



AMPLorr et TRANSSCRIPT : vers un corpus numérisé des actes médiévaux des princes lorrains et luxembourgeois

Christelle Balouzat-Loubet, Isabelle Guyot-Bachy

► To cite this version:

Christelle Balouzat-Loubet, Isabelle Guyot-Bachy. AMPLorr et TRANSSCRIPT : vers un corpus numérisé des actes médiévaux des princes lorrains et luxembourgeois. 2015. halshs-02907386

HAL Id: halshs-02907386

<https://halshs.archives-ouvertes.fr/halshs-02907386>

Preprint submitted on 27 Jul 2020

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.

AMPLorr et TRANSSCRIPT : vers un corpus numérisé des actes médiévaux des princes lorrains et luxembourgeois

Christelle Balouzat-Loubet (Université de Lorraine – CRULH)

Isabelle Guyot-Bachy (Université de Lorraine – CRULH)

L'étude des relations entre l'écrit et le pouvoir au Moyen Âge – champ de recherche au cœur de la médiévistique depuis les origines pour ses aspects diplomatiques et historiographiques – a connu récemment un regain d'intérêt par l'élargissement aux aspects mémoriels, identitaires et à l'étude des pratiques juridiques, administratives et de gestion. L'essor de la pratique de l'écrit aux derniers siècles du Moyen Âge est ainsi au centre de programmes de recherches récemment achevés ou en cours¹.

C'est dans ce champ d'études en plein renouvellement que s'inscrit AMPLorr (Actes médiévaux des Princes Lorrains : vers un corpus numérisé) mené depuis 2013 par l'Atelier diplomatique du Centre de Recherche Universitaire Lorrain d'Histoire.

La première partie de cette contribution voudrait tracer rapidement le contexte scientifique et de ce projet et en dessiner les objectifs ; dans une seconde partie seront exposés les choix techniques opérés, la méthodologie mise en œuvre et les perspectives de valorisation envisagées. Une troisième partie sera consacrée aux perspectives d'avenir et à l'élargissement de ce projet dans le cadre de l'ANR TRANSSCRIPT.

Le contexte scientifique d'AMPLorr

Le contexte scientifique d'AMPLorr est à situer entre passé, présent et avenir, au carrefour d'une réflexion méthodologique sur les pratiques de la numérisation et les enjeux suscités par le renouvellement des problématiques scientifiques.

Ce projet fait suite à l'opération EVADULOR (État Virtuel des Archives Ducales LORraines) conduite dans le cadre de la Maison des Sciences de l'Homme de Lorraine entre 2008 et 2013. Celle-ci avait pour objectif de proposer une reconstitution virtuelle du Trésor des chartes de Lorraine, dans l'état où la chancellerie de René II le laisse à la date de sa mort (1508). Cette reconstitution se fonde sur l'édition normalisée et structurée de l'inventaire analytique en 33 volumes des layettes du trésor des chartes, réalisé à l'intention du roi de France Louis XV sous la direction d'Antoine Lancelot, entre 1737 et 1740. Au terme de cette opération et de cinq années de financement, les premiers volumes sont désormais accessibles en ligne sur le site des archives départementales de Meurthe-et-Moselle². Mais l'outil, certes

¹ Sonderforschungsbereich 231 (1986-1999): Träger, Felder, Formen pragmatischer Schriftlichkeit im Mittelalter ; sous-projet D7 *Westmitteldeutsche und ostfranzösische Urkunden- und Literatursprachen im 13. und 14. Jahrhundert* du SFB 235 *Zwischen Maas und Rhein: Beziehungen, Begegnungen und Konflikte in einem europäischen Kernraum von der Spätantike bis zum 19. Jahrhundert* ; Centre de recherche *Pratiques médiévales de l'écrit*, Université de Namur ; GDR 3177 *Diplomatique* du CNRS ; *Cartul@rium* ; *L'écrit comme instrument de pouvoir au Moyen Âge* (32^e journée du RMBLF, 28 mai 2015).

² <http://ressourcesenligne.archives.cg54.fr/archives/classification-scheme>.

précieux, n'offre au chercheur que des analyses des documents réalisées au XVIII^e siècle, non les actes médiévaux eux-mêmes. Par ailleurs, le gigantisme de l'entreprise, amenée à s'étendre encore sur plusieurs années, a montré la nécessité d'adopter une méthode pragmatique, abordant directement le corpus médiéval mais en le fragmentant par périodes pour permettre une mise à disposition plus rapide et une exploitation efficace de la documentation.

Pour AMPLorr, le choix a donc été fait de fractionner le travail par règnes successifs des ducs de Lorraine et, très vite, Ferry III (1251-1303) s'est imposé comme le *terminus a quo* de l'étude : d'une durée exceptionnelle, c'est un temps et un champ d'observation privilégiés de l'essor quantitatif de l'écrit administratif et gestionnaire, de sa formalisation mais aussi de son ouverture aux transferts culturels, comme le colloque de 2007 sur le *Tournoi de Chauvency* l'avait déjà montré³. Or, malgré la richesse des sources disponibles, ce règne était demeuré largement ignoré des travaux d'érudition puis des programmes de recherches de la médiévistique⁴. Enfin, l'étude du règne de Ferry III constituait un angle d'attaque intéressant pour participer au renouvellement des problématiques d'une histoire régionale, fortement marquée et depuis longtemps par une historiographie qui replaçait systématiquement le duché de Lorraine dans une situation d'entre-deux, entre France et Empire, pensant son histoire comme principauté médiévale en termes identitaires modernes, c'est-à-dire à rebours, en partant des configurations et des intérêts de l'époque contemporaine.

Prétendre renouveler de manière significative la problématique et apporter des réponses précises autant qu'exhaustives nécessitait donc de repartir du corpus documentaire écrit, d'en revisiter les éléments mis à disposition par l'érudition depuis le XIX^e siècle, de le compléter et surtout de le rendre accessible aussi bien à la communauté scientifique (historiens, linguistes) qu'au grand public désireux d'accéder à sa dimension patrimoniale.

Un premier financement, accordé par l'Université de Lorraine et la Région pour l'année 2013, a permis de mener sur le catalogue des actes de Ferry III, réalisé au début du XX^e siècle par Jean de Pange⁵, trois opérations confiées et menées à bien par Kévin Echampard, chercheur contractuel au CRULH. La version numérisée de ce catalogue a été convertie en un fichier Excel, où ont été portés — et corrigés si besoin — les principaux éléments d'identification donnés par l'érudit pour chaque acte (cote, date, lieu de conservation, état original ou copie). Les difficultés subsistantes ont été signalées, afin de permettre des corrections ultérieures dans le cadre d'un travail collaboratif.

³ Mireille Chazan et Nancy Regalado Freeman (éd.), *Lettres, musique et société en Lorraine médiévale : autour du « Tournoi de Chauvency »* (Ms. Oxford Bodleian Douce 308), Genève, Suisse, Droz, 2012, vol. 1.

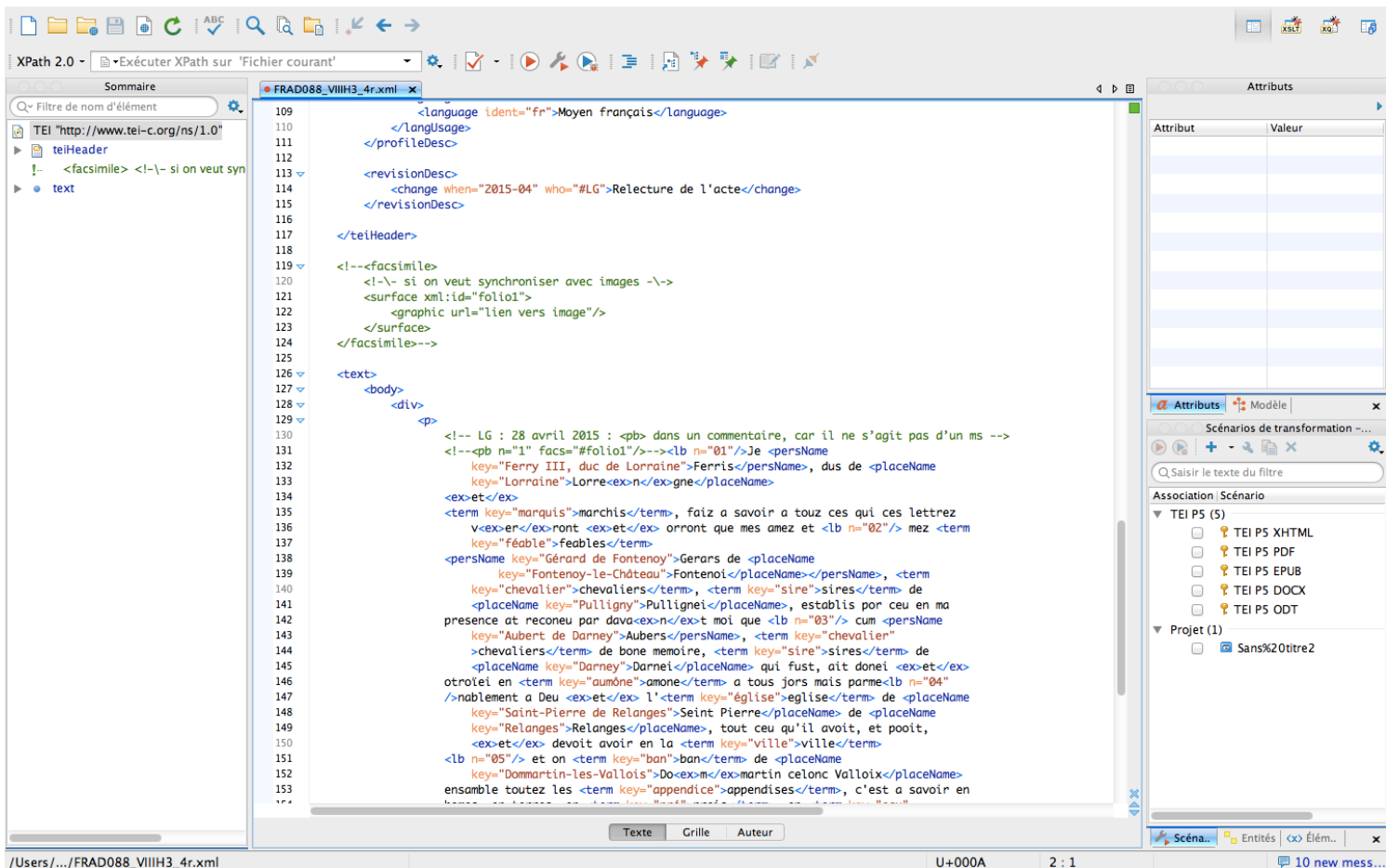
⁴ Une journée d'études consacrée au règne de ce duc a marqué le lancement du projet. Les contributions paraîtront dans les *Annales de l'Est* à la fin de la présente année.

⁵ Jean de Pange, *Introduction au catalogue des actes de Ferri III, duc de Lorraine (1251-1303)*, Paris, 1905 ; *Id.*, *Catalogue des actes de Ferri III, duc de Lorraine (1251-1303)*, Paris, 1930.

d’abord — la TEI a été fondée en 1987 et les premières *guidelines* ont été publiées en 1994 — est une preuve de sa pérennité. L’utilisation de syntaxes et de mécanismes XML standardisés par le W3C (World Wide Web Consortium) présentent par ailleurs de nombreuses garanties en termes d’interopérabilité et permettent d’utiliser les outils puissants disponibles pour manipuler des arbres XML. Les possibilités de personnalisation du modèle et de documentation du modèle grâce aux fichiers ODD, qui peuvent être ensuite convertis en RelaxNg, DTD ou XML Schema sont enfin un atout majeur pour le contrôle et l’uniformité de la saisie, ce qui permet de pallier le mouvement des personnels.

L’utilisation de la TEI ne prend pourtant tout son sens que si elle se place dans une réflexion scientifique globale. Avant d’encoder, il faut connaître et comprendre ses sources, effectuer des choix et définir des pratiques qui vont aboutir à l’élaboration d’un schéma qui servira de base et permettra de documenter le processus d’encodage. Ce schéma d’encodage, qui permet de le valider, est donc le témoignage de la réflexion scientifique qui doit se placer en amont. C’est une étape essentielle, qui consiste à sélectionner certains modules et éléments pour produire un sous-ensemble de la TEI (22 modules, 555 éléments possibles). Pour autant le schéma reste modifiable et peut être par la suite actualisé, au fur et à mesure que l’on affine ou renouvelle la problématique scientifique.

Dans le cas du projet présenté ici, le schéma élaboré pour l’encodage mêle éléments de description de la source et éléments sémantiques pour rendre au mieux la complexité du document.



Parmi les modules de TEI utilisés dans le cadre d'AMPLorr, il est possible de distinguer :

1. Les modules requis :

```
<moduleRef key="header"/>
<moduleRef key="core"/>
<moduleRef key="tei"/>
<moduleRef key="textstructure"/>
```

2. Les modules intégralement repris :

```
<moduleRef key="transcr"/> (Transcription of Primary Sources)
<moduleRef key="msdescription"/> (Manuscript Description)
<moduleRef key="namesdates"/> (Names, Dates, People, and Places)
```

3. Les modules dont seulement certains éléments ont été retenus :

```
<moduleRef key="textcrit"/> (Text Criticism)
<elementRef key="w"/> (Word)
<elementRef key="c"/> (Character)
<elementSpec ident="persName" module="namesdates" mode="change">
<attList>
<attDef ident="key" mode="replace">
<valList type="semi">
<valItem ident="Abert, curé de Saint-Epvre"/>
<valItem ident="Abrion de Denipaire"/>
```

Ce modèle pourra éventuellement être affiné au fur et à mesure du travail. La liste des entrées autorisées pour les noms propres, noms de personnes et noms de lieux est ainsi actualisée au fil de l'avancée du travail. Ce schéma a d'abord été élaboré dans une perspective linguistique, dans le cadre d'un dialogue noué avec l'équipe du *Dictionnaire du Moyen Français* (ATILF-CNRS), dialogue qui a abouti au versement de 452 textes (sur les 464 encodés à ce jour) dans la base Frantext de l'ATILF-CNRS⁶. Cette intégration dans Frantext enrichit avec profit le fonds existant des textes consultables, prolongeant en amont la chronologie du *Dictionnaire du Moyen Français* (1330-1500) et permet d'effectuer des recherches sur les formes et les lemmes présents dans le corpus. Très utile pour les analyses lexicométriques, cet outil est tout à fait pertinent pour les historiens, puisqu'il permet déjà d'interroger l'ensemble de ce corpus par les noms de personnes, de lieux ou de termes. Mais il ne permet pas d'associer les photographies des documents aux transcriptions correspondantes.

C'est pourquoi un autre partenariat a été développé avec ORTOLANG, un équipement d'excellence validé dans le cadre des investissements d'avenir⁷. 57 actes issus des archives départementales des Vosges devraient très prochainement être accessibles sur ce portail, qui, pour chaque document, associera photographie et transcription encodée, permettant de

⁶ <http://www.atilf.fr/dmf/ActesFerriIII/>

⁷ <https://www.ortolang.fr/#/market/home> C'est un lien sécurisé : n'est-ce pas imprudent de l'indiquer alors que nous n'avons pas encore l'autorisation de mise en ligne. Par ailleurs, le temps que ce colloque soit publié, nous aurons – j'espère – procédé à la mise en ligne et ce lien ne sera plus valable...

naviguer entre les différents documents. Le corpus sera également doté de trois index (termes, lieux et personnes) depuis lesquels il sera possible d'accéder aux textes.

Un laboratoire d'essai pour un projet plus ample (TRANSSCRIPT)

Le besoin de renouveler les problématiques, évoqué dans la première partie, n'est pas propre aux historiens de la Lorraine. Il est partagé par les chercheurs de tout l'espace correspondant au cadre territorial des principautés de l'ancienne Haute-Lotharingie, espace équivalant au cœur de l'actuelle Grande Région. Le dialogue engagé depuis deux ans avec les historiens médiévistes de l'Université du Luxembourg nous en a convaincus. Sur le plan méthodologique par exemple, nos collègues luxembourgeois, engagés dans l'opération EDITIO (édition « papier » des actes de Jean l'Aveugle)⁸, sentaient que la poursuite et la diffusion d'une telle entreprise requéraient aujourd'hui le moyen d'une édition en ligne, et rejoignaient ainsi les préoccupations de notre propre projet AMPLorr. C'est pourquoi nous avons choisi d'engager ensemble l'avenir autour d'un projet, TRANSSCRIPT, qui sera financé pour les quatre prochaines années conjointement par l'ANR (Agence Nationale de la Recherche - France) et le FNR (Fonds National de la Recherche – Luxembourg).

Le projet se situe au croisement de deux champs d'étude : d'une part le champ de la gouvernance des États naissants à l'échelon régional et d'autre part le domaine des transferts culturels entre régions situées de part et d'autre de la frontière linguistique. Pour explorer ces deux champs de recherche, le projet entend partir de la *scripturalité*, i.e. la production et l'usage des documents écrits, écrire et compter étant deux processus de la gouvernance des terres et des hommes. À travers leur étude l'historien peut prétendre approcher la réalité du pouvoir médiéval. Par leur aspect formel, écrire et compter sont perméables aux transferts culturels. En effet, loin d'être figés, les documents écrits, de quelque nature qu'ils soient, révèlent, à travers leur production et leur usage, par le biais du vocabulaire politique et juridique employé, à travers leur terminologie identitaire, les acteurs et les médiateurs (commanditaires, destinataires, témoins, scripteurs, messagers) d'une communication politique et sociale active et complexe dont ils portent le discours réel et idéologique. Sans omettre tous les éléments destinés à la visibilité de l'écrit (enluminures, initiales ornées, décor héraldique, symboles graphiques), qui participent aussi de la communication sociale et politique. Le projet permettra de mettre en œuvre une prosopographie des scribes qui, travaillant sur ces divers types d'écrits, participent à la porosité entre les différents types de documents ; une prosopographie des témoins pour éclairer la surface sociale et politique et le réseau d'alliances des commanditaires. Il contribuera aussi à la lexicométrie et rejoindra l'approche linguistique dans le contexte culturel du premier développement de l'écrit administratif et du passage progressif du latin aux langues vernaculaires.

Le corpus principal est celui des actes princiers du duché de Lorraine et du comté puis duché de Luxembourg, les actes princiers étant définis comme des actes (originaux et copies)

⁸ *Urkunden- und Quellenbuch zur Geschichte der altluxemburgischen Territorien*, begründet von Camille Wampach, Band XI: *Die Urkunden Graf Johans des Blinden (1310-1346)*, Teil 1: *Die Urkunden aus Luxemburger Archivbeständen*, hrg. v. Aloyse Estgen, Michel Pauly, Jean Schroeder, Luxembourg 1997, 155 S ; *Urkunden- und Quellenbuch zur Geschichte der altluxemburgischen Territorien*, begründet von Camille Wampach, Band XI: *Die Urkunden Graf Johans des Blinden (1310-1346)*, Teil 2: *Die Urkunden aus den Archives générales du Royaume in Brüssel*, hrg. v. Aloyse Estgen, Michel Pauly, Hérold Pettiau und Jean Schroeder, Luxembourg 2008.

dont le prince est le disposant. Afin de garantir la faisabilité du projet sur quatre ans, l'édition des actes princiers sera limitée pour la Lorraine à la période allant de 1251 à 1390, soit du règne de Ferry III à celui de Jean I^{er} ; pour le Luxembourg de 1247 à 1383 (du règne d'Henri V à celui de Wenceslas I^{er}). L'ensemble représente à ce jour 3276 actes. L'ampleur du corpus est heureusement compensée par la brièveté des actes, dont la plupart ont une longueur inférieure à la page. Les 758 actes des comtes/ducs luxembourgeois, déjà très largement édités, seront numérisés avant d'être encodés (pour les actes en français) ou lemmatisés (pour les actes en latin) ; les actes non édités seront systématiquement photographiés en vue de leur transcription mais aussi de leur future mise en ligne.

Sur le plan méthodologique, les verrous ne sont pas différents de ceux que nous avons détectés pour AMPLorr, mais des réponses se dessinent sur la base de l'expérience acquise dans le cadre de ce précédent projet :

1. Pour faire face à la dispersion extrême des fonds, nous allons procéder par mutualisation à commencer par la mise en commun des cotes, ce afin de limiter le nombre des missions de collecte et d'organiser les campagnes de photographies. À terme, cette mutualisation autour d'un projet commun devrait faciliter les négociations sur les droits de mise en ligne des photographies avec les institutions de conservation et favoriser la réalisation d'expositions physiques ou virtuelles en collaboration avec ces mêmes institutions.

2. Nos deux corpus sont traversés par la même diversité linguistique et une répartition à peu près similaire entre le moyen français, le latin et l'allemand, dont la progression sur la période est un des éléments significatifs des corpus. Cette diversité est fascinante par les possibilités d'enquêtes qu'elle offre sur les transferts culturels et la construction des langues. Mais son traitement informatique est une gageure. À partir des choix faits pour AMPLorr, les options de mise en ligne seront discutées en fonction de l'appartenance linguistique du document.

En effet, si les actes en français ont trouvé un hébergement et une première exploitation naturels grâce à Frantext, il reste à concevoir l'exploitation des documents en latin et en allemand. Diverses solutions sont envisagées. Celles qui permettront au chercheur le passage aisé d'un corpus linguistique à l'autre et les interrogations croisées seront privilégiées.

En cette fin d'année 2015, le projet AMPLorr, conçu comme une opération de transition, s'achève sur un bilan encourageant. Sur le plan de la connaissance historique, il a renouvelé de manière significative l'intérêt pour Ferry III, dont il apparaît que le règne fut un grand moment d'affirmation des pouvoirs princiers. Il a livré une matière abondante dont les chercheurs confirmés se sont emparés, mais aussi les jeunes chercheurs, comme en témoignent les premiers mémoires de master engagés à partir du corpus traité. Sur le plan méthodologique et technique, AMPLorr s'est efforcé de tirer toutes les leçons d'une expérience qui a marqué en son temps le paysage alors en formation des humanités numériques. La principale de ces leçons a sans doute été qu'il fallait lier étroitement le souci de livrer un corpus en ligne et les perspectives de son exploitation scientifique. Pour le dire brièvement, les partenariats avec les historiens luxembourgeois comme avec l'équipe de l'ATILF réunissent ceux qui seront les premiers utilisateurs du corpus en ligne. Les uns et les autres font le pari que le mûrissement conjoint des solutions techniques et des problématiques

scientifiques permettra de lever les verrous générés par la nature complexe du corpus.
L'avenir dira s'ils ont eu raison.