avoec, aveuc, ovuec, ovoec, oveuc* prép. “avec”: Somme, Pas-de-Calais 17,16%; Oise 20%; Aisne 59,55%; Nord 0%; Hainaut 0%; Wallonie 57,17%; Ardennes 18%; Marne 73,36%

DeesAtlas<sup>2</sup>, c. 477 *avoec, ovoec* ~ *avuec, avec, aveuc, oveuc, avec, oveuc* prép. “avec”: Somme, Pas-de-Calais 69,88%; Oise 60%; Aisne 29,09%; Nord 100%; Hainaut 66,67%; Wallonie 0%; Ardennes 40%; Marne 15,18%

DocLing, forme *avoec* (tot. 177): 127x chDo, 15x chFland, 10x chHain, 7x chOise

TEAF, s.v. *avuec*: la forme *avoec* est fréquente dans les textes picards (et agn.)

On remarque ici une certaine bipartition entre les formes en *-oe-* des *unica* (avec Hervis) et les autres textes; la VenjNS se distingue ultérieurement pour le choix de la forme diphtonguée *avuec*.

#### 4.1.13. Adverbe “illec”

	tot.	illec	alec	illuec	Illueques
<b>GenealTur</b>	0	0	0	0	0
<b>NoironsTur</b>	0	0	0	0	0
<b>SVou</b>	2	1 (50%)	1 (50%)	0	0
<b>MahomTur</b>	1	1 (100%)	0	0	0
<b>ProLoh</b>	3	0	2 (66,67%)	0	1 (33,33%)
<b>HermVal</b>	3	0	0	3 (100%)	0
<b>SFaniel</b>	10	0	0	7 (70%)	3 (30%)
<b>VenjNS</b>	21	0	0	14 (66,67%)	7 (33,33%)
<b>HervisH</b>	19	0	0	13 (68,42%)	6 (31,58%)

FEW 4,559a, s.v. İLLÖC

DeesAtlas<sup>2</sup>, c. 491 *illuec* ~ *ilueque* adv. “illec”: Somme, Pas-de-Calais 18,38%; Oise 56,67%; Aisne 11,95%; Nord 0%; Hainaut 0%; Wallonie 20,45%; Ardennes 17,14%; Marne 19,38%

DeesAtlas<sup>2</sup>, c. 493 *ilec* ~ *illuec, iloec, iloc, ileuc* adv. “illec”: Somme, Pas-de-Calais 4,28%; Oise 0%; Aisne 53,82%; Nord 0%; Hainaut 0%; Wallonie 25%; Ardennes 0%; Marne 28,57%

TEAF, s.v. *aluec*: la forme est principalement attestée dans des textes picards, parmi lesquels un document de Saint-Quentin de 1230 (cité par Gdf); AtreW (Ouest mil. XIII<sup>e</sup> s.) est dans un ms. hennuyer du 1<sup>er</sup> tiers du XIV<sup>e</sup> s.

<sup>66</sup> HervisH, p. XLVII, relève 4x *avoc* «à côté de *avec/awec*», ms. de base BnF, fr. 19160 (lorr. 2<sup>e</sup> tiers du XIII<sup>e</sup> s.).

Malgré le faible nombre d’occurrences, on remarque ici une bipartition assez claire entre les formes diphtonguées *-ue-* des autres textes et la forme réduite *-e-* des *unica*.

#### 4.2. Traits à faible pertinence

##### 4.2.1. Substantif “peuple”<sup>67</sup>

	tot.	<i>pueple</i>	<i>peul(le)</i>	<i>pul(le)</i>
<b>GenealTur</b>		0	0	0
<b>NoironsTur</b>	2	0	2 (100%)	0
<b>SVou</b>	1	0	1 (100%)	0
<b>MahomTur</b>	2	0	2 (100%)	0
<b>ProllLoh</b>	1	0	1 (100%)	0
<b>HermVal</b>	1	1 (100%)	0	0
<b>SFmanuel</b>	7	0	0	7 (100%)
<b>VenjNS</b>	5	0	4 (80%)	1 (20%)
<b>Hervis</b>	4	1 (25%)	2 (50%)	1 (25%)

FEW 9,187a, s.v. PÖPŪLUS: «*pule* (surtout wallon pic. ca 1180-XIV<sup>e</sup> s.), *peule* (pic. XIII<sup>e</sup> s.)»

DocLing: 6x *pueple* dans la chOise 162 (pic. mérid. 1281); 5x *peulle* chDo

TEAF, s.v. *pueple*: forme *peule* uniquement picarde et wallonne; forme *pulle* uniquement picarde

Les *unica*, avec SFmanuel et VenjNS, montrent une préférence pour la forme “picarde” diphtonguée *peulle*, contre la forme réduite *pulle*, picarde et wallonne.

##### 4.2.2. Substantif “lieu”<sup>68</sup>

	tot.	<i>lieu</i>	<i>leus</i>	<i>lius</i>
<b>GenealTur</b>	0	0	0	0
<b>NoironsTur</b>	1	1 (100%)	0	0
<b>SVou</b>	0	0	0	0
<b>MahomTur</b>	0	0	0	0
<b>ProllLoh</b>	0	0	0	0
<b>HermVal</b>	0	0	0	0
<b>SFmanuel</b>	3	0	0	3 (100%)
<b>VenjNS</b>	13	0	0	13 (100%)
<b>Hervis<sup>69</sup></b>	7	0	1 (14,29%)	6 (85,71%)

FEW 5,391b, s.v. LÖCUS

DeesAtlas, c. 168 *-ieu-*, *-iu-* ~ *-eu-* subst. “lieu”: Somme, Pas-de-Calais 98,12%; Oise 94,94%; Aisne 96,15%; Nord 98,68%; Hainaut 95,92%; Wallonie 96,67%; Ardennes 80%; Marne 29,41%

67 AubS, p. XVI, signale 3 occurrences de la forme *peule* (qu’il affirme être « caractéristique du wallon quoiqu’elle apparaisse en Picardie »), cohérent avec la section des *unica*, bien que dans Aub on y trouve des attestations de la forme *peuple*. Cf. Greub, Collet, *Variation* cit., § 2.20: «Les bilabiales des groupes P’L et B’L peuvent s’amuir (ou évoluer vers une fricative) dans des textes picards et wallons» (Picardie *peule, pule*; Wallonie *pule, pole*).

68 Cf. GodinM, p. LII, qui ne relève que la forme *liu* (+ 8x) à côté d’une forme *lieus*. Cf. Greub, Collet, *Variation* cit., § 1.33: «*Liu* [...] est presque exclusif dans le nord de la Picardie, très présent en Wallonie et est aussi attesté dans le reste de la Picardie et les Ardennes, mais pas au-delà».

69 HervisH, p. XLIV, relève 15x *leu(s)*, ms. de base BnF, fr. 19160 (lorr., 2<sup>e</sup> tiers du XIII<sup>e</sup> s.).

DeesAtlas, c. 169 *-iu-* ~ *-ieu-* subst. “lieu”: Somme, Pas-de-Calais 23,77%; Oise 39,09%; Aisne 95,67%; Nord 90,62%; Hainaut 97,87%; Wallonie 74,71%; Ardennes 40,97%; Marne 0%

GallRom: “ǭ[ OCU” (tot. 3770x), formes de LÖCU (3405x): la forme *liu* (233x) est ainsi distribuée, 1x docNorm, 45x chOise, 56x chDo, 23x chFland, 58x chHain, 20x chMa; les formes *lieu* et *leu* sont attestées dans presque tous les corpus

TEAF, s.v. *lieu*: attestations de *liu* surtout pic. et agn.

La seule occurrence de *lieu* dans NoironsTur n’est pas suffisante pour tirer des conclusions plus convaincantes, mais elle s’oppose néanmoins aux attestations plus nombreuses de *liu*, forme picarde.

#### 4.2.3. Adjectif “bon” (et dérivés)<sup>70</sup>

		<i>bon</i>	<i>boin</i>
<b>GenealTur</b>	3	2 (66,67%)	1 (33,33%)
<b>NoironsTur</b>	4	3 (75%)	1 (25%)
<b>SVou</b>	5	5 (100%)	0
<b>MahomTur</b>	3	3 (100%)	0
<b>ProlLoh</b>	3	3 (100%)	0
<b>HermVal</b>	23	1 ( <i>bom</i> ) (4,35%)	22 (95,65%)
<b>SFaniel</b>	23	23 (100%)	0
<b>VenjNS</b>	76	69 (90,79%)	7 (9,21%)
<b>Hervis</b>	206	199 (90,60%)	7 (3,40%)

FEW 1,433b, s.v. BONUS: «anam. apic. alorr. *boin*»

GallRom: “ǭ[” (tot. 16871x), formes de BÖNU (4884x): la forme *boin* (244x) est attestée dans l’Est de la France, mais avec une fréquence particulière dans chDo (120x), chFland (18x), chHain (25x) (cf. aussi 19x chMe)

TEAF, s.v. *bon*: les formes *boin* sont principalement picardes et lorraines.

La forme typiquement picarde *boin* est systématique dans HermVal mais très minoritaire, voire absente, de l’ensemble des autres textes.

#### 4.2.4. Substantif “diable”

	tot.	<i>diable</i>	<i>deable</i>
<b>GenealTur</b>	0	0	0
<b>NoironsTur</b>	1	1 (100%)	0
<b>SVou</b>	0	0	0
<b>MahomTur</b>	0	0	0
<b>ProlLoh</b>	0	0	0
<b>HermVal</b>	7	0	7 (100%)
<b>SFaniel</b>	4	4 (100%)	0
<b>VenjNS</b>	7	7 (100%)	0
<b>Hervis</b>	5	5 (100%)	0

DeesAtlas<sup>2</sup>, f. 176 *diable* [*diales*, *diaules*] ~ *deable* [*deable*, *deale*] subst. “diable”: Somme, Pas-de-Calais 83,25%; Oise 0; Aisne 40%; Nord 100%; Hainaut 66,67%; Wallonie 87,90%; Ardennes 55,56%; Marne 34,90%

<sup>70</sup> Greub, Collet, *Variation* cit., § 1.35: «En Picardie, une tradition graphique conduit à noter <boin> l’adjectif *bon*; cette graphie est rare ailleurs dans le domaine d’oïl [...] La forme *boin* est occasionnelle ou rare ailleurs».

TEAF, s.v. *diable*: la distribution des deux formes n'est pas claire.

La forme *deable*, typique du Nord-Est selon Dees, est la seule repérée dans HermVal, qui s'oppose ainsi aux autres textes.

#### 4.2.5. Voyelle dans le verbe "acheter" (et dérivés)

	tot.	ac(h)et-	ac(h)at-
<b>GenealTur</b>	0	0	0
<b>NoironsTur</b>	6	2 (33,33%)	4 (66,67%)
<b>SVou</b>	1	0	1 (100%)
<b>MahomTur</b>	0	0	0
<b>ProLoh</b>	0	0	0
<b>HermVal</b>	3	0	3 (100%)
<b>SFaniel</b>	2	0	2 (100%)
<b>VenjNS</b>	6	0	6 (100%)
<b>Hervis</b>	48	5 (10,42%)	43 (89,58%)

FEW 24,66a, s.v. \*ACCAPTARE

DeesAtlas<sup>2</sup>, c. 243 *achet-* ~ *achat-* verbe "acheter" et composés: Somme, Pas-de-Calais 0%; Oise 81,82%; Aisne 46,67%; Nord 0%; Hainaut 0%; Wallonie 4,17%; Ardennes 50%; Marne 41,67% (100% en frcomt. et 83% en orléanais, forme en -e- pratiquement absente du Nord)

DocLing: la distribution n'est pas aisée à cerner à partir des DocLing

Ce paramètre est peu significatif.

#### 4.2.6 Adjectif "meilleur"<sup>71</sup>

	tot.	millour	mildre	miudre	meillour
<b>GenealTur</b>	0	0	0	0	0
<b>NoironsTur</b>	1	1 (100%)	0	0	0
<b>SVou</b>	0	0	0	0	0
<b>MahomTur</b>	0	0	0	0	0
<b>ProLoh</b>	0	0	0	0	0
<b>HermVal</b>	1	0	0	0	1 (100%)
<b>SFaniel</b>	0	0	0	0	0
<b>VenjNS</b>	6	5 (83,33%)	1 (16,67%)	0	0
<b>Hervis<sup>72</sup></b>	19	11 (57,90%)	0	7 (36,84%)	1 (5,26)

FEW 6/1,665b, s.v. MĒLIOR: «aflandr. *meliour* BaudSeb, *millour* (ca 1325), apic. id. (fin XII<sup>e</sup> s.), *milleur*...»

DeesAtlas<sup>2</sup>, c. 116 *millour* ~ *meillor*, *moillor* adj. "meilleur": Somme, Pas-de-Calais 50%; Oise 40%; Aisne 25%; Nord 100%; Hainaut 0%; Wallonie 0%; Ardennes 100%; Marne 43,33%

DocLing: 13x chDo, 6x chHain, 1x chMa, 1x chMM, 2x chMo

Ce paramètre est peu significatif.

<sup>71</sup> Greub, Collet, *Variation* cit., § 1.14 : «*Mius* et *miudre* sont propres au picard, avec une concentration particulière dans les zones les plus nord-orientales»; cf. *milleur* et *millor* en Picardie, en Lorraine et en Champagne.

<sup>72</sup> HervisH, p. XLV, relève 6x de «*millor* ... à côté des formes aussi fréquentes *meillor*», ms. de base BnF, fr. 19160 (lorr., 2<sup>e</sup> tiers du XIII<sup>e</sup> s.).

#### 4.2.7 Verbe “somes”<sup>73</sup>

		<i>sons</i>	<i>soumes</i>	<i>seumes</i>	<i>somes</i>
<b>GenealTur</b>	0	0	0	0	0
<b>NoironsTur</b>	2	2 (100%)	0	0	0
<b>SVou</b>	3	1 (33,33%)	0	2 <sup>74</sup> (66,67%)	0
<b>MahomTur</b>	1	0	1 (100%)	0	0
<b>ProllLoh</b>	0	0	0	0	0
<b>HermVal</b>	5	0	0	0	5 (100%)
<b>SFmanuel</b>	2	0	0	0	2 (100%)
<b>VenjNS</b>	9	0	0	0	9 (100%)
<b>Hervis</b>	14	0	0	0	14 (100%)

DeesAtlas<sup>2</sup>, c. 298 *somes*, *sonmes* ~ *somes*

GallRom: “1ppl” (tot. 23128x), formes de SUMUS (747x): la forme *sons* est très rare (3x) et apparaît uniquement dans chMe, chMo et chV; les formes *soumes* (29x) et *sonmes* (16x) sont attestées dans 3x docNorm, 11x chOise, 10x chFland, 9x chHain; *seumes* est absent des DocLing.

On remarque une certaine bipartition entre les textes à tradition (où la forme “neutre” *somes* est systématique) et les *unica*, qui montrent une majeure variété de formes.

#### 4.2.8. Substantif “fils”<sup>75</sup>

	tot.	<i>fil</i>	<i>fix, fis</i>	<i>fiz</i>	<i>fils, filx</i>	<i>fius, fiux</i>	<i>fieus, fiex</i>
<b>GenealTur</b>	4	1 (25%)	3 (75%)	0	0	0	0
<b>NoironsTur</b>	7	3 (42,86%)	3 (42,86%)	0	0	0	1 (14,28%)
<b>SVou</b>	3	3 (100%)	0	0	0	0	0
<b>MahomTur</b>	5	1 (20%)	4 (80%)	0	0	0	0
<b>ProllLoh</b>	9	4 (44,44%)	5 (55,56%)	0	0	0	0
<b>HermVal</b>	10	7 (70%)	1 <sup>76</sup> (10%)	0	2 (20%)	0	0
<b>SFmanuel</b>	14	10 (71,43%)	4 (28,57%)	0	0	0	0
<b>VenjNS</b>	17	0	15 (88,24%)	0	2 (11,76%)	0	0
<b>Hervis</b>	168	54 (32,14%)	97 (57,74%)	6 (3,57%)	4 (2,38%)	6 (3,57%)	1 (0,60%)

FEW 3,521a, s.v. FILIUS

DeesAtlas, c. 150 présence de *-s*, *-z*, *-x* subst. cas sujet singulier “fils”: Somme, Pas-de-Calais 5,88%; Oise 0%; Aisne 0%; Nord 3,39%; Hainaut 0%; Wallonie 8,70%; Ardennes 2,78%; Marne 14,29%

DeesAtlas, c. 151 *-s*, *-z* ~ *-ls*, *-lz*, *-us*, *-uz* subst. “fils”: Somme, Pas-de-Calais 0%; Oise 0%; Aisne 0%; Nord 0%; Hainaut 0%; Wallonie 75,51%; Ardennes 32,38%; Marne 42,86%

DeesAtlas, c. 152 *-us*, *-uz*, *-x* ~ *-ls*, *-lz* subst “fils”: Somme, Pas-de-Calais 81,82%; Oise 100%; Aisne 93,78%; Nord 67,91%; Hainaut 67,31%; Wallonie 50%; Ardennes 63,64%; Marne 0

DeesAtlas<sup>2</sup>, c. 189 *fils*, *fius*, *fiz* ~ *fil* subst CRsg “fils”: Somme, Pas-de-Calais 3,90%; Oise 16,67%; Aisne 11,98%; Nord absent; Hainaut 0%; Wallonie 5,21%; Ardennes 4,17%; Marne 16,67%

<sup>73</sup> GodinM, p. LXXII, signale la présence de 2x *sons* à côté d’une attestation de *sonmes*.

<sup>74</sup> Une des deux attestations est pratiquement illisible mais on suppose qu’il s’agisse de *seumes*.

<sup>75</sup> Greub, Collet, *Variation* cit., § 1.25 indiquent qu’en on trouve Picardie uniquement *fius* et *fiex*, en Champagne *fiuz*.

<sup>76</sup> Une fois *fis* à la rime avec dans une lasse en *-is* (*fis*: *amis*: *espris* ...).

DeesAtlas<sup>2</sup>, c. 190 *fius, fieus ~ fils* subst “fils”: Somme, Pas-de-Calais 90%; Oise 100%; Aisne 50%; Nord 0; Hainaut 100%; Wallonie 0%; Ardennes 28,57%; Marne 36,99%  
GallRom: “fils” (tot. 2297x), la forme *fieu(l)s* est attestée dans les corpus picards (14x)

Ce paramètre est peu significatif.

#### 4.2.9 Substantif “ange”

La forme unique du ms. de Turin est *angle* (6x NoironTur, 1x HermVal, 7x SFanuel, 2x SVou, 2x MahomTur, 1x ProlLoh, 1x Hervis). La c. 150 de DeesAtlas2 *angle, angele ~ ange, angre* montre qu’il s’agit d’une forme picarde: Somme, Pas-de-Calais 100%, Aisne 100%, Hainaut 100%, Wallonie 87,5%, Ardennes 100%. Il manque donc la possibilité d’analyser la forme de manière contrastive.