

UNIVERSITÉ LIBRE DE BRUXELLES

Faculté de Lettres, Traduction et Communication

Étude stylométrique de l'œuvre de
Gerbert d'Aurillac

Guillaume QUINTIN

Travail préparatoire au mémoire présenté sous la direction de Sébastien DE VALERIOLO en vue de l'obtention du titre de Master en Sciences et Technologies de l'Information et de la Communication

Année académique 2021–2022

Table des matières

1	Introduction	1
2	Bibliographie commentée	1
2.1	Stylométrie	1
2.2	Gerbert d'Aurillac	3
3	Première approche du sujet	5
3.1	Gerbert d'Aurillac : sa vie, son époque, son œuvre	5
3.2	La stylométrie	7
3.3	Workflow : lemmatiser et tokeniser un texte	8
3.3.1	Texte brut	8
3.3.2	Texte tokenisé	8
3.3.3	Texte lemmatisé	10
4	Conclusion	13
4.1	Table des matières provisoire du mémoire	13
4.2	Plan de travail	14

1 Introduction

Dans ce travail préparatoire, je tâcherai d'exposer l'état actuel de mes recherches et d'effectuer un tour d'horizon de mon sujet. Cet exposé se veut provisoire, étant l'aspect figé à un moment T d'un processus en mouvement permanent.

Contrairement aux habitudes académiques, je commencerai par la bibliographie que j'ai déjà eu l'opportunité de consulter et je la commenterai. Je m'éloignerai davantage encore des normes en incluant les ouvrages qui, sans pour autant être cités dans le corps du travail, m'ont semblé importants pendant mes recherches. Ensuite, j'introduirai mon sujet à proprement parler : d'une part, l'homme qu'a été Gerbert d'Aurillac en le situant dans son temps; d'autre part, une première approche de la stylométrie accompagnée d'un cas pratique. Je conclurai, de façon temporaire, en évoquant l'avenir, c'est-à-dire en fournissant la table des matières de mon mémoire telle que je l'envisage en ce moment et en établissant un plan de travail qui reprend les prochaines étapes importantes dans le travail.

2 Bibliographie commentée

Ma bibliographie se découpe naturellement en deux sections : la stylométrie et des considérations méthodologiques, et le personnage de Gerbert d'Aurillac dans ses dimensions historiques et littéraires. Pour le moment, la première section est moins développée que la seconde, puisque je me suis concentré sur le personnage et que j'ai repoussé les considérations méthodologiques jusqu'au moment où j'effectuerai des analyses stylométriques à proprement parler. Toutefois, au vu des importantes listes de références mentionnées dans les quelques articles fournissant une vue d'ensemble que j'ai consultés, l'une et l'autre sections sont vouées à s'équilibrer.

2.1 Stylométrie

Maciej EDER. "Does Size Matter? Authorship Attribution, Small Samples, Big Problem". In : *Digital Scholarship in the Humanities* 30.2 (2015), p. 167-182. URL : <https://doi.org/10.1093/llc/fqt066> (visité le 09/01/2022).

Cet article explore la taille minimale des échantillons de texte pour obtenir des résultats fiables dans l'attribution de paternité d'une œuvre. Une des conclusions avancées par l'auteur m'a particulièrement intrigué : la langue latine est celle qui, parmi les langues étudiées dans l'article, demande le moins d'échantillons pour fournir des résultats fiables.

David I. HOLMES. “The Evolution of Stylometry in Humanities Scholarship”. In : *Literary and Linguistic Computing* 13.3 (1998), p. 111-117. URL : <https://academic.oup.com/dsh/article/13/3/111/933269> (visité le 09/01/2022).

Cet article, quelque peu daté, explore l'évolution de la stylométrie depuis le XIX^e siècle jusqu'aux portes du XXI^e.

M. KESTEMONT, S. MOENS et J. DEPLOIGE. “Collaborative Authorship in the Twelfth Century: A Stylometric Study of Hildegard of Bingen and Guibert of Gembloux”. In : *Digital Scholarship in the Humanities* 30.2 (2015), p. 199-224. URL : <https://academic.oup.com/dsh/article-lookup/doi/10.1093/llc/fqt063> (visité le 04/01/2022); Mike KESTEMONT et al. “Authenticating the Writings of Julius Caesar”. In : *Expert Systems with Applications* 63 (2016), p. 86-96. URL : <https://linkinghub.elsevier.com/retrieve/pii/S0957417416303116> (visité le 04/01/2022); Armando MARTINS et al. “Historia Augusta Authorship: An Approach Based on Measurements of Complex Networks”. In : *Applied Network Science* 6.1 (2021), p. 50. URL : <https://appliednetsci.springeropen.com/articles/10.1007/s41109-021-00390-7> (visité le 04/01/2022).

J'ai décidé de traiter ces trois articles ensemble puisqu'ils ont la même pertinence à mes yeux. En effet, ce sont des mises en pratique de l'analyse stylométrique appliquée à des sources anciennes et, plus encore, à des sources en langue latine. Le premier article est encore plus intéressant puisqu'il concerne une autrice qui écrit en latin médiéval. En bref, ces articles seront des guides bienvenus lorsque je me plongerai dans l'analyse stylométrique à proprement parler.

Mike KESTEMONT. “Function Words in Authorship Attribution. From Black Magic to Theory?” In : *Proceedings of the 3rd Workshop on Computational Linguistics for Literature (CLFL)*. Proceedings of the 3rd Workshop on Computational Linguistics for Literature (CLFL). Gothenburg, Sweden : Association for Computational Linguistics, 2014, p. 59-66. URL : <http://aclweb.org/anthology/W14-0908> (visité le 05/01/2022).

Cet article, du même auteur que les analyses stylométriques sur les œuvres de César et d'Hildegarde de Bingen précédemment mentionnées, est de nature plus théorique. C'est d'ailleurs ce que l'auteur revendique : du moins en 2014, il y avait un manque d'articles abordant des questions méthodologiques et théoriques en stylométrie.

Efstathios STAMATATOS. “A Survey of Modern Authorship Attribution Methods”. In : *Journal of the American Society for Information Science and Technology* 60.3 (2009), p. 538-556. URL : <https://onlinelibrary.wiley.com/doi/10.1002/asi.21001>

(visité le 05/01/2022).

Cet article complète celui de HOLMES précédemment mentionné, tout en étant quelque peu plus récent.

2.2 Gerbert d'Aurillac

Carlo BIANCHINI et Luca J. SENATORE. "Gerbert of Aurillac (c. 940/1003)". In : *Distinguished Figures in Descriptive Geometry and Its Applications for Mechanism Science : From the Middle Ages to the 17th Century*. Sous la dir. de Michela CIGOLA. History of Mechanism and Machine Science. Cham : Springer International Publishing, 2016, p. 33-51. URL : https://doi.org/10.1007/978-3-319-20197-9_2 (visité le 25/06/2022).

Brève présentation de Gerbert d'Aurillac, avec une liste concise de ses œuvres (ce qui est assez rare).

Catherine JACQUEMARD. "Erectio, inclinatio / erectus, inclinatus : de Vitruve à Gerbert d'Aurillac (à propos de l'expression de la distance angulaire fin Xe - début XIe siècle)". In : *Collection de l'Institut des Sciences et Techniques de l'Antiquité* 993.1 (2006), p. 157-162. URL : https://www.persee.fr/doc/ista_0000-0000_2006_act_993_1_2169 (visité le 30/01/2022).

Cet article tente de démontrer, à mon sens avec succès, qu'attribuer à Gerbert la capacité de rédiger le *Liber de astrolabio* sur base de quelques mots dans une de ses lettres, en considérant qu'il désignait alors l'alidade de cet astrolabe, n'est pas la seule interprétation possible de ces mots et, plus encore, n'en est pas la plus convaincante. Il fait écho au colloque organisé à Bobbio en 1983.

Flavio G. NUVOLONE, éd. *Gerberto d'Aurillac da abate di Bobbio a papa dell'anno 1000 : atti del congresso internazionale, Bobbio, Auditorium di S. Chiara, 28-30 settembre 2000...* Archivum Bobiense. Studia. Bobbio : Associazione culturale Amici di Archivum Bobiense, 2001.

Les actes d'un des trois colloques qui se sont tenues à Bobbio au sujet de Gerbert (en 2000 pour celui-ci, en 1983 pour celui mentionné ci-après et en 2004 que je n'ai pas encore eu l'opportunité de consulter). Le tableau comparatif réalisé par Marco MOSTERT, des ouvrages de Gerbert et de ceux d'Abbon de Fleury, un de ses contemporains et rivaux.

Alexandre OLLERIS. *Oeuvres de Gerbert, pape sous le nom de Sylvestre II*. Clermont-Ferrand & Paris, 1867. URL : <https://gallica.bnf.fr/ark:/12148/bpt6k408322c>

(visité le 05/02/2022).

Une édition, fort ancienne, de la plupart des œuvres de Gerbert. Elle a l'avantage d'être scannée et d'être passée par une reconnaissance optique de caractères (OCR).

Marek OTISK. "Gerbert of Aurillac (Pope Sylvester II) as a Clockmaker". In : *Teorie vdy/Theory of Science* 42.1 (2020), p. 25-49.

Cet article évoque les dispositifs que Gerbert évoque dans ses ouvrages, en particulier l'astrolabe. Très clair, il a permis au néophyte que je suis de comprendre le fonctionnement de ces appareils.

Pierre RICHÉ. *Gerbert d'Aurillac : Le Pape de l'an Mil*. Paris : Fayard, 1987.

La référence française, si pas mondiale, sur Gerbert d'Aurillac. Même si elle est un peu datée, il n'existe pas, à ma connaissance, d'autres monographies de même qualité qui traitent de l'ensemble du personnage.

Alain SCHÄRLIG. *Un Portrait de Gerbert d'Aurillac : Inventeur d'un Abaque, Utilisateur Précoce Des Chiffres Arabes, et Pape de l'an Mil*. Lausanne : PPUR Presses polytechniques, 2012.

Cet ouvrage est l'autre monographie dédiée à Gerbert, qu'on retrouve aux côtés de celle de Pierre RICHÉ lors d'une première recherche bibliographique superficielle. Elle n'est malheureusement pas de qualité similaire : le ton employé est loin du style scientifique habituel dans les monographies et plus proche du ton de la vulgarisation. De plus, certaines affirmations avancées par l'auteur, lorsque je les compare avec ce que j'ai appris sur Gerbert grâce au reste de ma bibliographie, me semblent exagérées. Une comparaison systématique et une évaluation complète de l'ouvrage excède toutefois les limites de mon travail et correspondrait davantage à une étude centrée sur une comparaison entre le personnage historique de Gerbert et la légende qu'il a laissée derrière lui.

Costantino SIGISMONDI. "Gerbert of Aurillac : Astronomy and Geometry in Tenth Century Europe". In : *International Journal of Modern Physics : Conference Series* 23 (2013), p. 467-471. arXiv : 1201 . 6094. URL : <http://arxiv.org/abs/1201.6094> (visité le 03/02/2022).

Cet article, écrit par un astrophysicien, explore l'astronomie au temps de Gerbert. Il émet notamment des doutes sur la paternité du *Liber de astrolabio*, évoquant le colloque tenu à Bobbio en 1983. Il m'a semblé intéressant de prendre en compte les travaux réalisés par des spécialistes de l'astronomie. L'auteur a écrit d'autres articles, la

plupart en italien, que je compte explorer davantage.

Michele TOSI, éd. "*Gerberto : scienza, storia e mito*" : *atti del Gerberti Symposium (Bobbio 25-27 luglio 1983)*. Archivum Bobiense. Studia, 2. Bobbio : Archivi storici bobbiensi, 1985.

Les actes du premier colloque organisé à Bobbio à propos de Gerbert. Intéressant pour de nombreuses raisons, il l'est dans le cadre de mes recherches grâce aux articles de F. NUVOLONE, E. POULLE et G. BEAUJOUAN. En effet, le premier évoque des doutes sur la paternité du *Sermo de informatione episcoporum*, en effectuant une étude entièrement philologique. Les deux suivants, quant à eux, en évoquent au sujet du *Liber de astrolabio*, principalement en se basant sur les connaissances que Gerbert aurait pu avoir à son époque et qui sont inférieures à celles présentes dans l'ouvrage.

3 Première approche du sujet

3.1 Gerbert d'Aurillac : sa vie, son époque, son œuvre

Cette section est basée en grande partie sur l'ouvrage de P. RICHIÉ,¹ complété par divers articles qui tracent souvent les grands traits de la vie de Gerbert dans leur introduction (voir par exemple l'article de BIANCHINI et SENATORE²)

Gerbert d'Aurillac est né, comme son nom l'indique, à Aurillac, en Auvergne. Les dates sont incertaines mais, selon RICHIÉ, il serait né entre 945 et 950.³ D'origine humble, il intègre un monastère bénédictin. C'est là qu'il commence son éducation, en apprenant la grammaire latine. Très vite, toutefois, en 969 il quitte son Auvergne natale pour se rendre en Catalogne, où il se trouve en contact avec les sciences arabes, ce qui sera le point de départ de nombreuses légendes à son égard. C'est là qu'il perfectionne son éducation dans le *quadrivium*.

Il revient en France en 972, où il devient écolâtre à Reims. Il y enseigne et est le secrétaire de l'archevêque, Adalbéron. En 983, il est nommé abbé de Bobbio, en Italie, par l'empereur Otton II. Cette abbaye était célèbre pour sa vaste bibliothèque et Gerbert en

¹Pierre RICHIÉ. *Gerbert d'Aurillac : Le Pape de l'an Mil*. Paris : Fayard, 1987.

²Carlo BIANCHINI et Luca J. SENATORE. "Gerbert of Aurillac (c. 940/1003)". In : *Distinguished Figures in Descriptive Geometry and Its Applications for Mechanism Science : From the Middle Ages to the 17th Century*. Sous la dir. de Michela CIGOLA. History of Mechanism and Machine Science. Cham : Springer International Publishing, 2016, p. 33-51. URL : https://doi.org/10.1007/978-3-319-20197-9_2 (visité le 25/06/2022).

³RICHIÉ, *Gerbert d'Aurillac*, p. 18.

profite pour continuer ses recherches. Néanmoins, les circonstances politiques difficiles qui entourent l'abbaye (les pressions des nobles locaux sont très fortes) le forcent à quitter Bobbio peu après la mort de l'empereur en 983. Il retourne à Reims, tout en restant abbé de Bobbio, et reprend son rôle d'écolâtre et de secrétaire de l'archevêque.

Il ne reste néanmoins pas le nez dans ses livres et devant ses élèves : de fait, il se mêle de politique. Ses lettres montrent qu'il a eu un impact significatif sur l'élection d'Hugues Capet au trône français, aux détriments des Carolingiens qui régnaient jusque-là. À la mort de l'archevêque, en 989, il espère être nommé archevêque par le roi. Néanmoins, ce dernier lui préfère un bâtard carolingien, Arnoulf. Gerbert, déçu, reprend son rôle de secrétaire, cette fois-ci auprès d'Arnoulf. Néanmoins, les intrigues politiques continuent et vont donner tort à Hugues Capet : Arnoulf le trahit en se ralliant à son oncle Charles de Lorraine qui espère reprendre le trône français. Charles est néanmoins vaincu et Arnoulf finit par être déposé de son archevêché lors du concile de Saint-Basle, en 991, dans lequel Gerbert joue un rôle prééminent.

Au terme du concile, Gerbert est élu archevêque de Reims. Depuis le concile de Saint-Basle, lors duquel il s'est opposé à la papauté, il est en conflit permanent avec celle-ci et est finalement démis de ses fonctions en 997 par le pape Grégoire V. C'est néanmoins pendant son archiépiscopat qu'il rencontre Otton III, empereur depuis la mort de son père en 983 mais encore mineur et sous la régence de sa mère puis de sa grand-mère. Une rencontre significative puisqu'après son départ de Reims, il devient le précepteur du jeune empereur qui le nomme archevêque de Ravenne en 998 et, enfin, pape en 999.

En tant que pape, il prend le nom de Sylvestre II et, avec son ancien élève, cultive le rêve de refonder l'empire romain : Otton III est le nouveau Constantin et Gerbert, le nouveau Sylvestre. Ils sont néanmoins chassés de Rome par une révolte en 1001, et Otton meurt très jeune en 1002. Gerbert le suivra dans la mort à peine une année plus tard, en 1003.

Gerbert a été un auteur prolifique, même si de certains ouvrages semblent lui avoir été attribués à tort. C'est d'ailleurs l'objectif de nos recherches d'affirmer ou infirmer cela. Pour la liste de ses œuvres, je me base sur la tableau réalisé par Marco MOSTERT en 2000⁴ :

- Dialectique : *De rationali et ratione uti*
- De nombreuses lettres

⁴Flavio G. NUVOLONE, éd. *Gerberto d'Aurillac da abate di Bobbio a papa dell'anno 1000 : atti del congresso internazionale, Bobbio, Auditorium di S. Chiara, 28-30 settembre 2000...* Archivum Bobiense. Studia. Bobbio : Associazione culturale Amici di Archivum Bobiense, 2001, p. 426-431.

- Droit canonique : actes du concile de Saint-Basle
- Des chartes pendant qu'il était pape
- Traités scientifiques : Gerbert s'est illustré en arithmétique, géométrie, musique et astronomie, au travers d'ouvrages et de lettres. Le *Liber de astrolabio*, de son nom médiéval *De utilitatibus astrolabii* se trouve dans cette catégorie.
- *Sermo de informatione episocoporum*

3.2 La stylométrie

Cette section est basée sur les articles de D. HOLMES⁵ et de E. STAMATATOS⁶

Je n'ai malheureusement pas eu l'opportunité d'étudier la question de la stylométrie de façon aussi complète que je l'aurais voulu. Je vais néanmoins tenter de donner une brève présentation et un état des lieux de la question.

La stylométrie a pour objectif d'analyser le style d'un texte grâce à des calculs statistiques : elle part du postulat que chaque auteur a son propre style, que le distingue de tous les autres. La façon concrète d'effectuer ces calculs est source à débat et il existe une large bibliographie qui traite de la question. Mon but sera de l'étudier pour alimenter avec des considérations méthodologiques les observations que je pourrai effectuer lors de mes propres analyses.

Une des positions dominantes, illustrée par M. KESTEMONT dans son article de 2014,⁷ est de considérer que ce sont les mots invariables (aussi appelés "foncteurs") qui reflètent le mieux le style d'un auteur. En effet, ce sont des mots qui sont employés inconsciemment et qui, par définition, apparaissent peu importe le genre des textes. Cette méthode est néanmoins améliorable, comme le soutient l'auteur de l'article : elle est naturellement moins efficace sur des langues qui utilisent moins ces mots invariables, comme le latin. Ainsi, et c'est aussi mon opinion, lors de l'analyse stylométrique d'une langue flexionnelle, il convient de prendre en compte plus que les mots

⁵David I. HOLMES. "The Evolution of Stylometry in Humanities Scholarship". In : *Literary and Linguistic Computing* 13.3 (1998), p. 111-117. URL : <https://academic.oup.com/dsh/article/13/3/111/933269> (visité le 09/01/2022).

⁶Efstathios STAMATATOS. "A Survey of Modern Authorship Attribution Methods". In : *Journal of the American Society for Information Science and Technology* 60.3 (2009), p. 538-556. URL : <https://onlinelibrary.wiley.com/doi/10.1002/asi.21001> (visité le 05/01/2022).

⁷Mike KESTEMONT. "Function Words in Authorship Attribution. From Black Magic to Theory?" In : *Proceedings of the 3rd Workshop on Computational Linguistics for Literature (CLFL)*. Proceedings of the 3rd Workshop on Computational Linguistics for Literature (CLFL). Gothenburg, Sweden : Association for Computational Linguistics, 2014, p. 59-66. URL : <http://aclweb.org/anthology/W14-0908> (visité le 05/01/2022).

invariables mais aussi la richesse morphosyntaxique de la langue (par exemple l'usage des cas).

3.3 Workflow : lemmatiser et tokeniser un texte

Pour illustrer la méthode utilisée pour passer d'un texte enregistré sous format brut à un texte traité et prêt à être utilisé dans une analyse stylométrique, j'ai décidé de prendre un extrait de la profession de foi que Gerbert a dû faire avant d'être élu archevêque de Reims et qui nous est parvenue parce que son auteur l'a retranscrite dans sa lettre 180.⁸

3.3.1 Texte brut

Incarnationem divinam non in Patre, neque in Spiritu sancto, sed in Filio tantum credo, ut, qui erat in divinitate Dei Patris Filius, ipse fieret in homine matris filius, Deus verus ex patre, homo verus ex matre;

Le texte est nettoyé : il faut retirer le plus d'erreurs d'OCR possible pour rendre l'analyse stylométrique plus efficace. Certes, noyée dans la multitude, l'une ou l'autre erreur n'aurait pas beaucoup d'impact, mais des erreurs systématiques dans la reconnaissance des caractères en auraient. Pour traiter un grand volume de données, il est possible d'utiliser des expressions régulières qui permettent d'automatiser la reconnaissance des chaînes de caractères concernées et leur modification. Cette étape n'a toutefois pas été nécessaire ici, puisque le texte était déjà très propre.

3.3.2 Texte tokenisé

Pour passer d'un texte brut à un texte tokenisé, c'est-à-dire dont les mots (quand bien même ce concept est très flou en linguistique) sont séparés par un retour à la ligne, j'utilise le script Perl `tokenizelatin-6.pl` (hébergé sur Glossaria). En quelques instants, ce script produit un nouveau fichier `.txt` où se trouve le texte tokenisé.

Incarnationem
diuinam
non
in
Patre

⁸RICHE, *Gerbert d'Aurillac*, p. 137.

,
neque
in
Spiritu
sancto
,
sed
in
Filio
tantum
credo
,
ut
,
qui
erat
in
diuinitate
Dei
Patris
Filius
,
ipse
fieret
in
homine
matris
filius
,
Deus
uerus
ex
patre
,
homo
uerus
ex

matre
;

Il est intéressant de souligner que les signes de ponctuation, en l'occurrence les virgules et points-virgules, sont considérés comme des mots et par conséquent ont droit à leur propre passage à la ligne.

3.3.3 Texte lemmatisé

Il est ensuite nécessaire de faire appel au logiciel TreeTagger, qui analyse chaque ligne du fichier tokenisé et identifie la nature du mot et retourne son lemme (la forme dictionnaire) ainsi que sa traduction. TreeTagger n'a pas été développé pour le latin médiéval mais est personnalisable. C'est ainsi que, de nouveau sur le site Glossaria, se trouvent des paramètres adaptés au latin médiéval qu'il suffit de renseigner à TreeTagger.

N'étant pas familier avec l'invite de commande, qui est la façon dont TreeTagger est normalement utilisé, j'ai décidé d'utiliser une interface graphique (figure 1) qui me permet de facilement contrôler tout un ensemble de paramètres.

Incarnationem SUB incarnatio incarnation, millésime selon l'ère de l'Incarnation
diuinam QLF diuinus divin, prophétique, excellent
non ADV non ne... pas, ne... point, non
in PRE in dans, en, sur, selon, envers, pour, contre, à propos de, vu, par, avec
Patre SUB pater père
, PON ,
neque CON nec et ne... pas, non plus, ne...pas ũ nec... nec : ni... ni ũ nec non
(necnon) : et aussi, et
in PRE in dans, en, sur, selon, envers, pour, contre, à propos de, vu, par, avec
Spiritu SUB spiritus souffle, inspiration, esprit, Saint-Esprit, être surnaturel
sancto QLF sanctusl saint, sacré
, PON ,
sed CON sed mais
in PRE in dans, en, sur, selon, envers, pour, contre, à propos de, vu, par, avec
Filio SUB filius fils, enfant, disciple
tantum ADV tantum autant, tellement, seulement, tout au plus. ũ tantum ut : à
condition que, pourvu que

credo VBE credo2 confier, croire, se fier à, décider de, faire confiance
 , PON ,
 ut CON ut comment, comme, combien ũ (+ indic.) : de même que, quand, vu
 que ũ (+ subj.) : que, à savoir que, pour que, de sorte que, tellement que, si bien
 que, à supposer que ũ (+ part. futur) : avec l'intention de, comme pour
 , PON ,
 qui PRO qui2 qui, quel, lequel, quelque, quelqu'un|quis qui, que, quelqu'un,
 quelque
 erat VBE sum être, exister, signifier ũ est : il est possible, il s'agit de
 in PRE in dans, en, sur, selon, envers, pour, contre, à propos de, vu, par, avec
 diuinitate SUB diuinitas divinité, perfection, théologie
 Dei SUB deus dieu, idole, seigneur
 Patris SUB pater père
 Filius SUB filius fils, enfant, disciple
 , PON ,
 ipse PRO ipse2 lui-même, elle-même
 fieret VBE fio se produire, se rencontrer, devenir, être créé
 in PRE in dans, en, sur, selon, envers, pour, contre, à propos de, vu, par, avec
 homine SUB homo être humain, homme ũ de corpore (capite) : dépendant
 non-libre
 matris SUB mater mère, supérieure, cathédrale|matre déesses tutélaires d'une
 contrée
 filius SUB filius fils, enfant, disciple
 , PON ,
 Deus SUB deus dieu, idole, seigneur
 uerus QLF uerus vrai, raisonnable
 ex PRE ex hors de, à partir de, après, en (matière), selon
 patre SUB pater père
 , PON ,
 homo SUB homo être humain, homme ũ de corpore (capite) : dépendant
 non-libre
 uerus QLF uerus vrai, raisonnable
 ex PRE ex hors de, à partir de, après, en (matière), selon
 matre SUB mater mère, supérieure, cathédrale|matre déesses tutélaires d'une
 contrée
 ; PON ;

Le fichier obtenu reprend le contenu du fichier tokenisé (le premier mot de chaque ligne). Le deuxième élément, en majuscules, est la nature du mot : ponctuation, substantif, verbe... Le troisième élément est le lemme que TreeTagger a reconnu ainsi que les traductions de ce lemme. Pour l'analyse stylométrique, la traduction n'est pas pertinente, mais le lemme l'est énormément : deux mots fléchis à des cas différents, par exemple, seront reconnus comme étant un seul et même mot. Certes, ce n'est pas pertinent pour les mots invariables, qui seront la première angle d'attaque que j'adopterai, mais cela permettra de diversifier les types d'analyses et d'identifier les plus efficaces.

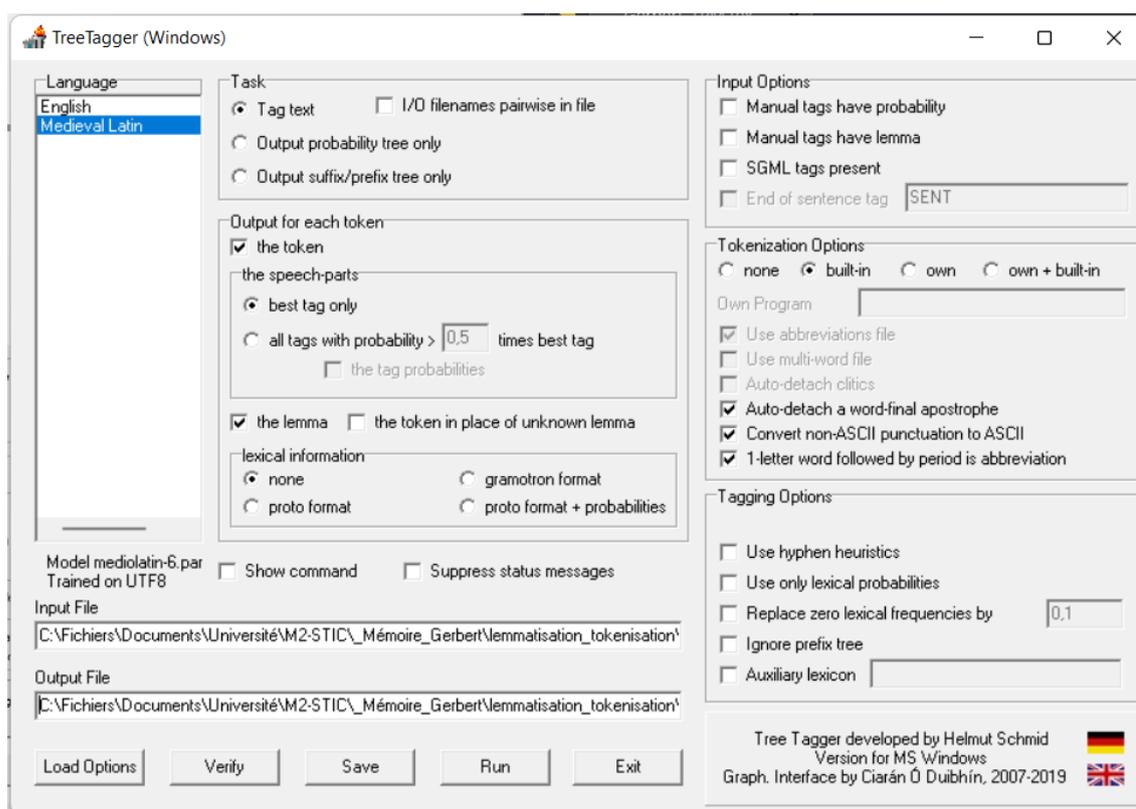


FIG. 1 : L'interface graphique (non-officielle) du logiciel TreeTagger

4 Conclusion

Le travail est encore loin d'être terminé, mais il est désormais plus aisé de voir dans quelle direction se diriger, éclairé par une contextualisation tant méthodologique qu'historique du sujet.

4.1 Table des matières provisoire du mémoire

La table des matières est vouée à être changée avec l'avancée de mes recherches. Il est néanmoins limpide qu'elle comportera les éléments suivants :

1. Contextualisation historique de Gerbert d'Aurillac
2. Méthodologie : stylométrie et paramètres d'analyses
3. Analyse stylométrique d'un texte dont l'attribution à Gerbert ne fait aucun doute, par exemple son corpus de lettre
4. Analyse stylométrique du *Liber de astrolabio* et du *Sermo de informatione episcoporum*
5. Conclusions méthodologiques et historiques

4.2 Plan de travail

Les prochaines étapes sont évidentes : d'abord, traiter l'ensemble des textes que je désire analyser pour avoir des fichiers lemmatisés prêts à l'emploi; ensuite, effectuer les analyses stylométriques à proprement parler, en prenant en compte le temps d'initiation que cela demandera; en parallèle, compléter la bibliographie en perfectionnant ma compréhension des enjeux autour des différentes méthodes d'analyse stylométrique; enfin, combiner l'ensemble des résultats et rédiger.

Voici comment j'envisage ce calendrier, qui devra bien entendu être adapté selon les circonstances de l'année prochaine :

Septembre Sélection et traitement des textes

Octobre – Décembre Premiers pas sur R et avancée dans la bibliographie

Février – Mars Suite des analyses stylométriques dans R

Mars – Mai ou Août Rédaction