

Formalisation des variantes à des fins computationnelles : vérification de l'hypothèse expérimentale sur un texte occitan

Maria Sofia Corradini

1. Enregistrement et classification des variantes et représentation graphique du *stemma codicum*

Le travail ici décrit s'insère dans le cadre de la philologie assistée par ordinateur¹ et constitue le *case study* pour la mise au point d'une des parties dont se compose la "Workstation" projetée à l'ILC-CNR de Pise par Andrea Bozzi,² qui développe un système de philologie numérisée.³

L'objectif visé avec le module *Enregistrement des variantes et représentation graphique du stemma codicum* était de produire un outil pour la mémorisation des variantes également capable de représenter graphiquement les relations de ressemblance et de dissemblance entre les manuscrits.⁴ On a donc essayé de concevoir et de réaliser un

¹ Abstraction faite de la déjà incommensurable bibliographie relative à l'application de l'informatique aux sciences humaines (pour laquelle, toutefois, je renvoie au moins aux répertoires de [LAN 88], de [ADA 94], de [SAB 94] et, seulement pour la période médiévale, de [BOU 82]), celle qui concerne plus spécifiquement le rapport entre informatique et critique du texte dans la dernière décennie se trouve surtout dans [FAU 91], [SHI 96], [MAR 96] et [MOR 99].

² Il n'est pas possible d'exposer dans cet article le programme du travail qu'il a coordonné, il y a quelques années, soit à l'intérieur des activités de l'Istituto di Linguistica Computazionale, soit dans le domaine des projets nationaux et européens. Voir, par exemple, le projet BAMBI LIB-3114 (*Better Access to Manuscripts and Browsing of Images*), conduit en collaboration avec de nombreuses bibliothèques, industries et centres de recherche et FAD (*Fondi e Archivi Digitali*), en phase de réalisation grâce à des financements du Ministero dei Beni Culturali. On peut lire des informations plus détaillées relatives au BAMBI dans [BOZ 97], [BOZ 99] et [CAL 99].

³ Il s'agit d'un système de "Philologie computationnelle" : ce terme a une valeur différente par rapport au plus connu "Philologie électronique" qui, même si pas toujours, se réfère à la production d'éditions critiques avec des supports magnétiques ou numérisés. La "Philologie électronique" a comme instruments de choix les systèmes hyper-textuels ou multimédias, tandis que la "Philologie computationnelle" ou "Philologie numérisée" se sert de procédures et de logiciels spécialisés pour aborder, dans des milieux intégrés, chaque phase du travail d'édition, à partir de la lecture des sources jusqu'à la production du *stemma*. Le terme "Philologie computationnelle" a été employé dans ce sens pour la première fois par Bozzi [BOZ 93].

⁴ Une anticipation de l'étude de faisabilité d'un module informatique pour la critique du texte se trouve au paragraphe 2 (*The textual criticism system tool*) de [BCR 02]. Le développement pratique de cette

support pour le spécialiste qui veut dresser une édition critique, lequel, en même temps, ne le remplace dans aucune des phases de décision les plus importantes pendant son travail de *restitutio textus*. À mon avis, en effet, on peut utiliser avec profit les nouvelles technologies seulement si les ordinateurs et les logiciels sont à même de résoudre les problèmes spécifiques sans avoir la présomption de se substituer à l'expérience humaine.⁵ Le philologue, en outre, doit avoir la possibilité de choisir la fonction qu'il veut sans connaître les éléments techniques et sans utiliser des procédures trop enchevêtrées de codages, mais en disposant de méthodes visuelles ; de plus, on ne doit pas lui imposer une participation trop complexe ou un travail préparatoire trop long.⁶

Le système qu'on va décrire permet :

- a) d'établir la corrélation entre les mots du texte et ceux de l'apparat critique et de produire le texte établi par l'éditeur ;
- b) de proposer des hypothèses de liens de parenté entre les manuscrits à partir des données insérées dans l'apparat critique ;
- c) d'interpréter les relations de parenté, les classant en hypothèses de dépendance jusqu'à la production du *stemma codicum*.

Voyons maintenant de façon plus précise les trois objectifs précédemment définis :

hypothèse de formalisation des variantes a été l'objet d'une thèse de maîtrise soutenue à l'Université de Sienna (1998-1999) par Duccio Fabbri.

⁵ Le scepticisme de la part des philologues vis-à-vis de la "New Philology" dérive justement de la place excessive que les méthodologies informatiques consacrent en général aux machines. On peut lire quelques opinions à ce sujet, à l'intérieur du domaine de la philologie romane, dans *Speculum* 65 (1990). Plus récemment, A. Varvaro et Ph. Ménard ont manifesté des réserves envers les nouvelles technologies quand elles ne tendent ni au rétablissement généalogique ni au rétablissement du texte d'origine, mais elles se bornent à structurer de différentes façons les données qui constituent la *varia lectio* ; voir [GLE 97]. Au-delà des tentatives de rétablissement automatique du *stemma codicum*, conçues et expérimentées pour *mainframes*, ainsi que les procédures de Dom Quentin et de Zarri (voir [QUE 26], [ZAR 68], [ZAR 69], [ZAR 77]), il y a eu d'autres tentatives de formaliser l'approche traditionnelle. C'est le cas, par exemple, de l'emploi d'algorithmes par Froger et par Buneman ; celui-là élabore une procédure dont les principes étaient nés dans le domaine des sciences biologiques (voir [FRO 64], [FRO 65], [FRO 68], [BUN 74]). Il y a quelque temps que les technologies modernes et les applications de l'informatique à l'élaboration du texte ont produit d'autres résultats, employant des systèmes différents ; un compte rendu des applications informatiques à la critique du texte mis à jour au 1989 se trouve dans [NAJ 89]. De plus récentes applications sont, par exemple, les travaux relatifs au *Canterbury Tales Project*, coordonnés par P. Robinson, qui a proposé une méthode pour établir les relations entre manuscrits selon des principes qui dérivent de l'analyse cladistique ; en particulier, voir [ROB 93]. La même procédure de base a été employée par Salemans [SAL 00]. En ce qui concerne l'emploi de la technologie multimédia pour l'édition critique des textes des variantes, voir [LEB 99].

⁶ Je me réfère, par exemple, à des méthodologies qui, pour être opérationnelles, nécessitent de disposer d'autant de transcriptions que les manuscrits d'un même texte. Le plus connu des systèmes de ce type, et sûrement l'un des premiers qui a dépassé la phase du prototype et qui est en effet adopté, c'est le TUSTEP (Tübinger System von Textverarbeitungs-programmen). Il a été réalisé par Wilhelm Ott au Zentrum für Datenverarbeitung de l'Université de Tübingen. Cette méthode, qui a l'avantage de paraître avec un résultat définitif pour imprimer le texte, les apparats et les index en photocomposition, a été employée, entre autres, par Hans Walter Gabler pour l'édition de l'*Ulysse* de Joyce [GAB 84] et par Heinrich Schepers pour l'édition de quelques œuvres de Leibniz [SCH 99].

a. *Corrélation entre les mots du texte et les informations de l'apparat critique et possibilité de produire le texte établi par l'éditeur.* Comme dans la pratique de la philologie traditionnelle, cette méthode est basée sur la comparaison systématique entre toutes les leçons des différents manuscrits par rapport à un témoin choisi comme base de la collation. Toutefois, dans son travail de comparaison des variantes des divers manuscrits dont il a fait la *recensio*, le spécialiste se trouve bientôt devant un nombre considérable d'annotations destinées à établir la généalogie de l'œuvre, même s'il analyse peu de témoins à la taille réduite ; au contraire, si les archives d'informations textuelles qui étaient traditionnellement sur papier sont introduites dans une base de données informatisée, il peut gérer avec plus de facilité⁷ et visualiser à n'importe quel moment, de façon rapide et intuitive, les caractéristiques saillantes de chaque manuscrit.

Un autre avantage est représenté par le fait que l'ordinateur peut automatiquement engendrer les textes de tous les manuscrits collationnés en employant le seul qui a été transcrit en entier, auquel chaque fois un logiciel remplace les différentes *lectiones* rapportées par les autres manuscrits et qu'on a enregistrées dans des champs spéciaux de l'apparat. Il s'ensuit que si celui-ci ne contient pas seulement un champ où insérer les leçons de chaque manuscrit, mais aussi un autre où le philologue peut mémoriser ses choix d'édition, on peut à tout moment produire le texte établi, dans lequel il est plus facile d'introduire des modifications dues à des réflexions apparues en cours de travail.

b. *Hypothèse de liens de parenté entre les manuscrits.* Dans le même temps, ce système permet d'accéder à des informations plus spécifiques et on peut aussi arriver à la formulation d'hypothèses de *stemma codicum* (voir le point c). Puisque le deuxième but de la recherche est de représenter graphiquement les relations qui existent entre les manuscrits, c'est-à-dire les distances mutuelles qui sont dues à une ressemblance plus grande ou plus petite, on demande à l'utilisateur d'exprimer un jugement codifié, analysant comparativement les leçons des manuscrits dont on a fait la *recensio*. À ce propos, il convient de préciser que, à l'intérieur de la dichotomie entre procédé inductif et procédé déductif dans la construction des relations généalogiques des manuscrits,⁸ la méthode que j'ai projetée se place dans la ligne inductive post-lachmannienne : en effet, je ne donne pas à l'ordinateur une série de caractéristiques avec lesquelles la machine a les informations et la connaissance pour déterminer et sélectionner ce qui dans le texte est reconnu comme "variante". Au contraire, c'est le philologue même qui : 1°) établit une grille de caractéristiques pour classer "toutes" les variantes d'apparat ; 2°) juge quelle est la classification la plus appropriée des variantes dans la tradition manuscrite qu'il est en train d'analyser. Contrairement à d'autres domaines où un critère de classification pour organiser les données prend une valeur si générale qu'il peut être considéré comme un standard, dans un système informatisé de critique textuelle, un critère de classification des variantes

⁷ L'affirmation « Senza dubbio il ricorso a strumenti informatici rende più agevole e spedito molto del lavoro filologico preliminare o collaterale » est exprimée, par exemple, par Alfredo Stussi : [STU 98], p. 43. Voir aussi [SEG 99].

⁸ À cet égard voir [SAL 99], en particulier p. 116.

ne peut pas être universellement valide, mais il l'est seulement d'un point de vue structurel. En effet, les classes sont étroitement connexes soit aux caractéristiques de chaque manuscrit, soit à celles de la tradition manuscrite dans l'ensemble : au-delà des variantes substantielles, les variantes de forme peuvent prendre une valeur complètement différente par rapport au contexte. L'affirmation de Margot Van Mulken : « L'important c'est de remarquer que chaque tradition manuscrite impose sa propre typologie de variance, et que, bien sûr, chaque philologue fera sa propre typologie »⁹ paraît, donc, tout à fait pertinente.

La méthode se base sur la détermination d'un index de dissemblance ou d'affinité entre tous les couples de variantes : une autre particularité qui la distingue des précédentes approches basées sur des critères statistiques, c'est le fait que le philologue n'établit pas seulement la typologie et, par conséquent, la classification des variantes, comme on l'a dit ci-dessus, mais il détermine aussi le degré de rapport entre les variantes. Cette procédure paraît bien plus fiable "philologiquement" par rapport à celles qui comparent les divers témoins d'une manière automatique, et donc en dehors du *iudicium*. Si dans l'ecdotique traditionnelle « pertengono al *iudicium* [...] le trascrizioni, le collazioni, l'individuazione degli errori, la messa a punto formale »,¹⁰ dans la méthode computationnelle décrite on propose d'étendre le *iudicium* aussi bien à la détermination de la typologie de la *varia lectio* et du différent degré de relation entre les variantes.

De cette façon, les données en entrée du système informatisé sont représentées non seulement par des chaînes de caractères (le texte des variantes), mais également par des classes de variantes qui sont organisées selon un poids numérique qui qualifie les différentes classes de ressemblances et de dissemblances. J'ai avancé, donc, la possibilité d'attribuer à chaque classe un poids numérique de sorte que l'ordinateur puisse exécuter des calculs sur les données qui dérivent de l'apparat critique. Je souligne que la responsabilité soit de la classification soit de l'attribution des poids est entièrement celle du philologue, tandis que la machine ne fait qu'exécuter des calculs, produisant des résultats. Suivant cette approche, on donne à l'utilisateur la possibilité de "peser" les divergences selon une échelle de valeurs établie initialement par lui-même. Si Bédier affirmait la nécessité de « peser les variantes » pour établir les liens aux bas niveaux du *stemma*,¹¹ notre approche computationnelle requiert du philologue de « peser les divergences entre les variantes » de toute la tradition manuscrite.

L'échelle comprend soit des valeurs négatives soit des valeurs positives pour exprimer, respectivement, les ressemblances et les différences entre les leçons des différents témoins, parce que, comme on l'a dit, l'un des objectifs que le module se propose est de formuler des hypothèses sur les liens de parenté entre les manuscrits. Plus la somme des valeurs est faible, moins il y aura de points de divergence entre les manuscrits, et, donc, plus petite sera la distance entre les manuscrits. Au

⁹ Voir [MUL 99], p. 105.

¹⁰ Voir [SEG 99], p. 11.

¹¹ Voir [BED 28].

contraire, plus élevé sera la somme des valeurs attribuées à chaque couple de variantes, plus grande sera la distance spatiale qui sépare les manuscrits auxquels ils appartiennent (voir Fig. 1).

c. *Production du stemma codicum*. La dernière activité, qui est au sommet des actions exécutables pendant le dialogue entre le philologue et la machine, est représentée par la production d'une ou plusieurs propositions de *stemma codicum*, à partir justement des informations structurées de l'apparat, grâce auxquelles le système a pu mettre en évidence les liens de parenté entre les manuscrits. Cela ne signifie pas que l'ordinateur produit un *stemma codicum* à titre autonome ; en réalité, il aide le spécialiste, lui indiquant les conséquences des choix d'édition que lui-même a effectués dans l'apparat critique, sur la base de la tradition manuscrite. On a été guidé par l'idée de pouvoir donner à l'ordinateur les données quantitatives qu'il a extraites des poids attribués à chaque couple de variantes, de façon à pouvoir interpréter les sommes obtenues au moyen d'une méthode de représentation graphique tridimensionnelle : de cette manière on concrétise l'image mentale de la tradition manuscrite que le philologue bâtit dans sa tête, au fur et à mesure que son travail critique avance et que les données à sa disposition s'accumulent. En d'autres termes, les sommes fournies par le système sur la base des valeurs attribuées aux variantes d'apparat, transformées en distances graphiques à eux proportionnelles, rendent possible la distribution des manuscrits dans un espace tridimensionnel. Parmi les solutions généalogiques possibles, le philologue sélectionne la plus réaliste ; chaque décision peut être sujette à des changements d'avis et l'arbre généalogique se modifiera automatiquement (voir Fig. 2). Il est important d'ajouter, en outre, que de cette manière on ne contraint pas le philologue à rétablir le texte d'un archétype perdu, quand il ne peut pas ou il ne veut pas le faire, c'est-à-dire, quand il est persuadé de l'opportunité de « valoriser la diversité des textes au lieu de l'inscrire dans la négativité d'un processus de dégradation à partir d'archétypes idéaux à jamais perdus ». ¹² À mon avis, cette technologie se révèle particulièrement intéressante parce qu'elle peut présenter de façon efficace toute la *mouvance* ¹³ qu'un texte a subi au cours de sa tradition.

2. Vérification de l'hypothèse expérimentale sur un texte occitan

Dans cette phase expérimentale de la recherche, qui sera ensuite soumise à une plus ample vérification, concernant tant l'extension de la tradition manuscrite que sa complexité intrinsèque, le choix s'est arrêté pour le moment sur une œuvre transmise par quatre témoins, que j'avais déjà analysée et éditée dans le passé selon les

¹² Voir [LEB 99] p. 127. La conviction que soit les différentes versions d'un texte soit les diverses rédactions produites par l'auteur même ont leur propre dignité a été exprimée, par exemple, par B. Cerquiglini [CER 89] et par J. McGann [McG 91].

¹³ La notion de *mouvance* a été formulée par P. Zumthor dans son étude sur la poésie médiévale française [ZUM 72].

méthodes traditionnelles.¹⁴ Le but de cette expérience est double : d'un côté, évaluer les aspects connexes à l'approche computationnelle des mêmes données ; de l'autre, comparer les résultats obtenus grâce aux deux différentes procédures.

L'œuvre en question est la vulgarisation occitane d'un herbier latin, le *De viribus herbarum* du médecin français Odo de Meudon, connu aussi sous le pseudonyme de Macer Floridus, qui vécut dans la première moitié du XI^e siècle. Si les deux mille hexamètres dont se compose le célèbre poème au sujet médico-pharmaceutique nous attestent la vitalité de la poésie métrique à côté de la poésie rythmique pendant l'époque scolastique, les nombreuses traductions et réélaborations faites ensuite dans différents idiomes vulgaires, romans ou non,¹⁵ nous rendent témoignage de la notoriété de cet herbier. Les quatre témoins occitans qui nous sont parvenus sont contenus dans les manuscrits suivants :¹⁶

– Princeton, Garrett 80, composé de deux parties distinctes, toutes les deux du XIV^e siècle. Les vertus de *arcemisa*, *alambroze*, *eixens*, *ortigua*, *alb*, *plantage*, en prose, sont contenues aux folios 15v-21v.

– Auch, Archiv. Dép. Du Gers I 4066, du XV^e siècle. *Lo livre que fec lo mege Arcemis de las vertutz de las erbas*, en prose, est aux folios 1r-14v.

– Chantilly, Musée Condé 330, du XV^e siècle. Une rédaction de l'*Erbario*, en vers, se trouve aux folios 46r-52v et une autre, en prose, aux folios 53v-59v.

Plusieurs éléments relatifs soit à l'aspect thématique, soit au linguistique, permettent de déterminer les liens de parenté entre les rédactions et, parfois, d'entrevoir aussi les changements qui ont au fur et à mesure caractérisé le texte original dans son passage d'un milieu à l'autre. Le texte de A, qui a une certaine originalité de contenu, confirme cette impression même pour son respect linguistique, parce qu'il montre des traits qui sont absents dans P, Ca et Cb. Parmi ces derniers, en outre, Ca et Cb sont les manuscrits les plus étroitement rapprochés.¹⁷ Le schéma qui en découle est celui de la Fig. 3.

La méthode computationnelle étudiée et adoptée pour l'analyse de ce texte est basée sur la comparaison systématique entre les leçons des quatre différents manuscrits. Dans les cas où au moins l'un diffère des autres, j'ai enregistré les leçons de tous les manuscrits les comparant deux par deux. De cette façon s'est vérifiée la correspondance de chaque leçon avec toutes les autres. L'archive des couples des variantes qui s'est ainsi constitué a ouvert la possibilité de rédiger une typologie en 9 classes : chaque couple de variantes a été classé avec sa classe d'appartenance (voir un exemple à la Fig. 4).

¹⁴ Voir [COR 91].

¹⁵ Entre le XIII^e et le XV^e siècle il y eut des traductions en danois, en allemand, en anglais. Quant au milieu roman, il y a des versions en italien, en français, en castillan, en aragonais, en catalan, en occitan. On pense aussi qu'une version hébraïque réduite vient également de la zone occitane. Voir [BOM 00].

¹⁶ Dorénavant les quatre rédactions occitanes de l'herbier seront respectivement indiquées comme P, A, Ca, Cb. Pour l'édition du texte latin voir [CHO 32].

¹⁷ Les éléments qui caractérisent les quatre versions de l'herbier appartiennent dans l'ensemble au Languedoc. À l'intérieur de cette zone, alors que le manuscrit A présente des traits qui se trouvent aux alentours de Moissac, les trois autres manuscrits sont caractérisés par des phénomènes typiques du sud de la Haute-Garonne. Voir [COR 97], en particulier pp. 76-100.

En ce qui concerne le texte choisi, deux sortes de données sont considérées comme fondamentales pour établir les relations de parenté entre les quatre rédactions : celles qui se réfèrent au contenu des recettes transmises et celles qui appartiennent au domaine linguistique. Du premier type font partie, par exemple, les modalités de formulation du précepte, la présence ou l'absence de tous les ingrédients ou des pratiques qui la rendent efficace, l'adjonction de particuliers. Quant aux données linguistiques, comme l'un des buts que l'étude de cette tradition manuscrite se propose est la localisation des différentes rédactions, apparaît comme plus significative, par exemple, une différence entre deux usages lexicaux qui viennent de deux milieux dialectaux différents, plutôt qu'une différence de type morpho-syntaxique, si les deux réalisations sont équivalentes dans ces mêmes milieux.

À cause des caractéristiques du texte pris en examen, donc, la typologie considérée comme la plus indiquée pour satisfaire le rapport entre les couples est représentée par les classes suivantes, dont je dresse une liste selon un ordre croissant de ressemblance :

1. présence / absence = le fait qu'entre deux manuscrits l'un des deux ne contient pas une expression qui, au contraire, caractérise l'autre, est considéré comme une grande divergence ;
2. différence lexicale sémantique = elle codifie la situation où deux manuscrits présentent deux différents lemmes de sens différent. Cette catégorie peut inclure même les cas de variantes indifférentes, de *lectio singularis*, de méprises ;
3. différence lexicale asémantique = elle codifie les lemmes qui, bien que synonymes, sont l'expression de *scriptae* différentes ;¹⁸
4. différence graphico-phonétique = cette codification est attribuée aux différences qui, dans le domaine linguistique, sont relatives aux diverses formes avec lesquelles un lemme se réalise, qui souvent sont produites par des *scriptae* différentes ;
5. différence morpho-syntaxique = elle codifie les situations dans lesquelles une leçon a une variation d'ordre morphologique ou syntaxique ;
6. différence lexicale insignifiante = il s'agit de synonymes qui n'ont pas d'importance dans la détermination de la *scripta* ;
7. renversement = cette codification est attribuée au renversement des termes qui composent une expression ;
8. identité = elle classe des lectures égales ;
9. identité dans la spécificité = cette codification est attribuée aux cas suivant : quand deux manuscrits ont la même leçon erronée (faute conjonctive) ou quand ils réfèrent à une expression qui manque dans tous les autres.¹⁹

Grâce à cette approche, le philologue a la possibilité de "peser" les diverses situations, en les réunissant dans une échelle de valeurs établie par lui-même. Dans l'hypothèse expérimentale, cette classification va d'une valeur maximale de dissem-

¹⁸ La même typologie (« équivalent sémantique mais lexicologiquement différent dans au moins une autre copie ») a été utilisée par van Emden dans l'analyse de quelques laisses de la *Chanson de Roland* ; [Voir EMD 02], p. 405.

¹⁹ Voir à ce propos l'affirmation de Segre dans [SEG 99], p. 14 : « non si può definire errore solo l'infrazione alla correttezza ; può essere errore anche una innovazione inopportuna rispetto al discorso del testo ».

blance correspondant à +5 (c'est le cas où un manuscrit omet toute une partie du texte) jusqu'à une valeur maximale de ressemblance correspondant à -5 (lorsque deux manuscrits retiennent la même leçon erronée). Quelques variantes suggèrent une possible origine commune des versions examinées (fautes conjonctives), et dans ce cas il est opportun d'employer un index de ressemblance exprimé par des valeurs négatives ; d'autres amènent à penser qu'on se trouve devant des branches diverses de la tradition (fautes séparatives), et elles comportent l'utilisation d'un index de dissemblance exprimé par des valeurs positives.

La combinaison de la typologie avec le poids numérique a été ainsi établie :

- +5 = présence / absence
- +4 = différence sémantique
- +3 = différence lexicale asémantique
- +2 = différence graphico-phonétique
- +1 = différence morpho-syntaxique
- 1 = différence lexicale insignifiante
- 2 = renversement
- 3 = identité
- 4 = (non utilisée pour cet exemple)
- 5 = identité dans la spécificité (faute conjonctive)

Analysant sur cette base les manuscrits, il est possible de classer tous les éléments indispensables pour suivre le parcours du texte et terminer avec la production du *stemma* montré à la Fig. 5, qui s'avère équivalent à celui obtenu avec le système traditionnel.

Je donne ci-après quelques exemples, qui se trouvent réunis dans l'Appendice.

Dans une recette relative à l'artémise et dédiée aux femmes, à côté de l'expression latine (*matrix si saepe*) *fovetur*, les témoins occitans récitent : *perfum* (P) ; *enfuman* (Ca) ; *affuma* (Cb) ; le manuscrit A omet cette expression. Appliquant la typologie plus haut décrite, A se lie à chacun des autres manuscrits par la relation de plus grande dissemblance "présence / absence" (+5), tandis que P, Ca et Cb sont réciproquement unis par la relation "différence lexicale non significative" (-1) : en effet, *enfumar*, *affumar*, *perfumar* sont des variantes lexicales non associables à des zones dialectales particulières, et donc, relativement à notre texte, des variantes indifférentes.

Le texte latin et les versions vulgaires rapportent une autre propriété de l'artémise, utile à ceux qui ont bu de l'opium : *opium nimium si quemquam laeserit haustum* ; (*per home que aia*) *maniat trop o begut* (P) ; (*si alcun a*) *opi begut* (Ca et A) ; (*a home que a*) *begut opi* (Cb). Dans ce cas Ca, A et Cb présentent la même expression, mais les termes qui la composent paraissent dans l'ordre inverse (relation de "renversement") ; P présente, à l'égard de chacun des autres trois manuscrits Ca, Cb et A une relation de "différence sémantique" (+4) parce qu'il formule une idée tout à fait différente.

La *Artemisia abrotanum* est appelé *alambroze* dans P, *brona* et *bronia* respectivement dans Ca et Cb, tandis que A omet la description de ses propriétés. Dans ce cas, ce dernier manuscrit est lié à chacun des trois autres par une relation de dissemblance

+5 (“présence / absence”), P est uni à Ca et à Cb grâce à une relation +3 (emploi d’un terme lexical qui a la même signification mais qui appartient à des milieux géographiques divers, “différence lexicale asémantique”), Ca et Cb montrent une “différence graphico-phonétique” (+2).

Le même simple a la qualité de réussir avec les nerfs. Le texte latin nous informe, en effet, qu’il *inuat nervos*, et les versions occitanes récitent : *profescha a noirir los nervis* (P) ; *en valon los nervis* (Ca) ; *contra los veris* (Cb). Le manuscrit de Princeton et la version rimée de C expriment la même idée, introduisant une variation lexicale dans l’usage verbal (“différence lexicale insignifiante” ; -1) ; au contraire, la version en prose de C fausse complètement le sens, introduisant une faute due vraisemblablement à une mauvaise interprétation de *nervis*, contenu peut-être dans l’antigraphe. La relation de Cb soit avec A soit avec P est donc une “différence lexicale sémantique” (+4).

Voilà une propriété de l’*Artemisia abrotanum* : éloigner les serpents ; à *fugat* latin correspond *fuig* (P), *fuch* (Ca), *fug* (Cb). Les trois versions occitanes sont réciproquement liées par une “différence graphico-phonétique” (+2).

Bibliographie

- [ADA 94] Adamo, G. 1994. *Bibliografia di informatica umanistica*. Roma : Bulzoni.
- [BCR 02] Bozzi A. et M.S. Corradini. 2002. « New trends in philology : a computational application for textual criticism », *Euphrosyne* XXX, pp. 267-85.
- [BED 28] Bédier, J. 1928. « La tradition manuscrite du *Lai de l’Hombre* : réflexions sur l’art d’éditer les anciens textes », *Romania* LIV, pp. 161-96 et 321-56.
- [BOM 00] Bos, G. et G. Mensching. 2000. « Macer Floridus : a Middle Hebrew Fragment with Romance Elements », *The Jewish Quarterly Review* XCI, pp. 17-51.
- [BOU 82] Bourlet, C. et Ch Doutrelepoint et S. Lusignan. 1982. *Ordinateurs et études médiévales : bibliographie I*, Montréal, Publications de l’Institut d’études médiévales de l’Université de Montréal.
- [BOZ 93] Bozzi, A. 1993. « Towards a Philological Workstation », *Revue informatique et statistique dans les Sciences humaines* 29 : 1-4, Université de Liège, pp. 33-49.
- [BOZ 97] Bozzi A., éd. 1997. *Better Access to Manuscripts and Browsing of Images. Aims and results of an European Research Project in the field of Digital Libraries*. BAMBILIB-3114, Bologna : Clueb.
- [BOZ 99] Bozzi, A. 1999. « Verso un sistema integrato di moduli computazionali per la filologia dell’era digitale », dans *Nuovi orizzonti*, pp. 253-85.
- [BUN 74] Buneman, P. 1974. « A characterization of rigid circuit graphs », *Discrete Mathematics* 9, pp. 205-12.
- [BUS 92] Busa, R. 1992. « Half a century of literary computing : towards a ‘new’ philology », *Literary and Linguistic Computing* 7, pp. 69-73.
- [CAL 99] Calabretto, S., J-M. Pinon et A. Bozzi. 1999. « BAMBILIB : système de gestion de manuscrits anciens pour historiens », dans *Document numérique* 2/1998, éd. I. S. Batori, W. Lenders et W. Putschke, Paris : Hermes, pp. 31-50.
- [CER 89] Cerquiglini, B. 1989. *Éloge de la variante*. Paris : Seuil.
- [CHO 32] Choulant, L. 1932. *Macer Floridus, De viribus barbarum*. Leipzig : Sillig, 1832.
- [COR 91] Corradini, M. S. 1991. « Sulle tracce del volgarizzamento occitanico di un erbario latino », *Studi Mediolatini e Volgari* XXXVII, pp. 31-132.
- [COR 97] Corradini, M. S. 1997. *Ricettari medico-farmaceutici medievali nella Francia meridionale*, I. Olschki Editore, Firenze.

- [EMD 02] van Emden, W. 2002. « Le Chef-d'œuvre épique », dans *L'épopée romane*, Actes du XV Congrès international Rencesvals, Poitiers, 21-27 août 2000, Université de Poitiers, pp. 395-412.
- [FAU 91] Faulhaber, Ch. 1991. « Textual criticism in the XXIst century », *Romance Philology* 45, pp. 123-48.
- [FRO 64] Froger, J. 1964. « La collation des manuscrits à la machine électronique », *Bulletin de l'Institut de Recherche et d'Histoire des textes* 13, pp. 135-71.
- [FRO 65] Froger, J. 1965. « La machine électronique au service des sciences humaines », *Diogenes* 52, pp. 108-44.
- [FRO 68] Froger, J. 1968. *La critique des textes et son automatisations*. Paris : Dunod.
- [GAB 84] Gabler, H. W. et alii, éd. 1984. *Ulysses : a critical and synoptic edition*. New York-Londres : Garland.
- [GLE 97] Gleßgen, M.D. et F. Lebsanft, éd. 1997. *Alte und neue Philologie*. Tübingen : Niemeyer.
- [LAN 88] Lancashire, I. et W. McCarty. 1988. *The Humanities Yearbook 1988*. Oxford : Clarendon Press.
- [LEB 99] Lebrave, J.-L. 1999. « L'édition critique au XXI^e siècle », dans *Nuovi orizzonti*, pp. 127-32.
- [MAR 96] Marcos Marín, F. 1996. *El comentario filológico con apoyo informático*. Madrid : Síntesis.
- [McG 91] McGann, J. 1991. « What is critical editing ? », *Text* V, pp. 15-29.
- [MOR 99] Morás, M. 1999. « Informática y crítica textual : realidades y deseos », dans *Filología e informática*, éd. J. M. Blequa et alii, Barcelona : Editorial Milenio i Universitat autònoma de Barcelona, pp. 189-210.
- [MUL 99] van Mulken, M. 1999. « Les changements de parenté dans le Cligès de Chretien de Troyes », dans *Nuovi orizzonti*, pp. 103-14.
- [NAJ 89] Najock, D. 1989. « Computer-assisted reconstruction of texts », dans *Computational linguistics : an international handbook on computer oriented language research and application*, éd. I. S. Batori, W. Lenders et W. Putschke, Berlin : Walter de Gruyter.
- Nuovi orizzonti = I nuovi orizzonti della filologia. Eadotica, critica testuale, editoria scientifica e mezzi informatici elettronici*, Roma : Accademia Nazionale dei Lincei ("Atti dei Convegni Lincei", n. 151),
- [QUE 26] Quentin, dom H. 1926. *Essais de critique textuelle (Eadotique)*. Paris : Picard.
- [ROB 93] Robinson, P. 1993. « Redefining critical editions », dans *The digital word : Text-based computing in the Humanities*, éd. G. P. Landow et P. Delany, Cambridge/Mass. – London : MIT Press, pp. 271-91.
- [SAB 94] Sabourin, C. F. 1994. *Computational character processing (character encoding, input, output, ...). A bibliography*. Montréal : Infolingua.
- [SAL 99] Salemans, B. J. P. 1999. « The old text-genealogical method of Lachmann updated with the help of cladistics and the computer », dans *Nuovi orizzonti*, pp. 115-25.
- [SAL 00] Salemans, B. J. P. 2000. *Building stemmas with the computer in a cladistic, neo-lachmannian way : the case of fourteen text versions of Lanselot van Denemerken*. Nijmegen : Nijmegen University Press.
- [SCH 99] Leibniz, G. W. 1999. *Philosophische Schriften*, herausgegeben von der Leibniz-forschungsstelle der Universität Münster, Berlin : Akademie Verlag.
- [SEG 99] Segre, C. 1999. « Prolusione », dans *Nuovi orizzonti*, pp. 11-17.
- [SHI 96] Shillingsburg, M. J. 1996. *Scholarly editing in the computer age*. Michigan University Press.
- [STU 98] Stussi, A. 1998. *Fondamenti di critica testuale*. Bologna : Il Mulino.
- [ZAR 68] : Zarri, G. P. 1968. « Linguistica algoritmica e meccanizzazione della collatio codicum », *Lingua e stile* 2, pp. 21-40.
- [ZAR 69] Zarri, G. P. 1969. « Il metodo per la recensio di Dom Quentin esaminato criticamente mediante la sua traduzione in un algoritmo per elaboratore elettronico », *Lingua e stile* 4, pp. 162-82.
- [ZAR 77] Zarri, G. P. 1977. « Some experiments on automated textual criticism », *Literary and Linguistic Computing* 5, pp. 266-90.
- [ZUM 72] Zumthor, P. 1972. *Essai de poétique médiévale*. Paris : Seuil.

Appendice

id	Man. 1	Man. 2	Variante 1	Variante 2	Pds	Relation
1	Ca	Cb	enfuman	affuma	-1	Diff. lexicale insign.
2	Ca	P	enfuman	perfum	-1	Diff. lexicale insign.
3	Ca	A	enfuman		5	Prés./absence
4	A	P		perfum	5	Prés./absence
5	A	Cb		affuma	5	Prés./absence
6	P	Cb	perfum	affuma	-1	Diff. lexicale insign.
1	Ca	Cb	ser e matin	ser e matin	-3	Identité
2	Ca	P	ser e matin	lo matin e i ser	-2	Renversement
3	Ca	A	ser e matin		5	Prés./absence
4	A	P		lo matin e i ser	5	Prés./absence
5	A	Cb		ser e matin	5	Prés./absence
6	P	Cb	lo matin e i ser	ser e matin	-2	Renversement
1	Ca	Cb	opi begut	begut opi	-2	Renversement
2	Ca	P	opi begut	maniat trop o begut	4	Diff. sémantique
3	Ca	A	opi begut	opi begut	-3	Identité
4	A	P	opi begut	maniat trop o begut	4	Diff. sémantique
5	A	Cb	opi begut	begut opi	-2	Renversement
6	P	Cb	maniat trop o begut	begut opi	4	Diff. sémantique
1	Ca	Cb	en valon los nervis	contra los veris	4	Diff. sémantique
2	Ca	P	en valon los nervis	profiecha a noirir los nervis	-1	Diff. lexicale insign.
3	Ca	A	en valon los nervis		5	Prés./absence
4	A	P		profiecha a noirir los nervis	5	Prés./absence
5	A	Cb		contra los veris	5	Prés./absence
6	P	Cb	profiecha a noirir los nervis	contra los veris	4	Diff. sémantique
1	Ca	Cb	fuch	fug	2	Diff. graph.-phonet.
2	Ca	P	fuch	fuig	2	Diff. graph.-phonet.
3	Ca	A	fuch		5	Prés./absence
4	A	P		fuig	5	Prés./absence
5	A	Cb		fug	5	Prés./absence
6	P	Cb	fuig	fug	2	Diff. graph.-phonet.
1	Ca	Cb	de bona guiza	de bona guiza	-3	Identité
2	Ca	P	de bona guiza	mout	-1	Diff. lexicale insign.
3	Ca	A	de bona guiza		5	Prés./absence
4	A	P		mout	5	Prés./absence
5	A	Cb		de bona guiza	5	Prés./absence
6	P	Cb	mout	de bona guiza	-1	Diff. lexicale insign.
1	Ca	Cb	pebre	vitre	4	Diff. sémantique
2	Ca	P	pebre	pebre	-3	Identité
3	Ca	A	pebre		5	Prés./absence

4	A	P		Pebre	5	Prés./absence
5	A	Cb		Vitre	5	Prés./absence
6	P	Cb	pebre	vitre	4	Diff. sémantique
1	Ca	Cb	colerincs	escoloris	4	Diff. sémantique
2	Ca	P	colerincs	qui a mal en lo budel maior	-1	Diff. lexicale insign.
3	Ca	A	colerincs		5	Prés./absence
4	A	P		qui a mal en lo budel maior	5	Prés./absence
5	A	Cb		escoloris	5	Prés./absence
6	P	Cb	qui a mal en lo budel maior	escoloris	4	Diff. sémantique
1	Ca	Cb	mieia semana	mieia semana	-3	Identité
2	Ca	P	mieia semana	tres iorns	-1	Diff. lexicale insign.
3	Ca	A	mieia semana		5	Prés./absence
4	A	P		tres iorns	5	Prés./absence
5	A	Cb		mieia semana	5	Prés./absence
6	P	Cb	tres iorns	mieia semana	-1	Diff. lexicale insign.
1	Ca	Cb	brona	bronia	2	Diff. graph.-phonet.
2	Ca	P	brona	alambroze	3	Diff. lexicale asém.
3	Ca	A	brona		5	Prés./absence
4	A	P		alambroze	5	Prés./absence
5	A	Cb		bronia	5	Prés./absence
6	P	Cb	alambroze	bronia	3	Diff. lexicale asém.
1	Ca	Cb	un test	una ola nova	4	Diff. sémantique
2	Ca	P	un test		5	Prés./absence
3	Ca	A	un test		5	Prés./absence
4	A	P			-3	Identité
5	A	Cb		una ola nova	5	Prés./absence
6	P	Cb		una ola nova	5	Prés./absence
1	Ca	Cb	ayga pur	vinaigre	4	Diff. sémantique
2	Ca	P	ayga pur		5	Prés./absence
3	Ca	A	ayga pur		5	Prés./absence
4	A	P			-3	Identité
5	A	Cb		vinaigre	5	Prés./absence
6	P	Cb		vinaigre	5	Prés./absence
1	Ca	Cb	en vaissel	en vaissel	-3	Identité
2	Ca	P	en vaissel	sus un vaissel	1	Diff. morph.-syn.
3	Ca	A	en vaissel		5	Prés./absence
4	A	P		sus un vaissel	5	Prés./absence
5	A	Cb		en vaissel	5	Prés./absence
6	P	Cb	sus un vaissel	en vaissel	1	Prés./absence

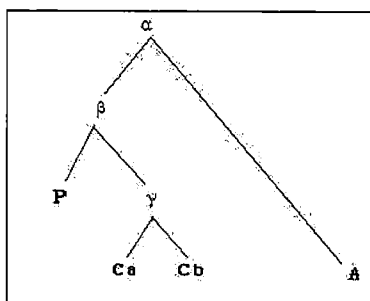


Fig. 3

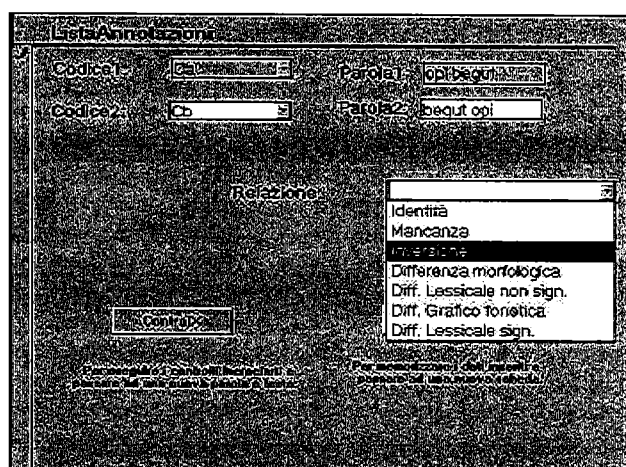


Fig. 4

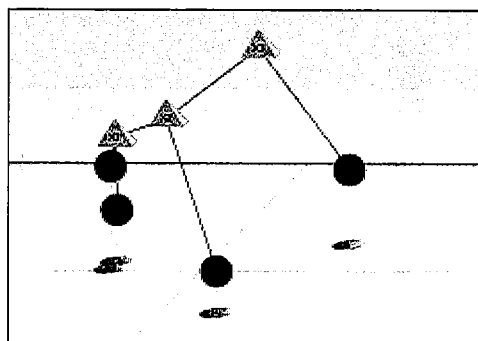


Fig. 5