



UNIVERSITÉ PARIS-SORBONNE  
ÉCOLE DOCTORALE I : Mondes anciens et médiévaux  
Laboratoire « Études et édition de textes médiévaux » (EA 4349)

## THÈSE

pour obtenir le grade de  
DOCTEUR DE L'UNIVERSITÉ PARIS-SORBONNE

Discipline : études médiévales

Présentée et soutenue par :

Jean-Baptiste CAMPS

*maître ès lettres*

le : 3 décembre 2016

# *La Chanson d'Otinél*

Édition complète du corpus manuscrit et prolégomènes à  
l'édition critique

M. Dominique BOUTET — **Sous la direction de :**  
— Professeur à l'Université de Paris-Sorbonne

M<sup>me</sup> Maria COLOMBO TIMELLI — **Membres du jury :**  
— Professeur à l'Université de Paris-Sorbonne  
M. Giovanni PALUMBO — Professeur ordinaire à l'Université de Namur  
M. François SUARD — Professeur émérite à l'Université de Paris Ouest Nanterre  
M<sup>me</sup> Françoise VIELLIARD — Professeur émérite à l'École nationale des chartes

## AVANT-PROPOS

Se rattachant à la geste du roi, la *Chanson d'Otinel* n'avait pas été rééditée depuis le travail pionnier de F. Guessard et H. Michelant en 1858. Partant des objets tangibles que sont les manuscrits, pour aller vers l'étude de la tradition et de l'œuvre, ce travail se propose de réexaminer l'ensemble des données disponibles touchant à cette geste et à sa circulation, en vue de permettre la restauration d'une œuvre qui a connu une diffusion importante dans l'Europe médiévale, mais que nous ne conservons qu'en l'état de vestiges épars (manuscrits lacunaires ou abrégés, fragments, traductions médiévales, chroniques et versions dérivées).

La thèse prend un parti résolument méthodologique, en cherchant à faire bénéficier l'édition des progrès épistémologiques engendrés, d'une part, par les contributions les plus récentes aux débats propres à l'ecdotique et à la critique textuelle, et, d'autre part, par ce qu'il est convenu d'appeler les « humanités numériques ».

Du point de vue de la critique textuelle, l'édition tente ainsi de dépasser le débat entre « nouvelle » philologie (*New Philology*) et philologie traditionnelle, qu'elle repose sur la méthode des fautes communes ou soit d'inspiration bédieriste, pour se placer dans la perspective d'une « quatrième voie » et d'une édition « tournée vers la tradition », qui fournisse également un accès au système graphique des différents témoins. Pour ce faire, sont sollicitées les techniques de l'édition électronique et de la philologie numérique, afin de fournir une édition complète du corpus manuscrit, qui, par des transcriptions « à couches », donne accès à différentes représentations du texte. Dans le même temps, l'édition vise également à la description et l'analyse des liens que ces témoins entretiennent entre eux, en mettant en place une méthode de représentation de la variance textuelle et en cherchant à appuyer l'analyse généalogique sur une prise en compte globale de la tradition, incluant les traductions médiévales (galloises, norroises, anglaises) et l'ensemble des versions dérivées.

Le travail de modélisation et de description des manuscrits et de leurs textes, formalisé par un modèle XML/TEI conçu pour les besoins de cette édition mais se voulant de portée plus générale, est très nettement tourné vers l'exploitation des données, dans une perspective d'analyse quantitative, doublée d'une approche plus traditionnelle, touchant notamment aux champs de la paléographie quantitative, de la scriptométrie et de la stemmatologie. Tout au long du travail, des méthodes relevant de la modélisation mathématique, de la statistique, de l'algorithmique et de l'intelligence artificielle sont mises en œuvre. Des traitements visant à permettre l'interopérabilité, la montée en masse et la systématisation du travail éditorial sont également déployés, en particulier dans le domaine de la reconnaissance optique de caractères, l'annotation linguistique et la collation assistée par ordinateur, ainsi que dans la mise en place d'une chaîne éditoriale faisant la part belle à l'exploitation des données ; d'un point de vue technique, le travail donne lieu à des développements principalement en XML/TEI, XSLT, R et Python.

### PREMIÈRE PARTIE LES MANUSCRITS

#### Chapitre premier Les *codices*

*Les manuscrits d'Otinel dans le corpus épique.* — Au sein du corpus des 206 manuscrits et 106 fragments des 132 chansons de geste que nous avons conservées, la *Chanson d'Otinel* se situe par sa tradition manuscrite, avec un fragment de la fin du XII<sup>e</sup> siècle, deux témoins plus complets, du dernier tiers du XIII<sup>e</sup> et du début du XIV<sup>e</sup>, ainsi qu'une citation de quelques vers dans un autre manuscrit, entre la moyenne et le quart supérieur. Par sa variété typologique et chronologique, elle fournit en outre un cas d'étude intéressant, qu'il convient de replacer dans le cadre plus vaste de la production manuscrite épique médiévale.

S'il est particulièrement délicat d'estimer la quantité de manuscrits et de textes perdus, certains indices permettent de soupçonner que la décimation a été plus forte pour les textes vernaculaires que latins, *a fortiori* pour la littérature récréative, et qu'elle s'accroît encore pour la production manuscrite moins onéreuse (*libelli*) et les témoins plus anciens (le nombre minime de manuscrits « de jongleurs » du XII<sup>e</sup> siècle conservés en

témoigne assez bien). Une projection permettant d'estimer globalement le nombre de témoins perdus est nécessairement dépendante du taux de décimation, et pourrait s'appuyer notamment sur une confrontation des attestations de gestes perdues et des attestations de geste conservées.

Un survol de la distribution géographique des manuscrits épiques confirme l'importance du foyer anglo-normand au XII<sup>e</sup> siècle, et, aux siècles suivants, du Nord-Est et de l'Est, ainsi que de l'Italie, plutôt que du « centre ». Il permet également de percevoir des distinctions de goût, avec notamment une apparente prédilection des terres de l'ancienne Lotharingie pour la littérature épique. Avec deux témoins anglo-normands (les plus anciens), un troisième qui, quoique de Saint-Brieuc, s'apparente par sa langue et ses sources au Nord-Est, et des attestations liégeoises et messines, *Otinél* paraît à nouveau assez représentatif.

Si l'opposition entre manuscrits luxueux, « de librairie », et manuscrits portatifs, « de jongleur », mérite d'être, et a été, nuancée, le corpus est néanmoins traversé par des distinctions typologiques, mais qui lui sont beaucoup moins particulières que ce que l'on a pu dire. Ainsi, l'opposition entre manuscrits de petit-format, pouvant circuler sous forme de cahiers non reliés (les *libelli*) – qui sont peut-être plutôt à comparer à la production scolastique – et manuscrits plus luxueux, avec des types intermédiaires, semble renvoyer à une réalité plus générale de la production et du commerce livresque médiéval.

L'étude des possesseurs des manuscrits épiques, même si elle ne devient riche de renseignements que pour les périodes tardives (à partir, surtout du XIV<sup>e</sup> siècle), est néanmoins une porte d'entrée intéressante sur l'histoire de la circulation et de la réception des textes épiques. Elle permet de mettre en valeur l'importance du public aristocratique, féminin peut-être même avant que d'être masculin, et ce jusqu'à des périodes assez récentes, phénomène pour lequel plusieurs explications sont proposées (appétence pour la lecture vernaculaire plutôt que latine, fonction éducative, ...). En outre, on peut tracer la possession de manuscrits épiques en français en Angleterre jusqu'à la fin du Moyen Âge et au-delà, même si le public paraît se rétrécir au fil du temps, et se limiter, à la fin du XV<sup>e</sup> et à l'aube du XVI<sup>e</sup> siècle, aux premiers cercles de l'entourage royal et de la plus haute aristocratie. Mais les possesseurs de manuscrits de chansons de geste se trouvent également assez abondamment dans les abbayes anglaises, dans lesquelles ces textes pouvaient fournir un contrepoint récréatif à des lectures latines plus savantes. Dans certains cas, une fonction mémorielle ou idéologique ne peut tout à fait être exclue. Sur le continent, on retrouve également des manuscrits épiques dans la bibliothèque de certains chanoines. La situation pour la bourgeoisie reste plus difficile à appréhender.

*Notices des manuscrits.* — Le ms. Paris, BnF, nouv. acq. fr. 5094 + (?) Clermont-Ferrand, Arch. dép., 1 F2 (témoin *M*) est à la fois le plus ancien témoin d'*Otinél* et le plus ancien recueil connu de chansons de geste. Conservé sous forme de deux fragments séparés, le second plus tardif que le premier, il contenait au moins *Otinél* et *Aspremont*, dans cet ordre, ainsi que des ajouts plus récents (fabliau du *Héron*, *Couplets sur le mariage*). De grand format, hésitant entre deux et trois colonnes assez comprimées verticalement, il fournit des exemples d'écritures pré-gothiques évoquant les types universitaires, et sa décoration se compose uniquement d'initiales filigranées ou peintes. D'origine anglo-normande et datant de la fin du XII<sup>e</sup> ou du début du XIII<sup>e</sup> siècle, par son format moyen-grand, son niveau d'exécution et la quantité de texte qu'il porte sur chaque feuillet, ce recueil organique évoque le type des manuscrits « de conservation » dans lequel prime l'intérêt pour le texte, revêtu d'une forme d'aspect « officiel ». Ses provenances notariales ultérieures (Lozère et Puy-de-Dôme), traçables à partir du XVII<sup>e</sup> siècle, pourraient rendre compte des vicissitudes des manuscrits des abbayes anglaises qui, après la Dissolution, furent notamment envoyés en masse sur le continent pour servir de matériau de récupération pour des usages divers.

Le ms. Paris, BnF, fr. 25408, *olim* Notre-Dame 273<sup>bis</sup> (*t*), assez secondaire dans la mesure où il ne fournit que quatre vers de la chanson, donne néanmoins une attestation datée de 1267. De petit format et de qualité médiocre, ce recueil assez personnel, peut-être possédé à date ancienne par un clerc et mêlant une variété de textes français et latins réunis par une nature didactique, morale ou pieuse, contient en outre une série d'additions, qui laissent supposer qu'il a été également utilisé pour des usages pédagogiques. Quoique lui-aussi d'origine anglo-normande, il a par la suite appartenu à Claude Fauchet (†1601), avant d'intégrer, vraisemblablement au XVII<sup>e</sup> siècle, les collections de Notre-Dame de Paris.

Le ms. Cologny-Genève, fondation Martin-Bodmer, cod. Bodmer 168 (*B*), témoin le plus complet, conte-

nant *Waldef*, *Gui de Warewic* et *Otinél*, est de format petit-moyen, et propose une mise en page sur deux colonnes et une écriture de type *Textualis Libraria* (*Textus rotundus*) qui partage des caractéristiques des écritures universitaires ou plus courantes. Sa décoration se limite à des initiales filigranées, et à quelques dessins postérieurs à la plume. Ce recueil organique, copié en Angleterre (peut-être en Est-Anglie ou dans le Yorkshire), vraisemblablement dans le dernier tiers du XIII<sup>e</sup> siècle, propose un bon exemple d'une production anglo-normande de qualité intermédiaire, qui brille surtout par sa lisibilité et la clarté de sa copie. Il laisse songer à un recueil dédié à une lecture peu savante, de loisir, pour un public disposant néanmoins d'une certaine aisance. Vraisemblablement commandité par une famille aristocratique anglaise, on le retrouve par la suite entre les mains de Jane Woodville, ép. Grey († av. 1492), sœur de la reine Elisabeth Woodville. Circulant ensuite sans doute parmi des membres de la famille de Grey, jusqu'à Katherine Grey (1540-1568), sœur de la « reine de neufs jours », il parvient au XVIII<sup>e</sup> siècle à William Fermor de Tusmore, puis passe à Richard Heber, Sir Thomas Phillips et ses descendants, avant d'être acquis par Martin Bodmer.

L'abbaye de Peterborough possédait jadis un manuscrit, Q.xv du *Matricularium librerie Monasterii Burgi sancti Petri* de la fin du XIV<sup>e</sup> (p\*), aujourd'hui perdu ou non identifié, décrit comme contenant *Gui de Bourgogne* et *Otinél*, précédés d'un texte qui n'est pas nommé. La composition et la copie de textes français est attestée à Peterborough au XII<sup>e</sup> siècle, de même que la réalisation ou la commande manuscrits enluminés aux scènes évoquant la littérature profane (*Jehan et Blonde*, thèmes courtois, ...), dans une facture oscillant entre *Court style* et *style « est-anglien »*. Le *matricularium* garde en outre la trace de nombreux textes français, épiques au premier chef.

Le ms. Rome, Bibl. ap. vat., Reg. lat. 1616 est un recueil factice, réunissant un recueil de traités musicaux latins du XII<sup>e</sup> siècle, un fragment d'un recueil de textes latins des IX<sup>e</sup>-XII<sup>e</sup> siècles contenant notamment des *Fables* de Phèdre de copie carolingienne, qui a été soustrait à l'abbaye de Fleury par Pierre Daniel (1530-1604), et un fragment d'un recueil de chansons de gestes françaises, copié à Saint-Brieuc en 1317. Ce dernier recueil, contenant *Otinél* (A) et *Fierabras* (V2 et V1), de petit format et copié vraisemblablement par quatre mains différentes, est de qualité médiocre. Sa décoration se limite à des initiales filigranées et il correspond assez largement à la définition traditionnelle des manuscrits « de jongleur ». Si son ordre a été très perturbé à la reliure et que le manuscrit comporte plusieurs lacunes (cahiers et bifeuillets perdus), on peut émettre l'hypothèse qu'il se constitue de *libelli* indépendants, sortis du même atelier, et reliés plus tard comme parties d'une recueuil cumulatif. Des indices pointent, pour cet unique témoin continental, vers des ancêtres du Nord-Est (Wallonie, Lorraine). Les trois entités codicologiques ont probablement été reliées ensemble alors qu'elles se trouvaient dans les collections de Paul (1568-1614) ou Alexandre Petau (1610-1672), d'où elles passeront à la reine Christine de Suède, puis au Vatican où le manuscrit demeure, un bref séjour à Paris excepté, jusqu'à ce jour.

## Chapitre II Les écritures

*Transcription et formalisation des données.* — L'acte de transcription d'un manuscrit médiéval, loin d'être mécanique, résulte de choix et d'une sélection opérée sur les données. Que l'on conçoive cette transcription comme une restitution partielle, une traduction ou une description, elle participe d'un processus sélectif, indissociable de sa dimension interprétative et critique. Dans le domaine de l'édition des anciens textes français, les règles élaborées par Paul Meyer et revues par Mario Roques, telles qu'elles sont énoncées par les manuels, sont souvent mises en œuvre sans faire l'objet d'une réflexion particulière, alors même qu'elles correspondent à un état de la recherche désormais ancien, et masquent des phénomènes qui sont aujourd'hui étudiés par la linguistique de l'écrit. L'édition électronique permet de rebattre les cartes, en autorisant l'encodage de phénomènes tels que l'allographie, les abréviations, la segmentation et la ponctuation ou les diacritiques médiévaux.

L'enregistrement des données allographétiques et abrégatives permet de nombreuses applications, tant du point de vue de la catégorisation générale des écritures et de l'étude de leur évolution, que dans la constitution

de profils scribaux, individuels ou collectifs, l'identification des mains et l'analyse linguistique du système graphique. En outre, il rend plus transparent le processus de résolution des abréviations, qui peut être interrogé et vérifié par l'utilisateur de l'édition, et, du côté de l'éditeur, rendu plus systématique. Enfin, un encodage de ce type permet de mieux cerner la nature et l'ampleur des interventions éditoriales, tout en fournissant au lecteur la possibilité de sélectionner pour lui-même les phénomènes qu'il souhaite étudier, et en rendant possible des approches qui se fondent sur une modélisation des faits observés, par opposition à celles qui cherchent dans les données la validation d'une théorie préexistante. Par ailleurs, l'encodage des interventions scribales – corrections, additions, repentirs... – permet l'étude des processus de copie et une forme de critique « génétique » des témoins médiévaux.

Une approche de ce type nécessite néanmoins un lourd travail de modélisation visant à sélectionner et catégoriser l'ensemble des faits que l'on souhaite enregistrer, et à en déterminer une représentation. Cette opération relève de dimensions à la fois scientifiques et conceptuelles, ainsi que techniques, qu'il importe de décrire et de documenter.

*Caractéristiques générales du corpus.* — Le phénomène allographétique peut être examiné sous différents aspects, à savoir ses sources – la provenance du stock de variantes de formes employé –, ses facteurs et ses fonctions. Le corpus des manuscrits d'*Otinél* permet de faire ressortir certaines évolutions chronologiques, connues ou moins connues, et de dégager des tendances générales qui peuvent être utiles à la datation et à la catégorisation des écritures vernaculaires livresques. L'analyse des facteurs menant à l'emploi d'une forme minoritaire, marquée, permet d'établir l'importance primordiale du contexte, entendu comme caractère précédant immédiatement l'allographe employé et comme position dans le mot ou dans la ligne, ainsi que le rôle plus mineur de la catégorie du discours ou de la valeur phonologique. Des explications liées tant à l'acte d'écriture et au parcours de la main sur la ligne, qu'à celui de lecture et au parcours de l'œil, ou bien encore à la compréhension et à l'espace sémantique du texte, c'est-à-dire à l'oreille, peuvent être proposées.

La question des « majuscules » se révèle assez nettement problématique. Si leur identification est souvent considérée comme acquise et si elles font parfois l'objet d'études autonomes les séparant du reste du phénomène allographétique, l'usage médiéval, pourtant, ne rentre guère dans les catégories et oppositions modernes. La notion de *litterae notabiliores*, moins restrictive, paraît plus proche de la situation observée dans les manuscrits, où des formes de la capitale ou de l'onciale, mêlées à des minuscules grossières ou distinguées par un quelconque artifice, peuvent remplir les mêmes emplois.

L'observation de la pratique abrégative, telle qu'on peut la faire dans les transcriptions des témoins d'*Otinél*, permet de confirmer ou d'infléchir certaines hypothèses. La plus faible polyvalence (ou ambiguïté) des signes abrégatifs dans les témoins anglo-normands semble confirmer l'hypothèse de G. Hasenohr relative à l'origine insulaire de leur mise en œuvre pour le vernaculaire d'oïl. Les scribes tendent également à respecter les frontières de mot et la règle de syllabation, avec néanmoins des exceptions qui deviennent plus prononcées dans le manuscrit le plus tardif. L'analyse quantitative permet en outre d'expliquer en partie l'utilisation d'un signe abrégatif par une série de facteurs dont les plus importants sont le nombre de jambages du mot, ainsi que la position du signe abrégatif dans le mot et du mot dans la ligne. À l'échelle de notre corpus, la densité abrégative paraît augmenter avec le temps, sans corrélation évidente avec la variété des abréviations utilisées, oscillant à la hausse ou à la baisse d'un témoin à l'autre, ou la faveur plus grande octroyée à un mode abrégatif donné. Le facteur le plus déterminant s'y révèle être l'extension du nombre de mots abrégés au moins une fois.

En ce qui concerne l'accent, dont l'emploi participe de la ponctuation médiévale, l'étude de notre corpus semble confirmer l'évolution d'un système relativement développé hérité du latin, enrichi d'expérimentations dans les plus anciennes copies anglo-normandes, puis réduit à une simple fonction de distinction des jambages. Seul *M* présente ainsi une utilisation plus étendue, dans laquelle l'accent peut aussi être utilisé pour marquer un hiatus ou une voyelle tonique, notamment un *e* non atone en syllabe finale ; ces usages sont déjà absents du fragment de Clermont-Ferrand. La ponctuation est particulièrement pauvre dans nos manuscrits, les scribes employant assez de signes pour nous laisser à voir qu'ils les connaissent et de manière suffisamment épisodique pour que nous comprenions que leur fonction reste marginale. S'il est vrai que le décasyllabe épique, dans

lequel se coule la syntaxe du texte, rend moins pressant le besoin d'une ponctuation extensive, on constate néanmoins une rareté même des emplois métriques (enjambements, rejets), l'usage principal étant le marquage des juxtapositions d'éléments de même niveau. L'étude de la segmentation de la chaîne graphique pose des difficultés dans la transposition d'un système qui ne se construit pas sur une opposition binaire (présence ou absence d'une espace) mais plutôt sur une gamme de pauses de différentes longueurs, de l'espace entre lettres à celle qui sépare des syntagmes ou propositions, qu'il peut être délicat d'identifier et représenter. Avec ces réserves, l'analyse du corpus permet néanmoins de dégager une tendance générale à l'agglutination des prépositions, plus rarement des conjonctions de coordination, déterminants, pronoms ou adverbes, avec les mots qui les suivent directement, surtout quand ceux-ci sont des noms, déterminants ou verbes, menant parfois à des séquences de plusieurs agglutinations successives. À l'inverse, l'élosion mène parfois à la déglutination de la voyelle initiale du mot suivant. Le corpus permet également de voir à l'œuvre les différents modes de correction employés par les scribes et d'observer, en dépit de préférences apparentes pour l'un ou l'autre d'entre eux, une utilisation de l'ajout interlinéaire ou de la surcharge pour des corrections de niveau inférieur au mot et des autres modes pour celles de niveau supérieur, à plus forte raison encore pour ce qui est de l'ajout marginal ou de la rature.

*Analyse des trois manuscrits.* — Répondant à un objectif double de modélisation – décrire les données observées, définir les classes retenues et en proposer une représentation – et d'analyse individuelle du système graphique, les analyses successives des fragments de Mende et des manuscrits Bodmer et du Vatican envisagent l'allographie, la pratique abrégative, la ponctuation, segmentation, et les corrections, dans le cadre des usages propres à chaque scribe ou groupe de scribe œuvrant de concert. À côté d'une critique paléographique plus traditionnelle, le recours aux méthodes quantitatives (écarts à l'indépendance, régressions, arbres de décision) permet d'étayer l'analyse des phénomènes individuels. Ces analyses confirment des constats paléographiques préexistants, mais révèlent également des spécificités dans la pratique de chaque scribe ou des tendances plus générales.

*Les mains, les lieux, les dates.* — En adoptant une approche plus large et plus englobante, des conclusions peuvent être tirées de l'application de différentes méthodes quantitatives à l'ensemble des données représentées. Des méthodes de partitionnement s'appuyant sur des calculs de distance similaires à ceux de la stylométrie permettent, après calibration, d'identifier, dans un même manuscrit, les parties redevables à chacune des mains (trois mains des fragments de Mende et Clermont-Ferrand, quatre mains du ms. du Vatican) et de déterminer celles qui sont les plus similaires entre elles. À partir de cette identification, des profils scribaux peuvent être créés, faisant ressortir les digrammes les plus caractéristiques de chaque main, ainsi que ceux les plus significatifs dans leur comparaison. Ces profils individuels peuvent être en partie compris comme représentant des tendances, chez un scribe, à enfreindre l'usage le plus courant ou à l'appliquer avec plus ou moins de constance. En élargissant encore la vue, l'étude de l'ensemble du corpus, qui mériterait d'être complétée par la prise en compte d'autres manuscrits vernaculaires, permet de rendre apparente l'évolution chronologique générale et les transformations collectives.

### Chapitre III La langue

*Datation et localisation de la scripta.* — La distinction des variétés régionales écrites du français médiéval nécessite une appréhension globale plutôt que l'étude de traits « isographes » isolés. Dans ce contexte, la scriptométrie (ou dialectométrie) ouvre de nouvelles possibilités. En conjoignant les textes d'*Otinél* à l'ensemble du *Nouveau Corpus d'Amsterdam* (296 textes) et en leur appliquant des méthodes de partitionnement, on peut ainsi faire ressortir des lignes de séparation principales, qui isolent en premier lieu les copies anglo-normandes et de l'Ouest, puis séparent les témoins du Nord, voire du Nord-Est, du reste du corpus. Ces analyses rendent également apparentes les attractions génériques, stylistiques ou autoriales, ainsi que les difficultés posées par la stratification linguistique des manuscrits littéraires ou le caractère relatif des localisations. Cela justifie, dans un second temps, de restreindre l'analyse à un corpus spécifique de chansons de geste (47 textes), dans

lequel on retrouve les mêmes subdivisions dialectales principales, mais qui permet également de mieux cerner les contours et subdivisions internes des *scriptae* de l'Est (Lorraine, Lorraine méridionale, Bourgogne), et de proposer une datation relative de la langue des témoins, en fonction de leur place dans le groupe auquel ils appartiennent. Il est ainsi possible d'affiner la localisation ou d'en proposer une nouvelle pour la langue de certains manuscrits de chanson de geste (*Destruction de Rome*, *Gui de Bourgogne*, *Floovant*, *Florence de Rome*, *Couronnement de Louis*, fragm. BnF, nouv. acq. fr. 934 du *Charroi*, ms. BnF, fr. 2493 de *Raoul de Cambrai*, *Florence et Octavien*, *Aye d'Avignon*), et de placer *B* et *M* dans un groupe anglo-normand ancien, et *A* dans un groupe burgondo-lorrain tardif. Des profils scriptologiques peuvent enfin être réalisés, donnant les formes les plus caractéristiques de chacun des groupes principaux et faisant apparaître, à côté de marqueurs dialectaux bien connus, d'autres plus étonnants.

*Description de la scripta des témoins.* — Le travail de lemmatisation, d'annotation morpho-syntaxique et flexionnelle du texte de l'ensemble des témoins d'*Otinél* (jeu d'étiquette CATTEX2009-max) permet la mise en œuvre d'une étude systématique de leur *scripta*, dans ses dimensions phonologique et graphématique, ainsi que morphologique et morpho-syntaxique. La syntaxe, moins sujette à la variation scribale que l'habillage graphique des textes, se montre particulièrement stable d'un témoin à un autre et permet déjà des hypothèses sur la langue de l'original ; son analyse sera complétée et affinée une fois que l'annotation syntaxique sera achevée. L'étude du lexique pourra s'enrichir par la poursuite du travail sur le glossaire et des recherches sur le caractère régional de certains lemmes.

Indubitablement anglo-normande et ancienne, la *scripta* de *M* se distingue également par des traits qui pourraient renvoyer à des substrats du Nord-Est ou de l'Est (Wallonie, Lorraine, Bourgogne). Également anglo-normande, la *scripta* de *B*, plus tardive (seconde moitié ou dernier quart du XIII<sup>e</sup> siècle), paraît plus influencée par des normes continentales. Soit en raison d'un plus grand nombre d'intermédiaires, soit à cause d'une tendance plus nette du copiste à neutraliser les formes qui lui paraissent trop marquées, on peine à distinguer, mis à part certains traits qui pourraient être redevables à un ancêtre anglo-normand commun à *M*, des traces des substrats de la langue de ce témoin. Quoique daté de 1317 à Saint-Brieuc, *A* présente une langue qui, sous de rares traits de l'Ouest et des picardismes peut-être dus à l'influence de la *koinè* épique d'alors, se marque par de nombreux traits de l'Est et du Nord-Est. La présence à la fois de traits wallons, lorrains et bourguignons peut amener à soupçonner plusieurs substrats légèrement différents, dus à des intermédiaires depuis l'archétype ou à l'héritage d'une langue déjà mixte.

Les principales observations en matière de syntaxe se rejoignent d'un manuscrit à l'autre et pointent principalement vers le Nord-Est (notamment Wallonie et Lorraine). Une comparaison avec un corpus, annoté en syntaxe, de chansons de geste, serait particulièrement utile pour affiner l'analyse.

*Langue des archétypes, langue de l'original.* — La synthèse proposée ici, provisoire, et qui devra être enrichie par l'étude de la versification de l'œuvre, et s'appuyer sur l'étude généalogique de la tradition, cherche à dégager certaines caractéristiques linguistiques des subarchétypes et de l'original de la chanson. L'ancêtre commun des versions insulaires et scandinaves possédait déjà un caractère anglo-normand marqué, à côté, peut-être, de traits continentaux du Nord ou du Nord-Est. On peut soupçonner, pour la langue de l'original, une provenance du Nord-Est (Wallonie ou Lorraine).

## DEUXIÈME PARTIE DE LA TRADITION VERS L'ŒUVRE

### Chapitre IV Prolégomènes à l'étude de la tradition

*La diffusion et les transformations d'Otinél.* — La diffusion en Europe de la chanson paraît avoir connu deux trajectoires, à première vue relativement indépendantes, vers l'Europe du Nord via les Îles britanniques, et en Europe continentale, vers l'Italie. À partir de manuscrits anglo-normands, plusieurs traductions ont été réalisées au cours des XIII<sup>e</sup> et XIV<sup>e</sup> siècles.

La traduction norroise (*NOt*), branche VI de la *Karlamagnús Saga*, connaît elle-même des dérivés danois ou féroéens. Au sein de la tradition de la *Karlamagnús saga*, les manuscrits de la rédaction  $\beta$ , réputés inférieurs, préservent parfois des leçons plus proches de l'original français que ceux de la rédaction  $\alpha$ . *NOt* reste particulièrement intéressant pour l'établissement du texte français, même si le traducteur adapte ou remanie parfois le contenu du texte, et commet également quelques contresens.

Ajout tardif à la compilation galloise sur l'histoire de Charlemagne (*Ystoria de Carolo Magno* ou *Campeu Charlymaen*), la version galloise, *WOt*, est conservée par trois manuscrits dont il est relativement aisé d'établir la généalogie. Parmi ceux-ci, ce sont tantôt les Peniarth 9 (*B*) et 4-5 (*W*, connu comme le Livre blanc de Rhydderch) qui ont raison contre le ms. Oxford, Jesus College, 111 (*R*, Livre rouge de Hergest), et tantôt l'inverse. L'établissement du texte critique de la version galloise, qui pourrait en partie s'appuyer sur le français, serait particulièrement utile à l'établissement du texte de la chanson, étant donné la très grande fidélité du traducteur gallois, et la proximité du manuscrit qu'il a utilisé avec *B*.

Les traductions anglaises, *Otuel a Knight* (*EOtA*) et *Otuel and Roland* (*EOtF*) d'une part, *Duke Roland and Sir Otuel of Spayne* (*EOtT*) de l'autre, sont d'un intérêt plus limité pour l'établissement du texte français, leur structure métrique propre ayant amené les traducteurs à un remaniement plus marqué.

Les reprises de la légende par Jean d'Outremeuse, Jacopo d'Acqui ou l'auteur du *Cantare dei cantari* valent surtout pour les attestations qu'elles fournissent de la circulation d'éléments de la geste.

*L'apport des œuvres dérivées à l'étude de la tradition de la chanson.* — La relative rareté des témoins d'oïl conservés, comme le nombre important de traductions et œuvres dérivées, rend la prise en compte de ces dernières indispensable à toute étude de la tradition du texte. S'il est clair qu'aucun des témoins d'oïl ne descend d'un des autres, on distingue clairement l'existence d'une branche anglo-normande (*BM WOt NOt EOtAEOtF EOtT*), et, au sein de celle-ci, on rencontre des erreurs communes à *BM*, voire *BMWOt*. Des points d'accords entre traductions (par exemple entre *WOt* et *NOt*) contre *B* soulignent la nécessité de prendre en compte de manière extensive ces traductions, sans se limiter aux cas où elles soutiennent un témoin français contre d'autres. Parmi les traductions anglaises, *EOtT* paraît avoir parfois gardé seul des leçons attestées uniquement en dehors de la famille anglo-normande. Le fragment *t* descend peut-être d'un archétype différent, mais est trop court pour être véritablement placé. L'approfondissement de ces études devrait néanmoins permettre de dégager le stemma de la chanson.

*Origine, date et composition de la chanson.* — L'onomastique, l'archéologie, de pair avec différentes attestations textuelles, permettent de dégager à grands traits la diffusion de la chanson, qui semble se construire autour d'un axe européen correspondant plus ou moins aux principales routes Nord/Sud, de l'Angleterre à Rome. Si les attestations onomastiques anglaises les plus anciennes sont suspectes, les attestations du domaine d'oïl continental pointent vers la Wallonie, surtout le pays liégeois, et la Lorraine, plus précisément Metz. La question des rapports de la geste d'*Otinél* avec l'Italie est particulièrement vaste, mais tant l'onomastique, que les fresques de Trévise et Sesto ou le *Cantare dei cantari*, semblent attester d'une diffusion partant de la Vénétie au dernier tiers du XII<sup>e</sup> siècle, et touchant ensuite la Toscane, au premier tiers du siècle suivant, puis l'Ombrie et la Lombardie. La question de l'existence d'un *Otinél* franco-vénète perdu reste posée. Enfin, la chanson est également attestée, assez marginalement, dans la péninsule ibérique.

Si la question des origines méritera des approfondissements ultérieurs, différents indices, concernant notamment le nom du héros, ainsi que des rapprochements possibles avec des textes comme la *Chevalerie Ogier*, semblent pointer vers le Nord-Est, peut-être plutôt la Wallonie que la Lorraine.

## Conclusions

L'étude de la généalogie, de la diffusion et de l'origine du texte, ébauchée dans ce travail, méritera d'être poursuivie. Un recours assez large aux traductions galloises et norroises devrait permettre, au moins pour la partie médiane du texte, de restaurer la substance, si ce n'est la lettre, d'un certain nombre de leçons qui demeurent encore problématiques. La reconstruction de l'ancêtre commun de *BM* sera relativement aisée, et pourra prendre appui sur le reste de la tradition. Éminemment souhaitable, la reconstruction du subarchétype



anglo-normand posera des difficultés bien plus sérieuses, notamment lorsque les traductions seules gardent trace de leçons qui ont été corrompues dans *BM*. Néanmoins, si elle s'avérait possible, la présentation de ce texte critique en synoptique de *A* permettrait de donner une idée assez nette de la tradition du texte et, vraisemblablement, de la leçon de l'archétype, même s'il est encore trop tôt pour dire si une tentative reconstructionniste de ce dernier, à la fois but ultime de l'éditeur et entreprise risquée par nature, s'avèrera possible.

## ÉDITION

*Note liminaire à l'édition.* — Les différentes visualisations proposées sont générées de manière automatisée, à partir de l'encodage en XML/TEI du texte des différents témoins. Elles proposent ainsi des sélections différentes et non exhaustives des données telles qu'elles ont été représentées par l'encodage, et cherchent à satisfaire aux besoins de différentes problématiques de recherche et disciplines.

## Transcriptions

*Transcriptions allographétiques.* — Des transcriptions allographétiques intégrales de *M*, *B* et *A* sont proposées. Elles ont pour finalité principale l'étude du système graphique, dans des perspectives relevant de la paléographie et de la linguistique de l'écrit, tout en pouvant également servir à l'étude du diasystème et des phénomènes propres aux processus de copie. Les informations d'ordre paléographique y sont imprimées en priorité : pas de normalisation des allographes, ni de résolution des abréviations ou de la ponctuation et des accents médiévaux ; letrines imprimées aux dimensions qu'elles occupent dans le manuscrit, en nombre de ligne de réglure ; pas de normalisation de la segmentation, ni d'emploi des diacritiques des règles de Meyer-Roques ; indication des ajouts ou suppression par des artifices typographiques. Elles sont complétées par deux étages de notes comportant des indications sur les difficultés de lecture ou la matérialité des manuscrits, ainsi que des notes de commentaire paléographique.

*Transcriptions graphématiques alignées.* — Ces transcriptions alignées des trois témoins, également intégrales, ont pour objet principal la critique textuelle et l'étude de la *scripta*. La présentation synoptique est conçue pour favoriser la critique des variantes. Les allographes sont normalisés, et les abréviations résolues, selon les principes exposés en introduction, les suppressions des copistes retirées et leurs ajouts intégrés ; une ponctuation éditoriale se substitue à la ponctuation médiévale, et les majuscules sont alignées sur l'usage moderne. Les diacritiques des règles de Meyer-Roques sont utilisés. L'apparat fait figurer, sur un premier étage, des notes concernant les additions et suppressions, les difficultés de lecture, les lacunes ou les entorses à la mesure ; sur un second, des notes de critique textuelle. Des signes complémentaires sont employés pour signaler, dans chaque témoin, du texte suppléé pour combler une lacune matérielle, les erreurs évidentes de faible portée généalogique, les variantes substantielles, les omissions, ajouts ou inversions.

## Annexes de l'édition

Les annexes de l'édition contiennent la liste des lemmes problématiques pour la résolution des abréviations. Les glossaires et l'index des noms propres, en cours d'élaboration, n'ont pas encore été intégrés.

## ANNEXES ET PIÈCES JUSTIFICATIVES

### Annexe A Pièces justificatives

Collation des variantes du passage concernant *Otinell* dans la *Cronica Imaginis Mundi* de Jacopo d'Acqui.

## Annexe B Annexes techniques de l'édition

*Préparation de l'édition et pistes d'automatisation.* — Si, pour cette édition, la transcription allographétique a été réalisée directement par l'éditeur, des expérimentations ont été menées avec des logiciels de reconnaissance optique de caractères. Un modèle pour *Ocopy*, qui utilise des méthodes d'intelligence artificielle, a été entraîné, et est parvenu jusqu'à 90% d'efficacité sur le manuscrit *B*. En ce qui concerne l'annotation linguistique, on a également eu recours à l'intelligence artificielle, et plusieurs modèles pour le logiciel *Pandora* ont été entraînés.

*Architecture globale de l'édition : de l'encodage à l'exploitation (analyse et publication).* — La chaîne éditoriale mise en œuvre pour la *Chanson d'Otinél* a été conçue pour laisser une large place à l'exploitation des données. Les documents sont encodés en XML/TEI, selon un modèle exprimé comme personnalisation des recommandations de la *Text Encoding Initiative* et formalisé par un ODD. À partir des documents, des scripts XSLT permettent des exports CSV pouvant être utilisés avec le langage d'analyse statistique R, et aux formats TEX (L<sup>A</sup>T<sub>E</sub>X) et HTML pour la visualisation imprimée ou numérique de l'édition.

## Annexe C Modèle de l'édition

Le modèle est présenté dans sa version actuelle (octobre 2016), quoique toujours en développement, et devant être élargi pour affiner, notamment, les descriptions codicologiques et l'encodage de l'apparat critique. Il est une personnalisation des *Guidelines* de la TEI, dont il reprend en partie l'ordre, et, bien qu'ayant été conçu pour l'édition d'*Otinél*, il se veut suffisamment généraliste pour pouvoir être appliqué à d'autres chansons de geste, depuis des numérisations d'éditions jusqu'à des transcriptions allographétiques à nouveaux frais ou des éditions critiques se fondant sur une tradition complète. Dans le cadre d'un projet de recherche en cours, il est déjà mis en œuvre pour l'encodage d'un corpus de chansons de geste.

Le modèle est lui-même modulaire et construit sur trois niveaux : un socle de base est commun à toutes les éditions, un second niveau s'y ajoute pour les transcriptions de manuscrits et un troisième pour les éditions critiques.

## Annexe D L'édition synoptique... et après ?

À partir d'une prise en compte de la situation particulière de la variance textuelle dans les chansons de geste et de l'étude des modèles existants pour l'encodage des variantes, des pistes sont tracées pour la modélisation. Celle-ci doit permettre la critique des fautes et l'établissement d'un catalogue organisé de ces dernières, ainsi que l'analyse stemmatologique, selon une méthode proposée par l'auteur dans un article séparé. Enfin, cet encodage permettra l'affichage synoptique du texte des témoins, et des textes critiques reconstruits.

## Annexe E *Corpora* et bases de données

Manuscrits épiques dans les bibliothèques anglaises médiévales (dépouillement des volumes du *Corpus of British Medieval Library Catalogues*). — Corpus de chansons de geste pour l'analyse de la *scripta* (sources et description).

## Indices et tables

Index des cotes anciennes. Index des cotes actuelles de manuscrits. Index des auteurs cités dans les références bibliographiques. Liste des tableaux. Table des figures. Table des matières.

The *Chanson d'Otinel*:  
Complete Edition of the Manuscript Corpus and Prolegomena to a Critical Edition

Part of the *geste du roi*, the *Chanson d'Otinel* had not been the subject of an edition since the pioneer work of F. Guessard and H. Michelant in 1858. Starting with the tangible objects that manuscripts are, and proceeding to the study of the tradition and work itself, this study aims to provide a new examination of all the available data, in order to enable a restoration of a song that has known an important diffusion in Medieval Europe, but whose scattered remains only are available to us.

This study is given a firmly methodological orientation, and searches to apply to the edition the epistemological progresses brought by recent contributions in the field of textual criticism and ecdotics, as well as by what is now called “Digital Humanities”. It aims at overcoming the debate between “New” and Traditional Philology, based on the common errors method or of Bedierist inspiration, to place itself in the perspective of a “Fourth Way” and of an edition “oriented towards the tradition”.

Digital scholarly editing and Digital Philology techniques are used in order to provide a full edition of the manuscript corpus, with “layered” transcriptions that give access to different representations and to the graphic system of all witnesses. The edition aims also at the study and description of the links between these witnesses, by suggesting a method of representation of textual variance and by rooting genealogical analysis in a global consideration of the tradition, including medieval translations (in Welsh, Norse, English) and derived versions.

The modelling and description of manuscripts and their texts – formalised by an XML/TEI model conceived for this edition but seeking to be of more general interest – is clearly oriented towards data mining, and computational as well as traditional analysis (Palaeography, Scriptometry, Stemmatology). Methods from the fields of mathematical modelling, statistics, algorithmic, and artificial intelligence are put to use, as well as processes seeking to allow interoperability, scalability, and systematisation of editorial work (optical character recognition, linguistic tagging, collation), in a workflow centred on analysis. Main used languages are XML (TEI, XSLT), R and Python.

**Keywords:** Romance Philology; Medieval Literature; chanson de geste; chanson d'Otinel; Circulation and Reception of *oïl* Literature; Medieval Translations; Ecdotics; Digital Editing; Digital Philology; Codicology; Text History; Palaeography; Scriptology; Stemmatology; Computational Methods; Digital Humanities

UNIVERSITÉ PARIS-SORBONNE  
École doctorale 1 – Mondes anciens et médiévaux  
Maison de la Recherche, 28 rue Serpente, 75006 Paris, France  
**DISCIPLINE** : Études médiévales