



HAL
open science

L'application des programmes de reconstruction phylogénétique sur ordinateur à l'étude de la tradition manuscrite d'un texte: l'exemple du chapitre XI de l'Ars Rhetorica du Pseudo-Denys d'Halicarnasse

Hossein Khonsari, Frédérique Woerther

► To cite this version:

Hossein Khonsari, Frédérique Woerther. L'application des programmes de reconstruction phylogénétique sur ordinateur à l'étude de la tradition manuscrite d'un texte: l'exemple du chapitre XI de l'Ars Rhetorica du Pseudo-Denys d'Halicarnasse. *Revue d'histoire des textes*, 2001, 31, pp.227-240. halshs-00199561

HAL Id: halshs-00199561

<https://shs.hal.science/halshs-00199561>

Submitted on 22 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

L'application des programmes de reconstruction phylogénétique sur ordinateur à l'étude de la tradition manuscrite d'un texte : l'exemple du chapitre XI de l'*Ars Rhetorica* du Pseudo-Denys d'Halicarnasse*

RESUME : La reconstruction phylogénétique, appartenant au domaine de la biologie, et le travail d'élaboration du *stemma* d'une tradition manuscrite sont fondés sur des principes très proches puisqu'on peut percevoir une analogie entre la modification des textes au cours de leurs copies successives d'une part et, d'autre part, l'apparition puis la propagation de mutations dans le génome des organismes vivants lors du processus évolutif. L'utilisation des programmes de reconstruction phylogénétique dans l'étude de la tradition manuscrite d'un texte permet au philologue de gagner un temps considérable grâce au traitement informatique des résultats obtenus par le travail de collation. Elle constitue ainsi une base de travail destinée à être discutée et comparée aux données fournies par les études codicologiques et historiques menées à partir des manuscrits.

I. Analogies entre l'étude de la tradition manuscrite d'un texte et la reconstruction de la phylogénie d'un groupe d'êtres vivants.

La phylogénie, terme introduit par Haeckel en 1866, peut se définir comme « le cours historique de la descendance des êtres organisés »¹. Dans la vision strictement géocentrique de la biologie, chaque être vivant est décrit par une séquence d'ADN, le génome, qui code la totalité de ses caractères observables². Au fil des générations, le génome est transmis par le biais de mécanismes de reproduction dont les « erreurs », appelées mutations, sont à l'origine de l'apparition puis de la divergence évolutive des grandes lignées du vivant. La systématique, science des classifications, établit des relations de parenté entre espèces en comparant soit leurs caractères observables qui sont les reflets directs de leur patrimoine génétique, soit des portions de leurs génomes eux-mêmes. Reconstruire une phylogénie, c'est se fonder sur la ressemblance entre individus, postulée comme intelligible en termes d'ascendance commune, pour inférer une phylogénie.

La tradition manuscrite d'un texte comme le chapitre XI de l'*Ars Rhetorica* du Pseudo-Denys d'Halicarnasse regroupe vingt-trois versions différentes d'un même « code » de départ, le texte d'origine, générées par un processus de copie de fidélité imparfaite³. Cette tradition présente par conséquent une double analogie avec une lignée animale

* Nous tenons à remercier Messieurs les Professeurs P. Chiron et P. Darlu pour les relectures qu'ils ont effectuées ainsi que pour leurs précieux conseils, et M. Harold Lopparelli pour son aide technique.

¹ Cf. P. Darlu et P. Tassy, 1993.

² L'ensemble des caractères observables d'un être vivant constitue son phénotype. Le phénotype est l'expression du génotype.

³ Dans le cas de la transmission du chapitre XI de l'*Ars Rhetorica* du Pseudo-Denys d'Halicarnasse, la consultation des textes sur microfilm a permis d'identifier ce processus de copie de fidélité imparfaite comme une succession de fautes de copie mécaniques (sauts du même au même, fautes d'iotacisme) et non comme des corrections dues à des copistes érudits, où un mot est remplacé par un mot différent, dont le sens et l'emploi dans le contexte aurait pu paraître plus adéquats au correcteur.

monophylétique⁴, au niveau du type d'information transmis (codée) et du mode de transmission (faillible). Cette analogie se retrouve lorsque l'on compare les méthodes de recherche de généalogie utilisées par les biologistes et par les philologues. Nous pensons que ces derniers peuvent tirer un grand profit des recherches théoriques actuellement menées par les phylogénéticiens, notamment en établissant une concertation visant à adapter les programmes informatiques de reconstruction phylogénétique à leur propre problématique⁵. Nous allons présenter ici les grandes lignes de la construction d'une classification phylogénétique, en soulignant les points intéressant le philologue.

La similitude est le concept fondateur de toutes les méthodes de reconstruction phylogénétique. Il s'agit d'une similitude de caractères ; les objets de l'étude, manuscrits ou êtres vivants, sont dans un premier temps décrits pour être réduits à un ensemble de caractères suffisant à les définir. Il s'agit du travail de collation, qui consiste, au sein du travail d'élaboration du *stemma* d'une tradition manuscrite, à faire la liste des « fautes » ou écarts présentés par les leçons des différents manuscrits par rapport à une édition de référence. Chaque caractère peut exister sous différents états ; dans notre étude, les caractères sont binaires : présence ou absence d'une modification par rapport au texte de référence sur une lettre, un mot ou une expression donnée.

On distingue au sein du concept de similitude l'homologie et l'homoplasie. Une homologie est une similitude héritée d'un ancêtre commun, tandis qu'une homoplasie est une similitude correspondant à un caractère qui n'existe pas chez l'ancêtre commun, et qui est par conséquent d'apparition plus récente. L'homoplasie peut correspondre par exemple à des fautes d'iotacisme, bien connues des philologues, et dues à la prononciation identique des lettres grecques i, u et h ; elles peuvent apparaître indépendamment les unes des autres dans des textes d'origine totalement différente. On parle alors de convergence.

Les deux grandes écoles de systématiciens, cladisticiens et phénétiens, sont en désaccord à propos du statut attribué aux homoplasies.

1. Les systématiciens phénétiens considèrent que seule la similitude globale, exprimée par des calculs de distance entre espèces, peut servir à regrouper les taxons⁶ les uns avec les autres. Dans le cadre d'une approche phénétique, tout le travail de collation effectué sur une tradition manuscrite aboutirait à l'analyse pour chaque manuscrit de l'ensemble des caractères en même temps, homologies et homoplasies mêlées. Il nous semble que cette approche est trop grossière, et que la perte d'information lors du calcul des coefficients globaux de similitude gomme une partie de la complexité des problèmes spécifiques de l'étude de la tradition manuscrite d'un texte.

2. Les systématiciens cladisticiens considèrent en revanche que les homoplasies ne sont pas discriminantes, et que, par conséquent, la similitude globale, qui considère au même titre homoplasies et homologies, ne permet pas de reconstituer la phylogénie d'un groupe. Ainsi, la cladistique est-elle fondée sur la distinction, au sein des caractères homologues uniquement, des états primitifs (ou plésiomorphes) et des états dérivés (ou apomorphes). Le partage d'un caractère resté à l'état primitif n'est pas un signe de parenté étroite, tandis que le partage de l'état transformé permet d'identifier des relations de parenté et de mettre en évidence les groupes monophylétiques. Un cladogramme, comme celui de la figure 1⁷, est une représentation graphique de la répartition de ces états dérivés. A chaque nœud du

⁴ Un groupe monophylétique est un groupe comprenant un ancêtre et tous ses descendants.

⁵ Cf. C.J. Howe, A.C. Barbrook, M. Spencer, P. Robinson, B. Bordalejo, L.R. Mooney, 2001.

⁶ Le taxon est l'ensemble des organismes contenus dans chaque niveau hiérarchique de la classification biologique.

⁷ Cf. *infra*.

cladogramme correspond la liste des caractères apomorphes partagés par les groupes frères en aval, les synapomorphies. Chaque taxon est défini par un ensemble de synapomorphies.

La définition des apomorphies et des plésiomorphies est relative et dépend du choix d'une référence donnant l'état ancestral de chaque caractère. Il s'agira d'une édition du texte qui sera considérée, lors de l'étude, comme référence, et qui, par définition, sera donc dénuée de fautes. Tous les caractères de cette référence seront par conséquent plésiomorphes.

Si ce texte de référence est présent sur le cladogramme, il aura le statut d'ancêtre de toute la lignée. L'arbre sera alors dit enraciné. Sinon, le texte ancestral peut rester externe à la tradition manuscrite et ne fournir que les informations sur l'état plésiomorphe des caractères. On parle alors d'extra-groupe. L'arbre de la figure 1 est enraciné, et l'édition de H. Usener et de L. Radermacher⁸ a été choisie comme ancêtre de la lignée. Notons que le choix de l'extra-groupe détermine la structure de l'arbre obtenu, et qu'il s'agit par conséquent d'une étape cruciale dans le déroulement de la reconstruction phylogénétique.

Le répertoire des fautes issu du travail de collation auquel s'ajoute l'extra-groupe choisi forment les données de base lues par les programmes informatiques de reconstruction phylogénétique comme Phylip⁹. Phylip regroupe les manuscrits selon les synapomorphies et permet d'obtenir comme résultat tous les arbres possibles qui répondent aux contraintes de la méthode cladistique. Pour trois manuscrits et un extra-groupe, il existe quinze arbres possibles, tous cohérents d'un point de vue cladistique ; ce nombre croît très rapidement avec le nombre de manuscrits. Il faut par conséquent définir un critère de choix pour obtenir un seul arbre, qui sera considéré comme l'arbre qui reflétera le mieux possible la phylogénie du groupe. C'est ici qu'intervient la procédure de parcimonie. Choisir l'arbre le plus parcimonieux, c'est choisir l'arbre qui demande le minimum d'hypothèses de transformation de caractères. La procédure de parcimonie revient par conséquent à considérer les homoplasies comme rares. Ainsi, en simplifiant, lorsque deux manuscrits présenteront un état de caractère commun, l'application de la procédure de parcimonie fera qu'ils seront apparentés dans l'arbre choisi. L'hypothèse d'une convergence (donc d'une homoplasie) sera écartée car elle met en jeu un nombre plus grand d'événements évolutifs.

L'application de la parcimonie pose le problème du nombre d'homoplasies. Si elles sont trop abondantes, l'arbre obtenu ne sera pas phylogénétique. Pour remédier à ce biais, il existe des méthodes statistiques de mesure de l'homoplasie¹⁰. Plus simplement, il est aussi possible de pondérer les différents caractères, ce qui augmente leur influence sur la répartition des groupes lors du calcul de tous les arbres possibles par le programme informatique. Ainsi, les fautes de type « iotacisme » pourront-elles avoir une faible pondération puisqu'elles peuvent comporter beaucoup d'homoplasies.

Inversement, des caractères apparaissant comme homoplasiques dans l'arbre obtenu à l'issue d'une étude bien conduite pourront faire l'objet d'une discussion car ils ont une certaine probabilité d'être effectivement de véritables homoplasies. On peut ainsi imaginer que des recoupements entre différents travaux permettraient de tracer un portrait assez fin des erreurs propres à certains copistes.

Reste un problème majeur qui se pose aux codicologues et qui a lui aussi un analogue biologique : le problème de la contamination. On dit qu'un manuscrit est contaminé lorsque le copiste se réfère, pour le produire, à plusieurs manuscrits-sources : issue de

⁸ H. Usener et L. Radermacher, *Dionysii Halicarnasei Opuscula*, vol. II, fasc. 1, Leipzig, 1904.

⁹ J. Felsenstein, 1995. Pour un exemple de ce répertoire appelé « matrice », cf. Tableau 5.

¹⁰ Cf. A.G. Kluge et J.S. Farris, 1969.

l'agencement de fragments d'origines différentes, la copie ainsi obtenue est un texte composite dont l'histoire ne peut pas être reconstruite s'il est étudié comme un tout. L'apport des méthodes informatiques ici est double. La rapidité du traitement des données permet d'une part de comparer différents arbres obtenus à partir de fautes relevées sur différentes parties du texte, afin de localiser les portions contaminées. De plus, il existe aujourd'hui des logiciels spécialisés dans la recherche des séquences virales parasites au sein des génomes étudiés¹¹. Un travail commun entre codicologues et phylogénéticiens permettrait d'adapter ces programmes aux problématiques spécifiques de l'étude des textes manuscrits.

La construction d'un arbre phylogénétique n'est finalement que la recherche de la meilleure interprétation de la matrice de caractères. Toute l'information est contenue dans les données de départ, et le choix des caractères, puis de l'extra-groupe, sont par conséquent les étapes primordiales de l'étude. Il existe des analogies à tous les niveaux entre les problématiques des études biologiques et celles des études codicologiques, reportées dans le tableau 1 suivant ; ces dernières études peuvent tirer un bénéfice important des recherches théoriques actuellement menées en systématique.

- | |
|---|
| <ol style="list-style-type: none"> 1. Délimitation du groupe d'intérêt : manuscrit et sa tradition 2. Choix d'un échantillon de manuscrits au sein de la tradition 3. Choix de l'extra-groupe et/ou de l'ancêtre : édition de référence 4. Choix des caractères : travail de collation 5. Construction de la matrice de caractères 6. Mesure de l'homoplasie et évaluation de la contamination : étape qui reste à développer en collaboration avec les biologistes 7. Utilisation d'un programme informatique pour calculer tous les arbres possibles 8. Choix d'un arbre par la procédure de parcimonie 9. Lecture et discussion de l'arbre avec l'aide de données extrinsèques (historiques, géographiques, etc.) |
|---|

Tableau 1 : Étapes de la construction d'une classification phylogénétique.

II. Utilisation du programme Phylip dans l'étude de la tradition manuscrite du chapitre XI de l'*Ars Rhetorica* du Pseudo-Denys d'Halicarnasse

Le texte de l'*Ars Rhetorica* du Pseudo-Denys d'Halicarnasse est contenu, d'après les recherches que nous avons réalisées, dans les vingt-trois manuscrits répertoriés dans le tableau ci-dessous (la pagination indiquée est celle du traité entier). Seuls deux manuscrits n'ont pas été pris en compte dans l'utilisation du programme Phylip : le manuscrit de Macédoine (n° 6), le plus récent de tous (il date du XVIII^{ème} siècle)¹² et dont la consultation n'est possible que sur place, dans la bibliothèque de Kozane, et le manuscrit n° 11 qui ne subsiste que dans un état fortement mutilé.

- | |
|--|
| <ol style="list-style-type: none"> 1) Cambridge, Queen's Coll., 20 (Horne 33), ff. 1r-36r. 2) El Escorial, Bibl. Mon., Gr. 114 (S, III, 15), ff. 1r-42r. 3) Firenze, Bibl. Med. Laur., Plut. 59.11, ff. 1r-49r. |
|--|

¹¹ Cf. E.C. Holmes *et al.*, 1999.

¹² Cette datation est fournie par le catalogue établi par A. Sigalas, *Ἰστορικὴ Πνευματικὴ Ἱστορία τῆς Βυζαντινῆς Θεολογίας*. Α'. Ἰστορικὴ καὶ βιβλιογραφικὴ ἐπιθεώρησις τῆς Ἰστορικῆς Πνευματικῆς Ἱστορίας τῆς Βυζαντινῆς Θεολογίας [Panepisthvmion essalonivkh~. Paravrthma D' tovmou jEpethrivdo~ Filosofikh~ Scolh~], Thessaloniki, 1939. La consultation du microfilm

- 4) Firenze, Bibl. Ricc., 15 (K.I.25), ff. 59r-96r.
- 5) Firenze, Bibl. Ricc., 68, ff. 1r-58v.
- 6) Kozane, Demotike Bibl., 77, ff. 1r-46r (pas pris en compte ici).
- 7) Milano, Bibl. Ambrosiana, B 164 sup. (gr. 158), ff. 1r-34v.
- 8) Milano, Bibl. Ambrosiana, M 92 sup. (gr. 537), ff. 1r-59r.
- 9) Napoli, BN, II E 4 (gr. 139), ff. 1r-49v.
- 10) Oxford, Bodl. Libr., Misc. gr. 230, ff. 2r-36v.
- 11) Panormitanus, Bibl. Communalis, 2QqB 123 (pas pris en compte ici).
- 12) Paris, BNF, gr. 1656, ff. 1r-27v.
- 13) Paris, BNF, gr. 1741, ff. 1r-37r.
- 14) Roma, Bibl. Angelica, 54 (D.5.8), ff. 63r-101v.
- 15) Vaticano, Bibl. Apost. Vat., Pal. gr. 66, ff. 1r-44v.
- 16) Vaticano, Bibl. Apost. Vat., Pal. gr. 277, ff. 4r-37v.
- 17) Vaticano, Bibl. Apost. Vat., Ross. 981, ff. 1r-19v.
- 18) Vaticani Graeci: Inventaire des Manuscrits, n° 1405, f. 46r- ?
- 19) Venezia, BN Marciana, 429 (coll. 861), ff. 4r-34r.
- 20) Venezia, BN Marciana, 508 (coll. 844), ff. 196v-202r.
- 21) Venezia, BN Marciana, gr. app. VIII, 10 (coll. 1340), ff. 1r-39r.
- 22) Wien, ÖNB, phil. gr. 60, ff. 1r-39r.
- 23) Wolfenbüttel, HAB, 14 fol. (Gud. gr.), 4201 (Gud. gr. 14), ff. 1r-37r.

Tableau 2 : Liste des manuscrits renfermant le texte du chapitre XI de l’*Ars Rhetorica* du Pseudo-Denys d’Halicarnasse.

Dans un premier temps, on choisit une édition de base qui servira de critère pour la définition des « fautes », c’est-à-dire des écarts présentés par les leçons des différents manuscrits. Que cette édition de base soit l’état du texte de l’un des manuscrits, une édition déjà établie du texte en question ou une édition émendée par les soins de l’éditeur lui-même, peu importe : il faut en tout cas que ce choix soit conscient et que l’on en tienne compte dans toutes les étapes du travail, et principalement lors de la lecture de l’arbre obtenu par l’application du programme de reconstruction phylogénétique. Nous avons pris le parti, comme il a été signalé plus haut, de choisir comme texte de référence l’édition de H. Usener et L. Radermacher.

On établit ensuite, à partir de cette édition choisie considérée comme base et, partant considérée comme dépourvue de « fautes », la liste des fautes discriminantes contenues dans chaque manuscrit. Chacune de ces fautes sera numérotée, décrite et reportée dans un tableau où elle figurera avec sa référence dans l’édition de base (page ou folio et numéro de ligne, par exemple).

A chacune de ces fautes sera attribuée une pondération. Ici aussi, la définition de la gravité ou de l’importance de la faute dépendra du jugement de chaque chercheur et la fourchette des pondérations pourra être plus large que celle que nous avons choisie, puisqu’elle va de 1 à 3 suivant l’importance relative de la faute. Voici les critères que nous avons définis et appliqués dans la classification des fautes apparaissant dans le chapitre XI de l’*Ars Rhetorica* :

<i>Pondération</i>	<i>Type de fautes</i>
1	Fautes mineures, imputables à la prononciation du grec par le copiste qui confond voyelles longues et voyelles brèves par exemple, et qui peuvent survenir indépendamment les unes des autres ; fautes d’iotacisme.
2	Absence ou ajout de particules de liaison, d’articles ou d’adverbe n’altérant pas de manière substantielle le sens du texte ; hésitation entre des préverbes tels que <i>periv</i> et <i>parav</i> , infléchissant le sens du verbe plus qu’ils ne le changent réellement.
3	Omissions (sauts du même au même ou autres types d’omissions) ; ajouts importants de texte induisant une altération du sens d’une manière conséquente, présence d’un mot à la place d’un autre ayant un sens totalement différent

Tableau 3 : Critères de pondération des caractères.

La plupart des manuscrits étudiés proposent des fautes de copistes traditionnelles, dues notamment à la prononciation ou à des sauts du même au même. Au cours de notre travail sur le chapitre XI du traité du pseudo-Denys, aucun manuscrit n'a témoigné d'une « faute » ou d'un écart dû à des corrections imputables à un humaniste.

On obtient alors le tableau des descriptions de fautes suivant, où figurent les fautes, leurs références dans l'édition considérée comme base, et leurs pondérations respectives. Voici les trois premières « fautes » parmi les 97 prises en compte dans notre étude :

<i>N° des fautes</i>	<i>Références (édition Usener-Radermacher)</i>	<i>Pondération</i>	<i>Description des fautes ou « écarts » par rapport à l'édition Usener-Radermacher</i>
1	171, 6	2	Perievcei à la place de parevcei
2	171, 21	2	Absence de la particule gavr après oujdevpote
3	172, 8	1	Misei` à la place de misoi`
....

Tableau 4 : Les trois premières fautes (caractères) issues de la collation de la tradition manuscrite.

Ces fautes serviront ensuite à la construction d'une « matrice », directement exploitable par le programme Phylip, où figurent en ordonnée les différents manuscrits (numérotés ici pour plus de commodité), et en abscisse le numéro que l'on a auparavant attribué à chaque faute, munie de leur pondération respective. Quand un manuscrit contient la faute n°1, on notera 1, quand il ne la contient pas, on notera 0. Ainsi, on pourra lire dans la tableau suivant que le manuscrit 3 ne contient pas les fautes 1, 2 et 4 (dont les coefficients respectifs sont 2, 2 et 3), mais qu'en revanche il comporte la faute 3, affectée du coefficient 1.

<i>Fautes</i>	1	2	3	4	5	6	7	8	9	10	11	12	...
<i>Pondération</i>	2	2	1	3	1	3	1	2	1	1	1	3	...
<i>Extra-groupe</i>	0	0	0	0	0	0	0	0	0	0	0	0	...
<i>ms 1</i>	0	1	0	1	0	0	0	1	0	1	0	1	...
<i>ms 2</i>	1	0	0	0	1	0	1	0	0	0	0	1	...
<i>ms 3</i>	0	0	1	0	0	0	1	0	1	0	0	1	...
...

Tableau 5 : Exemple de matrice de caractères telle qu'elle est lue par le programme Phylip.

Phylip lit ce tableau et construit son arbre en appliquant le critère de parcimonie. Pour ce faire, le programme prend trois des manuscrits, construit tous les arbres possibles et choisit le plus parcimonieux. Il rajoute ensuite un quatrième manuscrit sur cet arbre le plus parcimonieux, reconstruit tous les arbres possibles, et recommence l'évaluation par parcimonie. Et ainsi de suite, jusqu'à l'obtention du *stemma* complet¹³. La pondération des caractères revient à les prendre en compte plusieurs fois : une faute de pondération 3 sera artificiellement considérée par le programme comme apparaissant trois fois dans le tableau.

¹³ Dans ce processus, la forme de l'arbre dépend de l'ordre des manuscrits dans le tableau de données : il existe des options du programme permettant de passer outre ce biais, par exemple en modifiant le choix des trois premiers manuscrits de départ ; de plus, des algorithmes de réarrangements locaux et globaux des branches sont inclus dans le calcul, pour que le plus grand nombre de configurations de branches sur l'arbre soit évalué.

Phylip permet également d'évaluer la robustesse des branches du *stemma* : la méthode statistique utilisée pour ce faire est le bootstrap¹⁴. Dans la pratique, cette méthode permet d'attribuer à chaque branche un coefficient de robustesse, exprimé en pourcentage. Une valeur de plus de 50 % commence à être significative, et les informations données par l'arbre sont dignes de foi, sous réserve de cohérence historique.

Les logiciels de cladistique peuvent être d'un grand secours aux chercheurs, puisqu'ils permettent à ces derniers de considérer un nombre de fautes beaucoup plus important que celui qu'ils ont coutume de prendre en compte dans leurs travaux de collation. En outre, le programme Phylip permet, avec la méthode du bootstrap, de tester la robustesse des branches de l'arbre obtenu. Cette méthode, cependant, ne peut se dérober à la subjectivité du chercheur qui sélectionne les fautes discriminantes et qui, par le jeu des coefficients qu'il affecte à ces dernières, ne cesse pas d'influer sur la construction du *stemma*.

III. Remarques sur la lecture de l'arbre obtenu dans l'étude de la tradition manuscrite du chapitre XI de l'*Ars Rhetorica* du Pseudo-Denys d'Halicarnasse

La matrice répertoriant les fautes contenues dans le chapitre XI de l'*Ars Rhetorica* du Pseudo-Denys d'Halicarnasse et qui ont été précédemment définies a permis d'obtenir un arbre avec l'application du programme informatique choisi. Cet arbre issu de l'analyse cladistique ne fournit pas de réponses définitives aux problèmes posés par l'étude de la tradition manuscrite du texte choisi ici comme exemple, mais il forme néanmoins un support intéressant et une base de discussion qu'il faudra lire en tenant toujours compte des conditions de sa production, c'est-à-dire des données choisies pour la construction de la matrice de départ (nombre, définition et pondération de chacune des fautes ; choix du texte pris comme référence ; étendue du sondage des fautes dans le texte étudié, pour les problèmes de contamination). Ainsi, avec les fautes que nous avons définies plus haut et en prenant l'édition de L. Radermacher du chapitre XI comme texte de référence, le programme permet d'obtenir l'arbre suivant (figure 1) :

¹⁴ Inventé par Bradley Efron en 1979 (cf. bibliographie), cet algorithme construit un grand nombre de tableaux de données factices en tirant des colonnes avec remise (c'est à dire en choisissant au hasard à chaque tour une colonne parmi TOUTES les colonnes du tableau de données de départ). Les tableaux ainsi formés peuvent donc présenter plusieurs fois la même erreur, mais ont la même taille que le tableau de départ. On soumet chacun d'eux à une analyse cladistique avec hypothèse de parcimonie. Un résultat de statistique affirme que les différences de forme entre les arbres issus du bootstrap sont du même ordre de grandeur que celles qui seraient introduites en considérant de nouveaux jeux de données : une branche du *stemma* qui apparaît sur un grand nombre d'arbres construits par bootstrap est donc robuste car elle serait peu déplacée si on introduisait de nouvelles données.

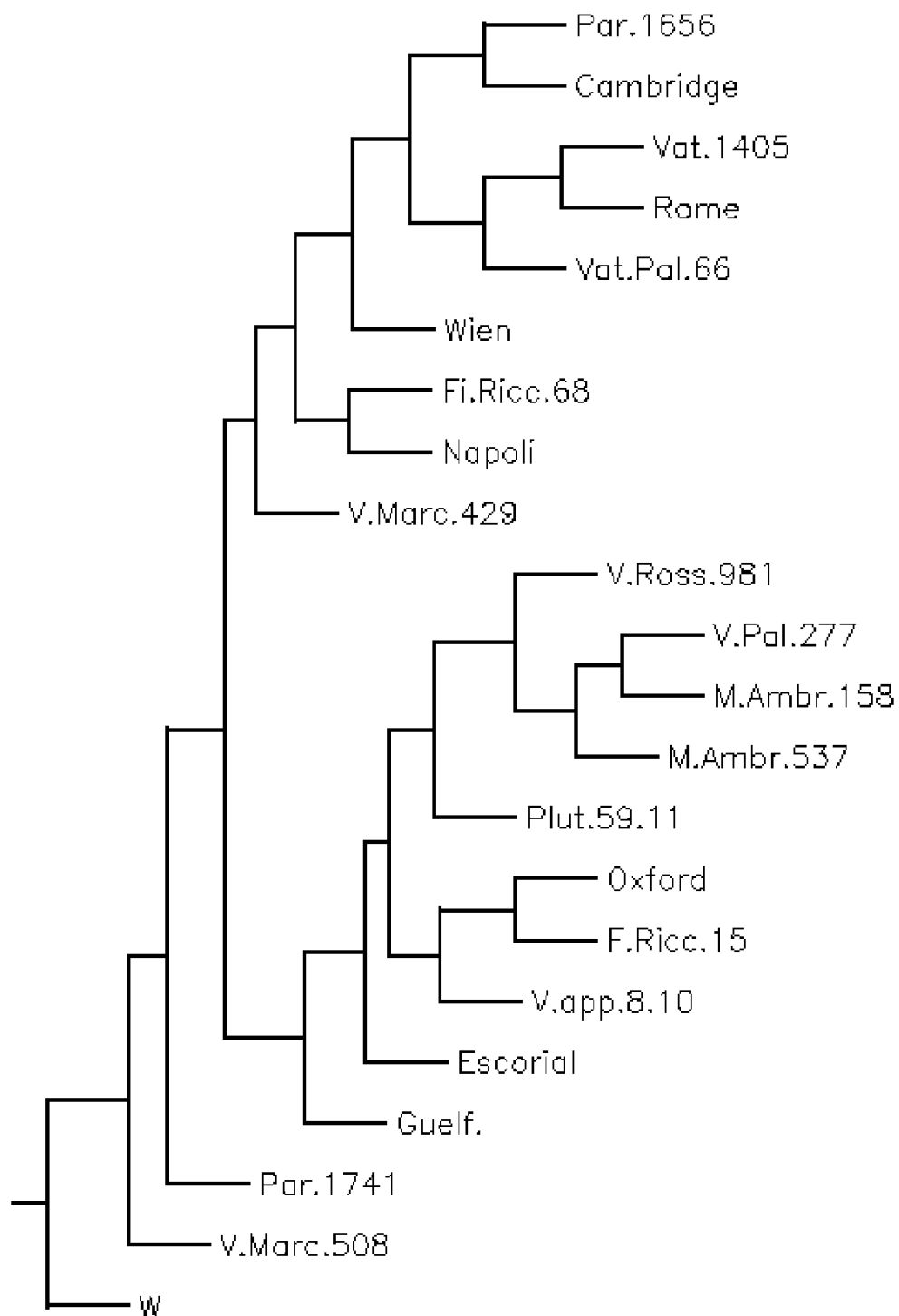


Figure 1 : Proposition de classification phylogénétique de la tradition manuscrite du chapitre XI de l'*Ars Rhetorica* du Pseudo-Denys d'Halicarnasse.

Cet arbre permet d'élaborer, toujours dans le cadre restreint du chapitre XI de l'*Ars Rhetorica* du Pseudo-Denys d'Halicarnasse, deux types de discussion. La première rend possible la confrontation des critères intrinsèques (les données fournies par l'arbre, c'est-à-dire les fautes communes de chaque manuscrit) avec les critères extrinsèques (c'est-à-dire les données que fournissent de leur côté les études codicologiques des manuscrits). La seconde permet de formuler de nouvelles hypothèses et de lancer de nouvelles pistes de recherche dans l'étude de la transmission du texte du chapitre étudié.

1. Certains éléments contenus dans l'arbre permettent de formuler des hypothèses que peut étayer l'étude codicologique et historique de certains manuscrits bien connus : c'est le cas du *Parisinus 1656* et du ms de Cambridge, *Queen's College. 20* (Horne 33). Ces deux manuscrits sont, d'après l'arbre obtenu, à l'intérieur d'un même groupe, du point de vue des fautes communes qui ont été prises en compte dans la construction de la matrice. Émettons donc l'hypothèse, suggérée par l'arbre, d'une liaison entre le *Parisinus 1656* et le manuscrit de Cambridge.

L'étude codicologique révèle d'autre part que la copie des folios 1 à 140v du ms de Cambridge est due à un certain Nicolaos¹⁵, actif aux XV^{ème} - XVI^{ème} siècles à Venise et à Padoue. Le *Répertoire des Copistes Grecs* indique en outre qu'il a collaboré avec les copistes Zacharias Kalliergès, Paulos, et peut-être aussi Konstantinos Mesobotes. La copie du *Parisinus 1656* est, quant à elle, attribuée, toujours d'après le *Répertoire des Copistes Grecs*, à Zacharias Kalliergès¹⁶, actif entre 1473 et 1524.

Ces indications permettent de rapprocher à juste titre, par le biais de Zacharias Kalliergès, les deux manuscrits. Les dates de copie des deux manuscrits n'étant pas établies précisément, on peut imaginer que Zacharias a emprunté à Nicolaos sa copie du texte (le ms de Cambridge), à moins que ce ne soit l'inverse ; mais on peut aussi imaginer l'existence d'un manuscrit intermédiaire disparu qui aurait servi de modèle à l'un et à l'autre copiste. Trois configurations sont donc possibles, les manuscrits ayant été copiés à quelques années d'écart, sans que l'on sache lequel a été copié en premier :

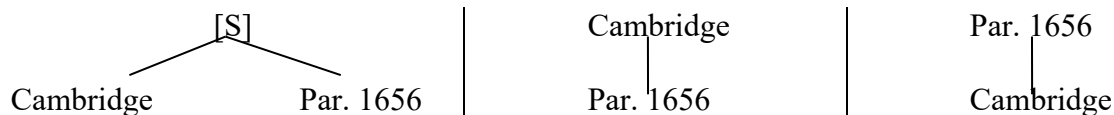


Tableau 6 : Hypothèses sur les relations de parenté entre le manuscrit de Cambridge *Queen's College 20* (Horne 33) et le *Par. 1656*.

La lecture de l'arbre permet donc de formuler des hypothèses de filiation ou, plus généralement, de regroupement entre manuscrits que peuvent étayer des études historiques et codicologiques. Les coefficients de solidité des branches de l'arbre que permet d'obtenir la méthode dite « du bootstrap » est un argument supplémentaire que l'on peut émettre dans le cadre d'un rapprochement entre ces deux manuscrits, puisque la branche qui les relie obtient, par l'application de cette méthode, un coefficient significatif proche de 100 % (91, 49 %).

¹⁵ Il s'agit, d'après le *Repertorium der griechischen Kopisten 800 - 1600, 1. Großbritannien*, édité par H. Hunger, Verlag der Österreichischen Akademie der Wissenschaften, Wien, 1981, du copiste étudié sous le n° 330 (pp. A 170-171 et B 139).

¹⁶ *Repertorium der griechischen Kopisten, 2.*, étudié sous le n° 156 (pp. A 75-76 et B 59).

2. Il est également possible, grâce à la lecture de l'arbre, de lancer de nouvelles pistes de recherche. Ainsi l'arbre obtenu permet-il de poser, ou plutôt reposer le problème du lien qui unit le *Parisinus 1741* et le *Marcianus 508*.

Ces deux manuscrits occupent une place importante dans les éditions de textes rhétoriques, car ils sont deux des témoins les plus anciens et ils présentent en outre une qualité de copie exceptionnelle, si bien que les éditions les considèrent toujours comme de précieux représentants. Comme ils ont en commun un certain nombre de traités¹⁷, la relation qui les unit a été étudiée par les éditeurs de ces textes, notamment H. Usener et L. Radermacher¹⁸ et G. Aujac¹⁹ pour l'édition des traités de Denys d'Halicarnasse, et P. Chiron²⁰ pour celle de Démétrios de Phalère. Parmi ces éditeurs, seul P. Chiron a opté pour une bifurcation, dans le *stemma* de tradition manuscrite du traité de Démétrios, entre les deux manuscrits, estimant que le *Parisinus 1741* et le *Marcianus 508* appartiennent à deux branches distinctes de la tradition, pour ce qui concerne du moins le texte de Démétrios. Sans reprendre dans le détail l'ensemble des arguments avancés par P. Chiron, il faut retenir que les deux manuscrits présentent un certain nombre de traités en commun, mais qu'ils apparaissent dans un ordre différent. En outre, le texte du manuscrit de Venise est souvent plus complet que celui du manuscrit de Paris.

Cette hypothèse rejoint celle que l'on peut formuler à partir de l'arbre obtenu avec l'application du programme de reconstruction phylogénétique. Cet arbre opère en effet une première bifurcation, à partir de l'ancêtre [w] entre le *Parisinus* et le *Marcianus*, ce qui signifie que, d'après les données retenues pour la construction de la matrice, les deux manuscrits appartiendraient à deux traditions manuscrites différentes. Cette première hypothèse lance ainsi une piste dans l'étude des liens qui unissent le *Parisinus* et le *Marcianus*, mais ne constitue en aucun cas une preuve : il s'agit seulement d'un indice supplémentaire permettant d'orienter la recherche du philologue. Une étude plus approfondie du texte sur les corrections opérées dans les copies, sur l'identité des correcteurs et les omissions importantes, mais aussi une étude historique sur le sort des manuscrits appartiennent à la suite de la recherche.

Ces deux exemples permettent de considérer dans quelle mesure la lecture d'un arbre obtenu à partir des programmes de reconstruction phylogénétique peut s'avérer utile : sans formuler des réponses claires et univoques, l'arbre permet d'orienter la recherche du philologue au cours de son étude sur l'histoire d'une tradition manuscrite. *A fortiori*, les exemples que nous avons pris ne permettent en aucun cas de tirer des conclusions sur la transmission de l'ensemble du texte du traité du Pseudo-Denys d'Halicarnasse, car les problèmes de contaminations, où plusieurs manuscrits sont à l'origine d'une seule copie du texte doivent toujours être pris en compte. Une recension complète des fautes sur l'ensemble du texte (du premier au dernier chapitre) est donc une étape indispensable pour commencer une recherche intelligente de l'histoire de la transmission du texte de l'*Ars Rhetorica*.

¹⁷ Les textes que le *Parisinus 1741* (P) et le *Marcianus 508* (M) ont en commun sont : *Ars Rhetorica* attribué à Denys d'Halicarnasse (P : ff. 1r-37r ; M : ff. 196v-202r), Peri; tw'n Qoukudivou ijdiwmavtwn (Lettre à Ammée) de Denys d'Halicarnasse (P : ff. 102v-106r ; M : ff. 204r-207r), Peri; sunqevsew~ ojnoma vtwn de Denys d'Halicarnasse (P : ff. 200r-225r ; M : ff. 174r-196v), Peri; eJrmhneiva~ de Démétrios de Phalère (P : ff. 226r-245v ; M : ff. 157r-173v), Peri; ajluytwn ajntiqevsewn de Maxime le Sophiste (P : ff. 294r-297v ; M : ff. 207r-210r), et Peri; mimhvsew~ de Denys d'Halicarnasse (P : ff. 299r-301r ; M : 202r-204r).

¹⁸ H. Usener et L. Radermacher, *Dionysii Halicarnasei Opuscula*, vol. II, fasc. 1, Leipzig, 1904.

¹⁹ G. Aujac, « Recherches sur la Tradition du peri; sunqevsew~ ojnoma vtwn », *Revue d'Histoire des Textes*, 4, 1974, p. 1-44.

²⁰ Démétrios de Phalère, *Du Style*, texte établi et traduit par P. Chiron, Paris, C.U.F., 1993, p. CXIX sq.

Phylip, et d'autres programmes analogues fonctionnant avec des algorithmes différents, peuvent par conséquent effectuer les calculs requis par l'étude de la tradition manuscrite d'un texte. Comme il n'y a pas aujourd'hui de logiciel parfaitement adapté aux besoins des codicologues et que l'arbre obtenu avec le programme Phylip ne possède aucune valeur chronologique mais classe les différents manuscrits en fonction du nombre de fautes qu'ils possèdent en commun, une concertation avec des informaticiens travaillant dans le domaine permettrait de choisir les méthodes statistiques adéquates parmi toutes celles utilisées en phylogénie. Un tel échange permettrait également d'introduire certaines modifications dans ces logiciels pour prendre en compte les problématiques spécifiques de l'étude de la tradition manuscrite d'un texte, comme par exemple des informations sur la datation relative des manuscrits. Il faut néanmoins souligner qu'un arbre issu d'une analyse cladistique n'est qu'un support de discussion pour des études ultérieures, et ne peut prétendre en aucun cas refléter *directement* la généalogie véritable des manuscrits. Le travail d'identification des copistes, des lieux de copie, l'ordre des textes contenus dans chaque *codex* sont des indices à part entière pour le travail d'enquête que doit mener le chercheur. C'est pour cette raison que le travail de ce dernier ne doit pas s'arrêter à la stricte application du programme, mais que les premiers résultats obtenus par l'application du programme ne seront toujours que des indices supplémentaires, appelant toujours à être confrontés aux résultats que l'examen codicologique proprement dit a permis de recueillir.

Hossein Khonsari et Frédérique Woerther

BIBLIOGRAPHIE SELECTIVE

- Browning R., « Byzantine Scholarship », *Past and Present*, 28, 1964, p.3-20.
- Darlu P. & Tassy P., *Reconstruction Phylogénétique, Concepts et Méthodes*, Collection Biologie Théorique, Masson, 1993.
- Don Cameron H., « The Upside-down Cladogram: Problems in Manuscript Affiliation », in : *Biological Metaphor and Cladistic Classification, An Interdisciplinary Perspective*, Hoenigswald and Wiener Editions, University of Pennsylvania Press, 1987, p.227-242.
- Efron B., « Bootstrap methods : an other look at the jackknife », *Ann. Stat.*, 7, 1979, p.1-26.
- Farris J.S., « Methods for computing Wagner trees », *Syst. Zool.*, 19, 1970, p.83-92.
- Felsenstein J. 1995. *PHYLP (Phylogeny Inference Package) version 3.57c*. Distribué par l'auteur. Université du Washington, Seattle. *Phylip* est disponible librement sur Internet.
- Haeckel E., *Generelle Morphologie der Organismen*, Georg Reimer, Berlin, 1866.
- Hennig W., *Grundzüge einer Theorie der Phylogenetischen Systematik*, Deutscher Zentralverlag, Berlin, 1950.
- Hennig W., *Phylogenetic Systematics*, University of Illinois Press, Urbana, 1966.
- Holmes E.C. *et al.*, « Phylogenetic evidence for recombinaison in dengue virus », *Mol. Biol. Evol.*, 16, 1999, p.405-409.
- Howe C. J., Barbrook A. C., Spencer M., Robinson P., Bordalejo B., Mooney L. R., « Manuscript Evolution », *Trends in Genetics*, 17, 2001, p.147-52.
- Kluge A.G. et Farris J.S., « Quantitative phyletics and the evolution of anurans », *Syst. Zool.*, 18(1), 1969, p.1-32.
- La Pratique des Ordinateurs dans la Critique des Textes*, Colloques internationaux du C.N.R.S. (N°579), Editions du C.N.R.S., 1979.
- Lecointre G. et Le Guyader H., *Classification Phylogénétique des Etres Vivants*, Belin, Paris, 2001.
- Maas P., *Textkritik*, 2. Auflage, Teubner, Leipzig, 1950.
- Reynolds L.D. et Wilson N.G., *Scribes and Scholars*, 2nd ed., Oxford University Press, New-York, 1974.
- Sandys J.E., *A History of Classical Scholarship*, 3 vols., Cambridge University Press, Cambridge, 1958.
- Schubart W., *Das Buch bei den Griechen und Römern*, 2. Auflage, De Gruyter, Berlin-Leipzig, 1921.
- Systématique : Ordonner la diversité du vivant*, Rapport de l'Académie des Sciences (R.S.T. N°11), Tec & Doc, Paris, 2000.
- Tassy P., *L'arbre à remonter le temps*, Christian Bourgeois Editeur, Paris, 1991.

Uthemann K.-H., « Editionstechnik ohne ‚Trenn- und Bindefehler‘? – Zur Rekonstruktion von Stammbäumen mittels des Computerprogramms von A. Dees », *Jahrbuch der Österreichischen Byzantinistik*, 39, 1989, p.49-54.

Uthemann K.-H., « Erster Bericht über griechische Editionen, Handschriften und Hilfsmittel », *Zeitschrift für antikes Christentum*, 1, 1997, p.17-41.