



King's Research Portal

DOI:
[10.25430/2724-0975/2](https://doi.org/10.25430/2724-0975/2)

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Ravenhall, H. (2020). L'expérience *The Values of French Language and Literature in the European Middle Ages*. In *Autour du Roman de Florimont. Approches multidisciplinaires à la complexité textuelle médiévale* (Vol. 2, pp. 151–163). Quaderni di Francigena. <https://doi.org/10.25430/2724-0975/2>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Autour du Roman de Florimont.
Approches multidisciplinaires
à la complexité textuelle médiévale

Direzione / Editors-in-chief

GIOVANNI BORRIERO, Università degli Studi di Padova
FRANCESCA GAMBINO, Università degli Studi di Padova

Comitato scientifico / Advisory Board

CARLOS ALVAR, Universidad de Alcalá
ALVISE ANDREOSE, Università degli Studi e-Campus
FRANCESCO BORGHESI, The University of Sidney
FURIO BRUGNOLO, Università degli Studi di Padova
KEITH BUSBY, The University of Wisconsin
ROBERTA CAPELLI, Università degli Studi di Trento
DAN OCTAVIAN CEPRAGA, Università degli Studi di Padova
CATHERINE GAULLIER-BOUGASSAS, Université de Lille 3
SIMON GAUNT, King's College London
BERNHARD HUSS, Freie Universität Berlin, Germania
JOHN HAJEK, The University of Melbourne
MARCO INFURNA, Università Ca' Foscari - Venezia
GIOSUÈ LACHIN, Università degli Studi di Padova
STEPHEN P. MCCORMICK, Washington and Lee University
LUCA MORLINO, Università degli Studi di Trento
GIANFELICE PERON, Università degli Studi di Padova
LORENZO RENZI, Università degli Studi di Padova
ANDREA RIZZI, The University of Melbourne
RAYMUND WILHELM, Alpen-Adria-Universität Klagenfurt, Austria
ZENO VERLATO, Opera del Vocabolario Italiano, CNR
LESLIE ZARKER MORGAN, Loyola University Maryland

Redazione / Editorial Staff

ALESSANDRO BAMPA, Università degli Studi di Padova
CHIARA CAPPELLI, Università degli Studi di Padova
RACHELE FASSANELLI, Università degli Studi di Padova, Editor in chief
MARCO FRANCESCON, Università degli Studi di Trento
LUCA GATTI, Università degli Studi di Padova
FEDERICO GUARIGLIA, Università di Verona
FABIO SANGIOVANNI, Università degli Studi di Padova
BENEDETTA VISCIDI, Università degli Studi di Padova

*Francigena is an international peer-reviewed journal with an
accompanying monograph series entitled "Quaderni di Francigena"*

ISSN 2724-0975

Dipartimento di Studi Linguistici e Letterari
Via E. Vendramini, 13
35137 PADOVA

info@francigena-unipd.com

Autour du Roman de Florimont.
Approches multidisciplinaires
à la complexité textuelle médiévale

A cura di Marta Materni

This work is licensed under <http://creativecommons.org/licenses/by/4.0/>

Realizzazione grafica a cura di Arun Maltese (biblioteca.bear@gmail.com)

Volume pubblicato nel 2020

DOI: 10.25430/2724-0975/2

L'EXPÉRIENCE *THE VALUES OF FRENCH LANGUAGE AND LITERATURE IN THE EUROPEAN MIDDLE AGES*

Henry Ravenhall

(King's College, London)

The Values of French Language and Literature in the European Middle Ages (désormais *TVOF*) est un projet financé par le Conseil Européen de la Recherche dans le cadre d'un Advanced Grant de cinq ans (2015-2020)¹. Abrité par King's College, Londres (KCL), et en étroite collaboration avec King's Digital Lab, le 'laboratoire numérique' de cette université, *TVOF* est en train de développer deux éditions numériques d'un texte important du treizième siècle: l'*Histoire ancienne jusqu'à César*. Tout en tenant compte d'un autre projet financé par le Arts and Humanities Research Council, *Medieval Francophone Literary Culture Outside France* (*MFLCOF*), lui aussi hébergé par KCL, *TVOF* a pour ambition non seulement de faire progresser nos connaissances à l'égard de ce texte fondateur de l'historiographie en langue vernaculaire, mais aussi d'entamer une réflexion sur les capacités des technologies numériques pour représenter la textualité médiévale².

L'*Histoire ancienne jusqu'à César* est une compilation qui raconte l'histoire universelle depuis la création biblique jusqu'au point où Jules César dirige ses armées dans le nord de la Gaule. On ignore qui est le compilateur du texte, même si la suggestion de Wauchier de Denain, proposée à partir des ressemblances stylistiques, semble fondée. Un indice biographique nous renvoie à Roger, châtelain de Lille (mort en 1230), qui est désigné dans le texte en tant que mécène de l'œuvre³. Datée d'environ 1214, l'*Histoire ancienne* se divise en onze sections selon le découpage de Paul Meyer. C'est ce dernier qui a donné le titre, l'*Histoire ancienne*, qui continue à s'utiliser au lieu des diverses appellations attestées par les manuscrits eux-mêmes (*Estories Rogier*, *Les Ansiennes Hystoires Rommaines*, *La Bible en françois*, *Le livre d'Orose*, *La Fleur des hystoires*, *Li livres dou commencement dou monde*, etc.). Selon le comptage d'Anne Rochebouet, il reste plus de 84 manuscrits de la première rédaction, 10 de la deuxième, et 4 de la troisième⁴. Le texte a connu un succès notable en dehors des limites modernes de la nation française, particulièrement dans les États latins d'Orient et en Italie. La tradition manuscrite, en raison de sa complexité, sa diversité et sa transmission 'internationale' ainsi que le luxe de ses enluminures, a attiré le regard intéressé des philologues, linguistes, littéraires et historiens de l'art.

¹ Sous la direction de Simon Gaunt (PI), l'équipe de projet consiste de Hannah Morcos, Maria Teresa Rachetta, Henry Ravenhall et Simone Ventura (N° de projet : 670726). Nous remercions le CER pour son soutien.

² Voir le site web de ce projet antérieur pour des notices des manuscrits de cinq traditions importantes du Moyen Âge, y compris celle de l'*Histoire ancienne* : <http://www.medievalfrancophone.ac.uk/> [cons. 25. III. 2020].

³ Cfr. Szkilnik 1986.

⁴ Cfr. Rochebouet 2016.

Deux manuscrits sont au cœur du projet : Paris BNF fr. 20125, considéré le témoin le plus complet de la première rédaction du texte, daté des années 1270 ou 1280 et confectionné soit dans le nord de la France, soit à Acre ; et London BL Royal D 20 I, témoin remarquable et original de la deuxième rédaction fabriquée vers 1340 à Naples auprès de la cour angevine de Robert d'Anjou⁵. La deuxième rédaction omet l'ouverture basée sur la Genèse et la section qui porte sur les conquêtes d'Alexandre le Grand en Orient, et y intègre un récit beaucoup plus développé sur la légende troyenne à partir de la cinquième mise en prose du *Roman de Troie* de Benoît de Sainte-Maure, dans laquelle ont été interpolées les *Héroïdes* d'Ovide en français (l'ouvrage ainsi constitué s'appelle *Prose 5*)⁶. Ces deux manuscrits-clefs ont voyagé à travers l'Europe francophone : le manuscrit parisien contient des annotations d'un lecteur catalan, alors que le manuscrit londonien a voyagé – selon l'hypothèse de Luca Barbieri – de l'Italie en Espagne avant d'être amené en France⁷. Le rôle de l'historiographie dans la transmission des valeurs à travers le temps et l'espace est donc un aspect central des recherches du projet, dont les intérêts se trouvent à l'intersection de l'identité, la langue et l'idéologie dans la création dite 'littéraire' au Moyen Âge⁸.

C'est également pour ces raisons que nous exploitons l'*Histoire ancienne jusqu'à César* comme laboratoire numérique. En nous penchant sur une tradition textuelle très complexe et étendue, nous sommes en position d'explorer les possibilités que nous offrent les nouvelles technologies par rapport à l'édition papier et aux bases de données plus traditionnelles. Dans la présente étude, je vais présenter trois outils numériques qui ont été développés dans le cadre du projet (*Text Viewer*, *Alignment*, *Lemming*) et je vais expliquer notre démarche opératoire pour que cela puisse servir comme modèle éventuel à d'autres projets numériques. Par conséquent, cela offrira une occasion de réfléchir – à partir de notre propre expérience – aux avantages de ces technologies et aux défis que nous, en tant que médiévistes, devons relever.

1. *Text Viewer*⁹

Sous la direction de Hannah Morcos, la filière '*digital editions*' de *TVOF* vise à fournir les premières éditions intégrales (et librement consultables en ligne) des première et deuxième rédactions de l'*Histoire ancienne jusqu'à César*, qui n'est jusqu'à présent que partiellement éditée d'après des sections individuelles du texte¹⁰. Nous avons choisi d'éditer deux manuscrits-clefs

⁵ Cfr. Gaunt 2016.

⁶ Pour notre édition de *Prose 5*, nous profitons de l'expertise de Luca Barbieri, qui est responsable de la version 'interprétative' de cette section de la deuxième rédaction.

⁷ Cfr. Barbieri 2005, pp. 10-12.

⁸ Les objectifs du projet sont répartis en quatre filières : 1) éditions numériques (Hannah Morcos), 2) nature du français (Simone Ventura), 3) histoire et littérature (Maria Teresa Rchetta et Henry Ravenhall), 4) identité européenne (Simon Gaunt). Pour en savoir plus, voir <https://tvof.ac.uk/about> [cons. 25. III. 2020].

⁹ <https://tvof.ac.uk/textviewer/> [cons. 25. III. 2020].

¹⁰ Pour une liste d'éditions partielles de l'*Histoire ancienne*, voir <https://tvof.ac.uk/k/bibliography/> [cons. 25. III. 2020].

de la tradition de l'*Histoire ancienne* (première et deuxième rédaction) : Paris BNF fr. 20125 et London BL Royal D 20 I. Il est, par conséquent, plus approprié de parler de l'édition de deux manuscrits. Nous mettons donc deux versions de chaque manuscrit à la disposition de la communauté scientifique : une version dite 'semi-diplomatique' et une version dite 'interprétative'. En ce qui concerne la première (l'édition semi-diplomatique) nous retenons les graphies, la division des mots, la ponctuation et la segmentation macro-textuelle des manuscrits. Pour la deuxième (l'édition interprétative), nous avons introduit la division des mots et la ponctuation modernes. Nous n'avons corrigé les manuscrits que dans les cas de lacunes et de fautes de scribe évidentes, notre but étant de pourvoir les lecteurs non spécialistes d'un texte aisément lisible. Les notes contiennent l'indication des sources ainsi que la référence à la leçon de nos manuscrits de contrôle lorsque ces derniers remplissent les lacunes de nos manuscrits. De cette façon, ces deux versions permettent aux utilisateurs d'accéder au texte qui convient le mieux à leurs objectifs de recherche, qu'ils soient philologiques, littéraires, linguistiques, et ainsi de suite. Il est aussi possible de visionner les deux versions des deux manuscrits en parallèle pour en faciliter la comparaison (Fig. 1).



Figure 1. Le *Text Viewer* chez TVOF. Cfr. <http://www.tvof.ac.uk/textviewer/> [cons. 25.III.2020]

Comme tout projet d'édition numérique, nous suivons les principes établis par la TEI (*Text Encoding Initiative*), qui a pour objectif d'atteindre une homogénéité entre les transcriptions numériques de divers projets en format électronique. Lors des premières réunions de TVOF avec les informaticiens de King's Digital Lab, nous avons décidé quels éléments à incorporer dans notre encodage (*mark-up*) et comment organiser nos fichiers numériques au niveau structurel¹¹. Pour ce qui concerne la subdivision microstructurelle (i.e. au niveau de nos 'para-

¹¹ Nous exprimons ici nos très vifs remerciements à Paul Caton, Ginestra Ferraro et Geoffrey Noël (King's Digital Lab) dont le travail exceptionnel et les conseils lucides ont été absolument essentiels pour le déroulement du projet.

graphes' et 'phrases'), nous avons essayé de respecter l'usage de nos manuscrits, notamment la présence et distribution des lettrines rehaussées de rouge, la ponctuation médiévale, etc. Nous avons accordé un identifiant XML à chaque unité de texte, par exemple :

a. *pour introduire une section*¹²

```
<milestone xml:id="tvof-section-1" unit="section" n="1" type="Genesis">
```

b. *pour introduire un paragraphe*

```
<div type="1" xml:id="edfr20125_00003" n="5ra">
```

c. *pour introduire une phrase*

```
<seg type="2" xml:id="edfr20125_00003_01">
```

L'identifiant contient une référence au manuscrit (edfr20125), un numéro de paragraphe (00003) et un numéro de phrase (01). '5ra' se réfère évidemment au folio du manuscrit, alors que 'seg type' oblige l'encodeur à donner une indication syntaxique concernant la présence/absence de subordination dans la proposition¹³. Pour le texte lui-même, nous avons également choisi d'inclure une sélection d'aspects discursifs, lexiques et paratextuels dans notre balisage du texte, à savoir :

rubrique

```
<head type="rubric">
```

ponctuation

```
<pc rend="1"> (punctus)
```

```
<pc rend="2"> (punctus elevatus)
```

```
<pc rend="3"> (punctus interrogativus)
```

discours rapporté (direct/indirect)

```
<said direct="true">
```

```
<said direct="false">
```

nom propre

```
<geogName> (topographie/géographie)
```

```
<persName> (personnage)
```

```
<placeName> (lieu)
```

¹² Il faut constater ici que les sections ne sont pas du tout homogènes dans les manuscrits, et chaque groupe a tendance à mettre certains passages en relief (par exemple, la partie du texte sur Ninus, vers la fin de la première section biblique, est souvent traitée comme une section à part).

¹³ Ce type de balisage syntaxique naît d'un compromis entre deux de nos objectifs : l'édition numérique du texte, d'un côté, et la description de l'architecture discursive et syntaxique de la prose de l'*Histoire ancienne*, de l'autre. Pour des réflexions sur l'utilité des outils numériques à une telle analyse linguistique, voir Ventura 2019.

substantif désignant un peuple

<name type="people">

substantif désignant un édifice

<name type="building">

chiffre

<num>

lexique de langue étrangère

<foreign>

référence métalinguistique

<rs type="metaling">

référence métatextuelle

<rs type="metatext">

glose

<gloss>

intervention du narrateur

<q type="NI">

citation

<quote source="X" xml:lang="Latin">

intervention du scribe (S) ou éditeur (E) médiéval

<add place="left_margin" hand="S"> (addition, marge gauche)

<add type="overwritten" hand="E"> (réécriture)

<del rend="strikethrough" hand="S"> (texte barré)

répétition du mot

<choice>

 <sic>*ne par sa force ne par sa force*</sic>

 <corr>*ne par sa force*</corr>

</choice>

Pour des raisons d'espace, je ne peux pas fournir tous les détails concernant les critères pour l'établissement de chacun de ces aspects¹⁴. Dans tous les cas, ce balisage facilitera d'un côté la

¹⁴ Il est évident que la définition de quelques-unes de nos catégories demeure controversée. Par exemple, ce qui constitue une intervention du narrateur pourrait dépendre du sens inféré par l'encodeur. Pour des renseignements ultérieurs, je renvoie à notre guide TEI qui sera mis en ligne lorsque les éditions seront finalisées.

création d'un glossaire et d'un index de noms propres de personne et de lieu et de l'autre côté des recherches plus pointues dans les domaines de la linguistique du discours (discours rapporté), des études littéraires, de la ponctuation médiévale, et ainsi de suite. Vu de cette perspective, il est clair qu'il vaut mieux cerner les éléments du balisage au stade de conception initiale du projet.

Passons à l'aspect pratique. Nous avons divisé le manuscrit en morceaux d'environ dix folios et chaque partie correspond à un fichier dans notre Dropbox. Ceci permet aux encodeurs de travailler sur le même manuscrit en même temps, pourvu qu'un seul encodeur soit responsable d'un fichier. De plus, cela a encouragé la participation de personnes hors de l'équipe de projet – puisqu'il est possible de travailler sur les fichiers (en consultant les numérisations disponibles en ligne) à distance – et nous avons profité de la contribution des étudiants de niveau supérieur, qui de leur part avaient été formés dans nos principes TEI, avec la transcription et le balisage de nos manuscrits¹⁵. Il va sans dire que les contrôles des données sont absolument nécessaires dans ce système collaboratif. Par conséquent, tous les fichiers ont été vérifiés par un membre de l'équipe de projet afin d'assurer l'homogénéité du balisage à travers l'édition. Toute

```

18 <body>
19 <milestone xml:id="tvof-section-6" unit="section" n="6" type="Eneas"/>
20 <div type="1" xml:id="edfr20125_00588" n="148rb">
21 <head type="rubric">Ci comence d<choice>
22 <orig/>
23 <reg>'</reg>
24 </choice><persName>eneas</persName> qui <choice>
25 <seg type="semi-dip">separti</seg>
26 <seg type="crit">se parti</seg>
27 </choice> de <placeName>troies</placeName> [et] coment il s<choice>
28 <orig/>
29 <reg>'</reg>
30 </choice>en ala en <placeName>itale</placeName><pc rend="5"/><anchor
31 type="sourceNote" corresp="#edfr20125_0588_peach"/></head>
32 <pb n="148v"/>
33 <cb n="a"/>
34 <ab>
35 <seg type="5" xml:id="edfr20125_00588_01"><q type="NI">Segnors [et] dames</q><choice>
36 <orig/>
37 <reg>,</reg>
38 </choice> quant <placeName>troies</placeName> la grande fu arse [et] destruite<choice>
39 <orig/>
40 <reg>,</reg>
41 </choice> ne mie encore tote<pc rend="1"/><choice>
42 <orig/>
43 <reg>,</reg>
44 </choice> mais ta[n]t que li <name type="people">gri<choice>
45 <orig>i</orig>
46 <reg>j</reg>
47 </choice>ois</name>

```

Figure 2. Extrait de notre encodage sur *Oxygen* : l'ouverture de la section sur Énée dans Paris BNF fr. 20125. <choice> introduit le texte propre à la version 'semi-diplomatique' (<seg type="semi-dip">) et à la version 'interprétative' (<seg type="crit">).

¹⁵ Nos remerciements vont à Gabrielle Imbert, Matthew Siôn Lampitt et Edward Mills.

modification dans nos fichiers XML est transmise automatiquement à un autre fichier ‘agrégé’ (*aggregated file*) où se trouve l’intégralité du texte de la version ‘semi-diplomatique’. Pour la version ‘interprétative’, une autre couche d’encodage (*shorthand*) s’ajoute au fichier converti (*converted file*) où sont actionnées, par exemple, des altérations des orthographes ou de la division des mots (Fig. 2). C’est à partir de ce fichier converti que notre système de lemmatisation doit fonctionner (*Lemming*, voir dessous). Ce processus de conversion automatique (*workflow*) se fait par intervalles de deux heures, ce qui aboutit à une constante mise à jour en ligne de notre texte.

La première étape de notre travail éditorial a été alors de transcrire le texte dans son intégralité, avant de commencer les sections de la version ‘interprétative’, qui apparaissent sur le site au fur et à mesure qu’elles sont prêtes. D’un côté, cette mise en ligne graduelle a pour but de solliciter des réponses critiques des utilisateurs du site en vue de la parution de la section suivante, et de l’autre côté, n’étant pas soumise aux mêmes contraintes de l’édition papier, la version numérique se lit ‘en temps réel’ et dès qu’elle est disponible. Pour les lecteurs qui le souhaitent, l’édition est téléchargeable et imprimable. Nos fichiers XML seront également mis à la disposition de la communauté scientifique.

2. *Alignment*¹⁶

Alignment (‘alignement’) est un outil numérique conçu pour donner une vue d’ensemble des éléments structurels de la tradition manuscrite de l’*Histoire ancienne*. *Alignment* indique la présence ou l’absence des paragraphes, des rubriques, des passages en vers ou des sections entières, ainsi que des informations codicologiques supplémentaires, dans cinq manuscrits de l’*Histoire ancienne jusqu’à César* :

1. London BL Add. 15268 (Acre, c. 1280),
2. London BL Add. 19669 (France du nord, c. 1270),
3. London BL Royal D 20 I (Naples, c. 1340),
4. Paris BNF fr. 17177 (Soissons, c. 1290),
5. Paris BNF fr. 20125 (Acre/France du nord, c. 1280).

Au cours des mois à venir, nous y ajouterons les données d’autres manuscrits importants de la tradition. Pour le moment, l’utilisateur peut comparer des sections (en cochant celles qui l’intéressent) dans les différents témoins pour avoir une meilleure idée de comment se forment les groupes textuels de l’*Histoire ancienne* (‘version longue’ *vs.* ‘version abrégée’, première *vs.* deuxième rédaction). Trois ‘visionnements’ sont actuellement possibles sur le site TVOF (par colonnes, par tableau, et par barres) et chacun correspond à des besoins particuliers de la part du chercheur. En cliquant sur l’icône *Read*, l’utilisateur sera également réorienté vers l’édition numérique de Paris BNF fr. 20125 (Fig. 3).

¹⁶ <https://tvof.ac.uk/histoire-ancienne/alignment> [cons. 25. III. 2020].

The Values of French / Alignment					
	Add 1526B	Add 1966B	Fr 17177	Fr 20125	Royal 20 D 1
Eneas					
fr20125_00588	Location: 129b Rubric: Coment enneas se parti detroies por aies en ytalie	Location: 92va Rubric: Ci coumence de enneas qui separti detroie [e]c[om]ment il san ala en ytale	Location: 74b Rubric: Ci co[m]mence de Eneas co[m]ment il se p[ar]t detroie [et] [co]mment il ene ala en ytale p[ar] la [co]m[un]ement a ses deus	Location: 148b Rubric: Ci comence denneas qui separti de troies [et] coment il san ala en italie	Location: 184ra Rubric: Ci come[n]ce de Eneas qui se parti detroies et ala en ytalie
fr20125_00589					
fr20125_00590					
fr20125_00591					
fr20125_00592					
fr20125_00593					
fr20125_00594	Location: 131va Rubric: Con[en]t enneas ust la cite de cartage q[ue] la royne dido faisoit encores faire Variation: Partial material lacuna	Location: 93vb Rubric: Que enneas ut pransierement lache decartage q[ue] dido faisoit faire	Location: 75b Rubric: Co[m]ment enneas ut p[ri]miers la cite decartage que dydo faisoit faire	Location: 150vb Rubric: Q[ue] enneas ut p[ri]miers la cite de cartage que la royne dido faisoit encores faire	Location: 185v
fr20125_00595					
fr20125_00596					

Figure 3. Capture d'écran tirée de <https://tvof.ac.uk/histoire-ancienne/alignment>, visionnement par colonnes. Le violet signale une lacune matérielle partielle, alors que l'orange indique une lacune matérielle complète.

L'encodage d'*Alignment* se fait à travers plusieurs fichiers qui sont alloués chacun à un manuscrit de l'*Histoire ancienne*. Par exemple, dans celui du manuscrit London BL Royal D 20 I, tous les paragraphes du manuscrit Paris BNF fr. 20125 – le témoin le plus complet de la tradition – sont associés à (ou alignés avec) leurs équivalents respectifs dans le manuscrit de Londres. Nous avons aussi noté le contenu de la rubrique (*rubric*), le numéro de folio (*location*), et des informations supplémentaires (*note*). Voilà un extrait de cet encodage :

```
<div type="alignment" xml:id="fr20125_00645">
  <ab corresp="#edfr20125_00645" type="ms_instance">
    <seg type="ms_name">Fr20125</seg>
    <seg type="rubric">Des assirieins qui de tot le mo[n]de orent la poeste [et] la
    segnorie.</seg>
    <seg type="location">177ra</seg>
    <seg type="note"/>
  </ab>
  <ab corresp="#edRoyal20D1_00560" type="ms_instance">
    <seg type="ms_name">Royal_20_D_1</seg>
    <seg type="rubric">Des assyriens qui de tout le monde orent la poueste.</seg>
    <seg type="location">213rb</seg>
    <seg type="note"/>
  </ab>
</div>
```

Comme équipe de recherche, nous avons déjà tiré des bénéfices remarquables de l'usage de cet outil numérique. Non seulement permet-il d'avoir accès à des vastes informations sur les manuscrits individuels, mais on peut aussi reconnaître l'absence ou le déplacement de para-

graphes, ce qui par la suite suggère des filiations très nettes entre certains témoins¹⁷. Tout cela nous offre un aperçu des processus d'abrègement et de réécriture qui caractérisent fortement l'*Histoire ancienne*. Étant donné que le nombre de manuscrits numérisés en ligne est nettement à la hausse, il nous semble que ce type d'instrument comparatif et de navigation dans les mailles de textes longs et de traditions touffues ne peut que rendre service aux chercheurs.

3. Lemming¹⁸

Lemming, notre lemmatiseur, est le fruit d'une collaboration avec Stephen Dörr et Marcus Husar du *Dictionnaire Étymologique de l'Ancien Français (DEAF)*. Ce logiciel développé par Marcus Husar nous permettra d'intégrer les données lexicales très riches de l'*Histoire ancienne* dans le *DEAF*. La lemmatisation n'étant possible qu'à l'aide de textes où les mots sont séparés selon les conventions lexicologiques modernes, nous travaillons ainsi sur nos versions 'interprétatives'. Nous avons évoqué plus haut le processus de conversion automatique (*workflow*) : celui-ci transforme le fichier converti en fichier 'tokénisé' (*tokenised file*) – c'est-à-dire un numéro s'assigne à chaque unité lexicale – pour que le logiciel *Lemming* puisse cerner les mots individuels (*key word in context*) à associer avec un lemme. Lorsque l'on lemmatise, on peut donc voir le mot à lemmatiser dans son contexte (les mots qui l'entourent), ce qui est essentiel à l'identification correcte du lemme (Fig. 4).

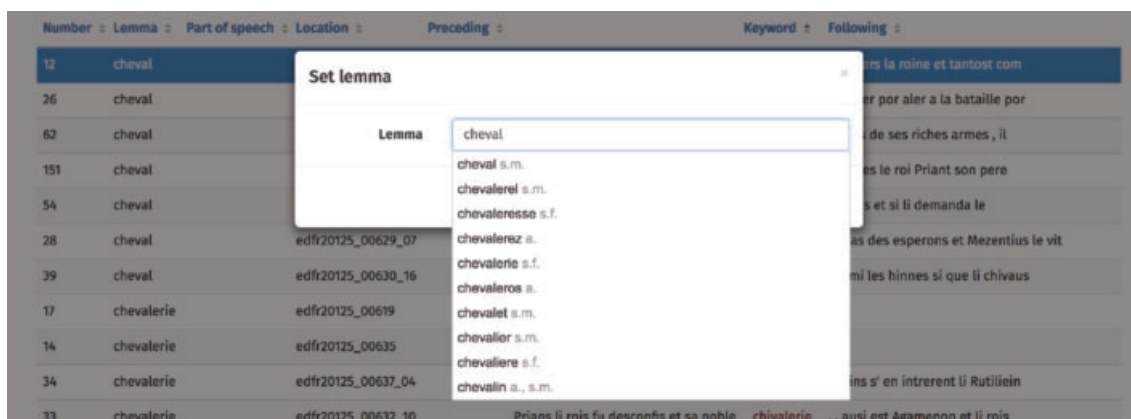


Figure 4. Le logiciel *Lemming*. Exemple d'entrée du lemme *cheval*.

Une page sur le site *TVOF* sous la rubrique 'Search' permettra aux utilisateurs de chercher soit par forme (le mot tel qu'il apparaît dans le manuscrit) soit par lemme (le mot tel qu'il apparaît dans le dictionnaire) (Fig. 5). La fréquence du terme y sera également indiquée, ainsi que la section et le manuscrit auxquels il appartient. En cliquant sur la forme, l'utilisateur sera

¹⁷ Voir, par exemple, Rachetta 2019.

¹⁸ *Lemming* est toujours en voie de développement.

rapporté à la phrase correspondante dans l'édition numérique. Nous envisageons la possibilité d'effectuer des recherches par les éléments discursifs, lexiques et paratextuels que nous avons inclus dans notre balisage du texte. Dans ce cas de figure, il serait possible, par exemple, de compiler une liste de toutes les occurrences de discours directs ou de substantifs désignant un peuple, ce qui faciliterait des analyses linguistiques, littéraires et historiques plus approfondies.

Figure 5. La page 'Search' sur le site prototype de TVOF, qui est maintenant disponible aux utilisateurs. <https://tvof.ac.uk/search/>? [cons. 25.III.2020]

Le *DEAF* dépouillera de manière exhaustive les données lexicales de l'*Histoire ancienne*, un texte qui présente sans doute plusieurs attestations originales, sinon très anciennes, des mots en ancien français. Mais TVOF mettra ses textes à la disposition d'autres dictionnaires pour que ces données puissent y être également intégrées.

4. Quelques défis?

Même si le potentiel de ces outils numériques est bien évident, ils s'accompagnent de défis à la fois logistiques et théoriques. Commençons par les défis logistiques. Tout d'abord, l'investissement financier est crucial parce qu'il faut collaborer avec les spécialistes des technologies numériques, ce qui est surtout le cas dans un environnement informatique où tout change à une vitesse exponentielle. Les soi-disant 'nouvelles' technologies deviennent démodées aussi vite qu'elles s'inventent. Si les technologies changent, les compétences requises pour opérer ces innovations doivent le faire aussi. Cependant, l'accent qu'il faut mettre sur les compétences numériques ne devrait pas réduire l'importance d'une formation plus traditionnelle. Les compétences dites 'dures' (philologiques et linguistiques, paléographiques et codicologiques) demeurent incontournables et devraient être encadrées dans le contexte du numérique : à titre d'exemple, la manière dont on transcrit un manuscrit selon un schéma philologique se baserait

sur comment baliser un manuscrit dans un encodage XML. Cela dit, alors que l'aspect collaboratif des projets numériques, comme nous l'avons vu, tend vers l'efficacité et l'optimisation des ressources, il est aussi vrai qu'il nécessite de contrôles de cohérence parce qu'avec plusieurs personnes travaillant sur un même fichier, des petites divergences présentes au début du projet pourraient en devenir des grandes, faisant partie d'un processus qui se nomme en anglais *mission creep* ('enlissement de mission').

Toutefois, il convient de constater qu'aucune représentation numérique de manuscrit médiéval n'est irréprochable. Quelques problèmes théoriques s'imposent lorsqu'on essaie de transférer l'objet, avec toutes ses idiosyncrasies et ses confusions, dans un nouveau contexte immatériel. Tout acte de transcription est d'abord un acte d'interprétation : la division des mots en est l'illustration parfaite. Selon le choix binaire que nous présente un logiciel d'encodage numérique (nous utilisons *Oxygen*, mais il y en a d'autres), il faut soit mettre un espace soit le supprimer. En réalité, cependant, il s'agit souvent d'un cas limite où, en tant qu'encodeurs, nous risquons d'y imposer nos propres conventions grammaticales ou linguistiques (par exemple, comme nous le savons, dans nos manuscrits les articles définis ou tout autre particule clitique tendent à se souder avec leur mot d'appui). Certaines lettrines, tant qu'elles sont exécutées dans le manuscrit, ne sont, à proprement parler, ni majuscules ni minuscules, et présentent donc une troisième catégorie qui brouille les possibilités fournies par le logiciel. Malgré des efforts concertés de notre part de capturer la nature palimpseste de nos deux manuscrits – en indiquant dans notre balisage la présence des additions effectuées par des mains ultérieures –, il y a toujours quelque chose qui se perd dans le transfert. Dans le manuscrit London BL Royal D 20 I, par exemple, un éditeur médiéval a 'corrigé' un grand nombre de graphies où 'i' se prononce /dʒ/ (*saie*→*sage*, *cartaie*→*cartage*, *ient*→*gent*, etc.) ou fonctionne comme pronom (*i auroit*→*y auroit*)¹⁹. Nous devrions alors favoriser soit la forme originale soit la forme 'corrigée' en dépit du fait que toutes les deux sont visibles simultanément au lecteur. Le manuscrit est un objet polytemporel, complexe, muable, et il est trop facile de lui accorder une structure génétique, fixe, rigide. Sans doute, les images nous aideraient à visualiser les particularités du manuscrit et nous rattacher à la matérialité du témoin, mais il est important de nous rappeler de la subjectivité inéluctable qui sous-tend tout effort de balisage.

5. Conclusion

C'est un moindre prix à payer, peut-être, pour autant de potentialités de recherche. Tous les fruits du projet ne sont pas, au moment d'écriture de cette contribution, encore récoltés. Mais les outils numériques que nous avons développés ont fourni une vue d'ensemble de l'*Histoire ancienne* qui n'avait pas été possible jusqu'à présent. La promesse d'*open access*, ainsi que les numérisations de manuscrits librement disponibles sur Internet, encouragent les utilisateurs

¹⁹ Voilà comment nous représentons une telle particularité dans notre encodage : `<mod hand="E"><add>g</add>i</mod>ent`. 'E' signale, bien entendu, le travail d'un éditeur médiéval.

à revenir aux sources, à comparer l'objet matériel avec son incarnation numérique. Une philologie 'transparente' pourrait alors s'en déduire : les éditions numériques sont capables de démontrer le montage et le *making of* en même temps, permettant ainsi une compréhension supérieure du texte et du manuscrit, ce qui aurait évidemment une valeur didactique incontestable, tout particulièrement avec une tradition aussi complexe que celle de l'*Histoire ancienne*.

Nul ne saurait douter que les outils numériques aient une place privilégiée dans le futur des études médiévales. Mais le problème de comment assurer la survie de ces outils – lorsque les financements seront écoulés et la technologie qui les soutient ne fonctionnera plus – est bien évident. Qui peut s'occuper de ces projets numériques dans la longue durée? Comment éviter qu'ils tombent dans l'oubli? Nous espérons que ce colloque déclenchera une discussion à ce propos qui s'avère de plus en plus nécessaire.

I. Manuscripts

London BL Add. 15268	London, British Library, Additional 15268
London BL Add. 19669	London, British Library, Additional 19699
London BL Royal D 20 I	London, British Library, Royal D 20 I
Paris BNF fr. 17177	Paris, Bibliothèque Nationale de France, français 17177
Paris BNF fr. 20125	Paris, Bibliothèque Nationale de France, français 20125

II. Bibliographie

Barbieri 2005

Luca Barbieri, *Le Epistole delle dame di Grecia nel Roman de Troie in prosa: la prima traduzione francese delle Eroïdi di Ovidio*, Tübingen – Basel, Francke, 2005.

Gaunt 2016

Simon Gaunt, *Philology and the Global Middle Ages: British Library Royal MS 20 D 1*, dans « *Medioevo Romanzo* », 40 (2016), pp. 27-47.

Rachetta 2019

Rachetta, *Sull'Historie ancienne jusqu'à César: le origini della versione abbreviata e il codice Wien ÖNB cod. 2576. Per la storia di una tradizione*, dans « *Francigena* », 5 (2019), pp. 27-57.

Rochebouet 2016

Anne Rochebouet, *De la Terre Sainte au Val de Loire : diffusion et remaniement de l'Historie ancienne jusqu'à César au XV^e siècle*, dans « *Romania* », 134 (2016), pp. 169-203.

Szkilnik 1986

Michelle Szkilnik, *Écrire en vers, écrire en prose. Le choix de Wauchier de Denain*, dans « Romania », 107 (1986), pp. 208-230.

Ventura 2019

Simone Ventura, *Digital Editing and Linguistic Analysis: The First Redaction of the Histoire ancienne jusqu'à César*, dans « Textual Cultures », 12 (2019), pp. 33-56.