

LE CLASSEMENT DES MANUSCRITS
PAR LA STATISTIQUE ET LA PHYLOGENETIQUE:
LES CAS DE GREGOIRE DE NAZIANZE ET DE BASILE LE MINIME.¹

Quel que soit le type de tradition manuscrite auquel le philologue est confronté, sa tâche première est la même dans tous les cas: trier et classer les témoins. Pour réaliser cette étape, le philologue dispose désormais de moyens techniques qui peuvent le soulager de la part mécanique que comporte la procédure de comparaison (comptage des ressemblances et des différences), en augmenter la rapidité et diminuer le risque d'erreurs qui y est lié. Encore faut-il avoir expérimenté ces techniques, mesuré leur degré de fiabilité et éprouvé leurs limites. Nous voudrions présenter ici deux méthodes complémentaires de classement, l'une purement statistique, l'autre inspirée de la phylogénétique. Ces méthodes, qui ont déjà trouvé des applications dans le monde de la philologie, recourent à des programmes informatiques qu'il est relativement facile de se procurer et d'utiliser. Nous les avons appliquées à deux traditions manuscrites différentes, celle de Grégoire de Nazianze (IV^e s.) et celle de son commentateur, Basile le Minime (X^e s.)², dont nous avons été amenés à préparer les éditions³. Les bons résultats obtenus, grâce à ces méthodes, sur deux traditions manuscrites tellement différentes, nous encouragent à penser que notre expérience peut être profitable à d'autres philologues.

I. PRESENTATION DES CAS

A. La tradition manuscrite des discours de Grégoire de Nazianze est un exemple typique de ce que E. Amand de Mendieta appelait une « tradition manuscrite surabondante »⁴. A ce jour, elle n'a pas encore été entièrement explorée. Seules les collections dites complètes de ces discours ont fait l'objet d'un examen exhaustif, par T.

¹ Cette recherche a été menée conjointement par deux philologues et un informaticien, dans le cadre d'une action de recherche concertée au sein de l'Institut orientaliste de l'Université catholique de Louvain à Louvain-la-Neuve. Les auteurs tiennent à remercier le Prof. B. Coulie, directeur du Centre d'études sur Grégoire de Nazianze, ainsi que Leigh-Ann Bedal, Fellow in Landscape Architecture à Dumbarton Oaks (Washington, DC), qui a pris une grande part dans la mise au point finale des graphiques pour la publication.

² Evêque de Césarée, Basile le Minime est un personnage peu connu, auteur de Commentaires à l'ensemble des discours de Grégoire de Nazianze; pour les rares renseignements biographiques dont on dispose à son sujet, voir R. CANTARELLA, *Basilio Minimo, Scolii inediti con introduzione e note*, I, in *Byzantinische Zeitschrift*, t. 25, 1925, p. 293-295 (= CANTARELLA, *Basilio Minimo*).

³ L'édition du *Disc.* 38 de Grégoire de Nazianze est en cours de préparation par C. Macé; le commentaire de Basile le Minime à ce discours est récemment sorti de presse: Th. S. SCHMIDT, *Basilii Minimi in Gregorii Nazianzeni Orationem XXXVIII commentarii* (*Corpus Christianorum. Series Graeca*, 46. *Corpus Nazianzenum* 13), Turnhout / Leuven, 2001 (= SCHMIDT, *Basilii Minimi*).

⁴ E. AMAND DE MENDIETA, *Un problème d'ecdotique. Comment manier la tradition manuscrite surabondante d'un ouvrage patristique*, in *Texte und Textkritik. Eine Aufsatzsammlung*, herausgegeben von J. DUMMER (*T.U.*, 133), Berlin, 1987, p. 29-42.

Sinko⁵ au début de ce siècle et par V. Somers⁶ ensuite. Jusqu'au travail de V. Somers, les collections complètes des discours étaient classées en deux familles, appelées M et N, d'après le nombre et l'ordre (ou acolouthie) des discours dans la collection⁷. Ce classement, traditionnel depuis les travaux de T. Sinko, a été remis en cause par V. Somers, notamment grâce à la découverte de nombreux manuscrits contenant une collection complète des discours, agencés selon un ordre propre à chacun des témoins, appelés X. Mais aucun autre classement de ces collections n'a été proposé jusqu'ici. Le *Disc. 38*⁸, qui va nous servir d'exemple, est conservé dans 691 manuscrits grecs antérieurs à 1550⁹, ainsi que dans des traductions anciennes en latin, en arménien, en syriaque, en géorgien et en arabe¹⁰. La tradition manuscrite est coupée de son origine par plus de quatre siècles, puisque Grégoire meurt en 390 et que les plus anciens témoins grecs conservés datent du IX^e s. Outre les collections complètes, le *Disc. 38* est également transmis dans des collections partielles (contenant une sélection de discours) et des recueils hagiographico-homilétiques. Seules les collections complètes seront considérées dans le cadre de cet article, parce qu'elles sont représentatives de l'ensemble de la tradition, dont elles constituent une partie relativement homogène. Le classement de ces collections pose des problèmes complexes¹¹. Les 67 manuscrits¹² étudiés ici sont datés du IX^e au XIV^e s. Ils contiennent de 14 à 53 pièces (les 45 discours plus un certain nombre de pièces assimilées¹³), selon leur composition et leur état

⁵ T. SINKO, *De Traditione Orationum Gregorii Nazianzeni I (Meletemata Patristica, II)*, Cracovie, 1917 (= SINKO, *Traditione*).

⁶ V. SOMERS, *Histoire des collections complètes des Discours de Grégoire de Nazianze*, Louvain-la-Neuve, 1997 (*Publications de l'Institut Orientaliste de Louvain*, 48), (= SOMERS, *Collections complètes*).

⁷ SINKO, *Traditione*, p. 149-150; et SOMERS, *Collections complètes*, p. 69.

⁸ Le *Disc. 38* est édité dans la collection des *Sources chrétiennes*, sur base de 19 manuscrits grecs: *Grégoire de Nazianze, Discours 38-41*, introduction, texte critique, traduction et notes par C. MORESCHINI, Paris, 1990 (*Sources chrétiennes*, 358), p. 104-149.

⁹ Voir J. MOSSAY, *Repertorium Nazianzenum. Orationes. Textus Graecus*, I-VI (*Studien zur Geschichte und Kultur des Altertums*, N.F., 2. Reihe: *Forschungen zu Gregor von Nazianz*, Bände 1, 5, 10, 11, 12, 14), Paderborn / Munich / Vienne / Zurich, 1981-1998 (= MOSSAY, *Repertorium*, I-VI). Voir aussi le répertoire des manuscrits grecs et orientaux de Grégoire de Nazianze consultable en ligne, sur le site Nazianzos: <http://nazianzos.fltr.ucl.ac.be/>.

¹⁰ La traduction latine de Rufin d'Aquilée a été éditée par A. ENGELBRECHT, *Tyrannii Rufini Orationum Gregorii Nazianzeni novem interpretatio*, Iohannis Wrobellii copiis usus edidit et prolegomena indicesque adiecit A. ENGELBRECHT, Vienne / Leipzig, 1910 (*Corpus Scriptorum Ecclesiasticorum Latinorum*, 46). Les traductions orientales de Grégoire de Nazianze sont en cours d'édition dans le *Corpus Nazianzenum* (*Corpus Christianorum. Series Graeca*). Pour un dernier état de la question, voir B. COULIE, *Avant-propos*, in *Studia Nazianzenica*, I, B. COULIE (ed.), Turnhout / Leuven, 2000 (*Corpus Christianorum. Series Graeca*, 41. *Corpus Nazianzenum*, 8), p. V-XI.

¹¹ Voir SOMERS, *Collections complètes*, p. 17-41, pour un exposé synthétique sur les différentes tentatives de classement par les éditeurs des *Sources chrétiennes*.

¹² Ces manuscrits ont été partiellement traités par V. Somers, dont nous utilisons les sigles (SOMERS, *Collections complètes*, p. 63-65). Nous avons écarté N 7, N 27 et X 41, trop lacunaires; ainsi que X 15, manuscrit composite dont la partie qui contient le *Disc. 38* est en fait un morceau de collection de XVI (SOMERS, *Collections complètes*, p. 426-429); ce dernier manuscrit apparaît néanmoins dans les phénogrammes. En revanche, nous avons retenu N 11 (restauré au XI^e s.), N 31 (*Cesena, Malatest. gr.* 40), X 35, X 36, X 38, X 39 et X 48 (*Paris, BNF, gr.* 1153), dont V. Somers n'avait pas tenu compte (SOMERS, *Collections complètes*, p. 197-199). Le manuscrit *Moscou, Syn. gr.* 145 n'est toujours pas disponible aux chercheurs.

¹³ SOMERS, *Collections complètes*, p. 8-9.

de conservation. Comme nous l'avons dit, ils se distinguent selon le critère externe de l'acoulouthie en M, N et X¹⁴. Les rapports directs de filiation clairement identifiables sont extrêmement rares : la méthode du « zéro caractéristique » de J. Mogenet¹⁵ a permis d'en repérer un seul, indubitable¹⁶, entre N 6 et N 10. Le nombre de témoins conservés donne une idée du nombre plus considérable encore de manuscrits perdus et donc d'intermédiaires disparus qu'il faudrait postuler pour reconstituer un véritable *stemma codicum*. La difficulté de cette tâche est accrue encore par la quantité d'informations non pertinentes générées par les coïncidences de variantes¹⁷, les contaminations et le fait que la transmission du texte n'est pas aussi mécanique que dans le cas d'un traité scientifique¹⁸. Plutôt que de tenter de retracer des filiations, difficilement prouvables, il a semblé plus judicieux de chercher à déterminer la structure générale de la tradition manuscrite.

B. Face à la complexité du cas de Grégoire de Nazianze, la tradition manuscrite des *Commentaires* de Basile le Minime présente le double avantage d'être beaucoup moins abondante et de permettre une première classification objective des manuscrits sur la base de critères externes. A ce jour ont été identifiés 84 manuscrits transmettant, sous diverses formes, les *Commentaires* de Basile¹⁹. Parmi ceux-là, 53 contiennent le commentaire au *Disc.* 38 qui nous retiendra ici. Dans son inventaire des manuscrits de Basile, J. Sajdak, suivi par R. Cantarella, avait proposé une classification en trois catégories : (a) les manuscrits qui contiennent uniquement le texte des *Commentaires*, (b) ceux qui présentent des scholies marginales de Basile, et (c) ceux qui offrent le texte de Basile dans la version

¹⁴ SOMERS, *Collections complètes*, p. 69-81.

¹⁵ J. MOGENET, *Autolycus de Pitane. Histoire du texte, suivie de l'édition critique des traités de la sphère et des levers et couchers*, Louvain, 1950 (*U.C.L., Recueil de travaux d'histoire et de philologie*, 3^e série, 37), p. 63-64. Seule la technique de J. Mogenet a été appliquée, sans tenir compte de son système de classification des variantes.

¹⁶ J. Bernardi avait déjà vu que N 10 était copié sur N 6 et l'avait signalé dans une lettre du 6 mars 1970 adressée au directeur des *Sources chrétiennes*, citée dans *Grégoire de Nazianze, Discours 27-31*, introduction, texte critique, traduction et notes par P. GALLAY, Paris, 1978, (*Source chrétiennes*, 250), p. 18.

¹⁷ Par exemple, sur les 305 variantes non isolées recensées dans les manuscrits de collection complète du *Disc.* 38, 48 sont attestées par deux manuscrits, formant 46 couples différents (c'est-à-dire deux couples seulement sont attestés plus d'une fois); les 23 triplets constitués sont tous différents et trois couples seulement apparaissent plus d'une fois dans ces triplets.

¹⁸ M. West a par ailleurs souligné le fait que les traités techniques et scientifiques étaient souvent librement remaniés par les copistes, ces traités étant généralement copiés par ceux qui en avaient l'usage (M.L. WEST, *Textual Criticism and Editorial Technique*, Stuttgart, 1973, p. 16-17). Par contre, pour les textes à caractère sacré, le respect de l'intégrité du texte doit être plus grand, en même temps que sa connaissance est plus vivante.

¹⁹ Le premier inventaire des manuscrits de Basile le Minime est dû à J. SAJDAK, *Historia critica scholiastarum et commentatorum Gregorii Nazianzeni (Meletemata Patristica*, I), Cracovie, 1914, p. 37-89 ; il fut repris avec de légères modifications par R. CANTARELLA, *Basilio Minimo*, p. 295-298, et complété récemment par Th. SCHMIDT, *Les Commentaires de Basile le Minime: liste révisée des manuscrits et des éditions*, in *Byzantion*, t. 70, 2000, p. 155-181 (= SCHMIDT, *Liste révisée*) ; les résultats de la révision de Th. Schmidt ont servi à la création d'une base de données électronique des manuscrits de Basile le Minime, accessible sur le site Nazianzos mentionné plus haut. Pour l'identification d'un nouveau manuscrit de Basile depuis la parution de cette liste, voir SCHMIDT, *Basilii Minimi*, p. XII, n. 14.

dite de la « syllogè »²⁰. L'étude approfondie du commentaire au *Disc.* 38 a toutefois montré que cette classification ne tient pas. Sans entrer dans les détails de la discussion, il est possible de distinguer trois versions du texte de Basile: (a) une version longue (à laquelle nous attribuons le sigle L), qui constitue à nos yeux le commentaire de Basile dans sa forme définitive ; (b) une version courte (de sigle B), qui est soit l'abrégé, soit la version primitive, de la première ; (c) une version « syllogè » (S), qui reprend presque intégralement les scholies de la version longue en les juxtaposant à celles d'un commentaire dû à Georges Mocénos. Les 53 manuscrits contenant le commentaire au *Disc.* 38 de Basile se répartissent donc tout naturellement en trois grandes familles en fonction de la version du texte qu'ils transmettent, à savoir 11 manuscrits pour la version longue (L), 12 pour la version courte (B) et 16 pour la version de la « syllogè » (S), auxquels il faut ajouter 14 manuscrits isolés (I) ne se rattachant à aucune des familles précitées²¹. Pour le classement des manuscrits, nous avons fait abstraction des isolés (qui présentent en général des versions radicalement différentes des trois autres et qui demeurent de ce fait « inclassables ») et de ceux qui sont des copies avérées d'autres manuscrits²², ainsi que d'un manuscrit quasi illisible de la version courte²³. Parmi les 32 manuscrits ainsi sélectionnés et classés selon les trois grandes familles²⁴, certaines sous-familles ont en outre pu être distinguées grâce à l'examen de cinq critères externes (titre initial, en-tête de la lettre dédicatoire qui précède le commentaire, titre final, configuration et attributions marginales). Bien qu'incomplets, les résultats de ce classement préalable sont un outil précieux pour tester la validité de ceux qui se dégageront de l'analyse statistique et phylogénétique ci-dessous²⁵.

II. PRESENTATION DE LA METHODE STATISTIQUE

L'utilisation de méthodes statistiques pour le classement des manuscrits n'est pas une idée neuve²⁶. La méthode que nous avons employée ici est largement utilisée en

²⁰ Depuis l'étude de V. PUNTONI (*Scolii alle orazioni di Gregorio Nazianzeno*, in *Studi di filologia greca*, E. PICCOLOMINI (éd.), t. I², Turin, 1883, p. 133-180 ; I³, Turin, 1884, p. 207-246, le terme « syllogè » désigne, à tort, un recueil de scholies dans lequel les explications de Basile seraient mêlées à celles de quatre autres scholiastes. En réalité, la version « syllogè » est la juxtaposition de deux commentaires, celui de Basile et un autre de Georges Mocénos (voir SCHMIDT, *Liste révisée*, p. 157-158).

²¹ Pour la liste complète de ces manuscrits par famille, voir SCHMIDT, *Basilii Minimi*, p. XLVI - XLV.

²² Ainsi, nous avons écarté les trois manuscrits *Paris, BNF, suppl. gr. 1344*, *Leiden, Ruhnck. 58* et *Leiden, Voss. gr. in-f° 45* (qui dépendent tous directement ou indirectement de *Paris, BNF, gr. 573*), ainsi que *Athos, Pantel., gr. 762* (copie de *Athos, Pantel., gr. 7*), *Oxford, Bodl. Libr., Auct. T.1.3* (copie de *Venise, Marc., append. gr. II.43*) et *Paris, BNF, suppl. gr. 842* (copie de *Escorial, ψ.III.3*).

²³ Il s'agit du manuscrit *Oxford, Bodl. Libr., Baroc. gr. 190*.

²⁴ Pour des raisons de clarté, nous avons attribué à chacun des manuscrits un sigle et un numéro en fonction de la famille à laquelle il appartient ; la liste des sigles et des manuscrits se trouve dans SCHMIDT, *Basilii Minimi*, p. LXXIX.

²⁵ Pour le détail de ces résultats, nous renvoyons le lecteur à l'introduction de SCHMIDT, *Basilii Minimi*, p. XXIX - XL.

²⁶ Voir en annexe une bibliographie succincte à ce sujet.

économie et en psychologie et appelée « Multidimensional Scaling » (MDS). Cette technique permet d'analyser les similarités et dissimilarités des données dans un groupe d'objets : dans notre cas, les données sont des variantes et les objets, des manuscrits. Elle modélise les données sous forme de distances entre des points dans un espace géométrique multidimensionnel²⁷, à partir d'une matrice de distances²⁸. Il existe de nombreuses variétés de MDS, qui se distinguent par la technique employée, les buts qu'on leur assigne, les façons de les interpréter et surtout le type de données et leur présentation. Dans notre cas, le but est purement exploratoire : il s'agit d'obtenir une représentation des différences relatives des manuscrits sur base des variantes et de pouvoir ainsi procéder à un premier classement. La matrice de distances est calculée à partir du nombre de désaccords entre les manuscrits deux à deux. Cette matrice permet de calculer les coordonnées des objets, afin de les positionner sur le schéma. L'algorithme utilisé pour cela doit faire en sorte que les distances calculées sur base des coordonnées correspondent le plus possible aux distances réelles. Mais la méthode est par nature approximative. En effet, les coordonnées des points permettent de les placer à l'intérieur d'un espace, déterminé par l'orientation de ses axes et le nombre de ses dimensions. Premièrement, un résultat sans approximation nécessiterait que le nombre de dimensions de l'espace soit égal au nombre d'objets moins un. Dans la pratique, on est obligé de réduire ce nombre, généralement à deux (ou parfois trois) dimensions, pour pouvoir représenter l'espace. Deuxièmement, l'orientation des axes, telle qu'elle découle du calcul du MDS, est arbitraire, toute rotation des axes préservant les distances. La lecture de cette information comporte donc une certaine dose de subjectivité et le philologue doit faire intervenir d'autres paramètres dans son interprétation et dans son estimation de la validité des regroupements obtenus sur base du MDS. Le schéma est en quelque sorte l'expression graphique de l'information contenue dans la matrice des distances.

III. APPLICATION DE LA METHODE STATISTIQUE

a. Préparation des données

La matrice des distances est calculée à partir d'un tableau synthétisant le comportement de chaque manuscrit sur chacun des lieux variants du texte, c'est-à-dire sur chaque « passage du texte, pour lequel tous les manuscrits ne sont pas d'accord; en tout lieu

²⁷ I. BORG, P. GROENEN, *Modern Multidimensional Scaling. Theory and Applications* (Springer Series in Statistics), New York / Berlin / Heidelberg, 1997, p. VII.

²⁸ Nous avons utilisé des distances symétriques, qui donnent le même poids à un accord 0-0 qu'à un accord 1-1. Des distances dissymétriques seraient sans doute plus correctes, parce qu'elles donnent plus de poids à la présence d'un caractère qu'à son absence, mais cela impliquerait que l'on puisse distinguer chaque fois la leçon originale et la leçon secondaire.

variant, les différents manuscrits présentent au moins deux variantes »²⁹. Les variantes ont été collectées à partir de la collation des manuscrits sur un texte de base. A ce stade, intervient une notion importante : la délimitation du lieu variant. Il existe plusieurs manières de procéder, la plus simple et la plus objective étant de donner au lieu variant sa définition minimale : à tout mot du texte correspond un lieu variant. Cette méthode est la plus rigoureuse, mais non la plus économique. Elle permet une pondération quantitative implicite, puisque une variante est comptée autant de fois qu'elle comporte de mots. Cependant, cette méthode multiplie le nombre de variantes par mots, ce qui ne correspond pas nécessairement à la réalité (par exemple, une omission doit être traitée globalement, elle n'est pas l'accumulation d'omissions successives d'un mot). Pour des raisons de commodité, nous avons décidé d'opérer un encodage négatif, c'est-à-dire de n'encoder que les variantes (et non tout le texte), et d'autre part, de conserver comme lieux variants les lieux variants réels, en décomposant toutefois les variantes complexes (mettant en jeu plusieurs types de transformations du texte), mais sans diviser les omissions, ce qui signifie que deux omissions se recouvrant en partie sont considérées comme deux variantes différentes.

Pour des raisons liées aux circonstances de notre travail, la préparation des données s'est réalisée d'une manière différente pour le *Disc.* 38 et pour le commentaire de Basile.

A. Le texte de base utilisé pour la collation du *Disc.* 38 est l'édition du texte réalisée par les Mauristes au XVIII^e s. et reproduite dans la *Patrologie grecque* de J.-P. Migne³⁰. Le travail d'encodage des données récoltées par la collation s'est déroulé en deux temps. Les variantes des manuscrits ont été encodées dans une base de données, dont la structure est telle que chaque variante constitue une fiche, sur laquelle sont indiqués tous les manuscrits qui présentent la variante. Ce type d'encodage est donc doublement négatif, d'une part parce que seules les variantes sont encodées, et d'autre part parce que le comportement d'un manuscrit n'est connu avec certitude que si ce manuscrit porte la variante. Autrement dit, seule la présence du manuscrit sur la fiche est une information univoque ; son absence est interprétée comme l'attestation du texte de base par ce manuscrit, mais il s'agit là d'une simplification (le manuscrit peut présenter une lacune à cet endroit ou une autre variante). Les variantes orthographiques ou phonétiques ont été écartées. Cet encodage a été traduit ensuite sous la forme d'un tableau, présentant les variantes (et non les lieux variants) en

²⁹J. NORET, *Deux Vies de Saint Athanase l'Athonite*, Turnhout, 1982 (*Corpus Christianorum. Series Graeca*, 9), p. LXI. Voir aussi *Table Ronde sur les problèmes de sélection et d'utilisation des variantes 1. Rapport de J. IRIGOIN*, in *La pratique des ordinateurs dans la critique des textes. Actes du Colloque international, Paris du 29 - 31 mars 1978*, J. GLENISSON (éd.), Paris, 1979, p. 266 : « Variante. Le mot s'emploie quand la tradition offre deux leçons ou davantage en un même endroit du texte dénommé lieu variant. »

³⁰Gregorii Nazianzeni Theologi, *Opera*, Paris, 1858 (*Patrologia Graeca*, 36), col. 312 A 1-333 A 7.

lignes et les manuscrits en colonnes. La présence du manuscrit sur la fiche de variante a été symbolisée par un « 1 » dans le tableau, son absence par un « 0 ». Le fait que les variantes n'aient pas été regroupées sous un même lieu variant, en cas de lieu variant multiple³¹, entraîne une distorsion de la réalité, en donnant un poids double aux désaccords des manuscrits. Cela n'aura pratiquement pas d'incidence sur le calcul, étant donné que les lieux variants multiples sont rares.

Quatre MDS ont été réalisés, à partir de données chaque fois modifiées. Pour le premier MDS, toutes les variantes ont été utilisées (473 variantes). Ensuite, seules les variantes non isolées (172 variantes) ont été retenues, puis 71 variantes sélectionnées comme les plus significatives, c'est-à-dire comme porteuses d'une information généalogique. Ces variantes sont celles qui, selon des critères philologiques³², ont été jugées suffisamment importantes pour que leur présence dans deux ou plusieurs manuscrits ne puisse être due au hasard. Enfin, 30 variantes seulement ont été conservées, sur base d'un critère externe, à savoir l'accord des trois versions les plus anciennes (latine, arménienne et syriaque) sur ces variantes. Cette dernière sélection appelle quelques explications. Comme nous l'avons dit, les versions sont plus anciennes que tous les témoins directs conservés du texte de Grégoire. Il a été montré ailleurs que ce témoignage était plus important encore que nous l'avions cru au départ³³. Il a donc semblé intéressant d'utiliser les versions comme critère de sélection des variantes, pour deux raisons : premièrement parce qu'il s'agit d'un critère purement externe et objectif, tout à fait différent des autres critères de sélection employés auparavant ; et deuxièmement parce que cette sélection permet d'introduire les versions en bloc, au même titre qu'un manuscrit grec, et de les comparer aux manuscrits sur une base identique.

B. Vu l'existence de trois versions bien distinctes pour le commentaire au *Disc.* 38, le choix d'un texte de base pour la collation des manuscrits devait porter logiquement sur un représentant de la version longue, puisque cette dernière englobe l'intégralité de la version courte et se trouve elle-même, à quelques ajouts près, à la base de la version « syllogè ». Le meilleur candidat nous a semblé être le manuscrit *Paris, BNF, gr.* 573, du XI^e siècle, qui, sans être le plus ancien, est du moins le plus complet et le plus lisible. Après la collation, complète ou partielle, des 32 manuscrits retenus pour le classement, 150 lieux variants ont été sélectionnés dans la partie du commentaire se rapportant aux six premiers chapitres du *Disc.* 38. En principe, le choix des lieux variants s'est fait sans tentative de pondération philologique ; toutefois, nous n'avons pas retenu (a)

³¹ On appelle « lieu variant multiple » un lieu variant pour lequel il existe plus de deux variantes.

³² Sur la définition et l'application de ces critères, voir C. MACE et C. SANSPEUR, *Nouvelles perspectives pour l'histoire du texte des Discours de Grégoire de Nazianze. Le cas du Discours 6 en grec et en arménien*, in *Le Muséon*, t. 113, 2000, p. 377-416 (= MACE-SANSPEUR, *Discours 6*).

³³ Voir MACE-SANSPEUR, *Discours 6*, p. 414-416.

les variantes orthographiques et phonétiques, (b) les variantes isolées, (c) un très petit nombre de variantes mineures, (d) les variantes portant sur les lemmes ou sur un ajout de la « syllogè » sauf si elles distinguent une sous-famille³⁴. Les 150 lieux variants et leurs variantes respectives ont été traduits dans un tableau sous la forme de chiffres : « 0 » pour le texte de base, « 1 » pour une première variante, « 2 » pour une deuxième, et ainsi de suite pour chaque variante supplémentaire. En outre, deux signes spéciaux ont été utilisés, l'un (#) pour indiquer les cas, très rares, où un manuscrit présente une lacune sur un lieu variant, l'autre (\$) pour signaler l'absence de texte à l'endroit en question, en fonction de la version transmise par le manuscrit. Ce dernier cas concerne en premier lieu les manuscrits de type court, qui omettent un nombre important de scholies, mais aussi certains représentants de la version « syllogè » et occasionnellement ceux de la version longue.

Pour la comparaison des manuscrits et l'établissement des MDS, les # n'ont évidemment jamais été retenus dans les calculs. En revanche, la présence des \$ nous a conduits à distinguer trois méthodes de calcul. Dans la première, les \$ n'ont jamais été pris en compte et les manuscrits n'ont donc été comparés que sur les lieux variants où ils présentent un texte. D'un point de vue philologique, c'est la méthode qui paraît *a priori* la plus satisfaisante, mais elle présente le désavantage de négliger un nombre relativement important de lieux variants, surtout parmi les manuscrits de type court, et risque donc de donner une image déformée des relations entre les manuscrits. Dans la deuxième méthode, nous avons modifié la perspective en tenant compte des accords entre \$, sans toutefois les considérer comme des variantes du lieu variant en question, c'est-à-dire sans les opposer aux autres variantes, puisqu'une comparaison à proprement parler n'est pas possible vu l'absence de texte. L'effet prévisible de ce deuxième calcul sera d'accentuer les similitudes entre sous-familles sans détruire les rapprochements de texte entre familles établis selon la première méthode. A notre sens, il devrait correspondre le mieux à la réalité. Enfin, dans la troisième méthode, nous avons considéré les \$ comme de véritables variantes, c'est-à-dire en tenant compte non seulement des accords entre eux, mais aussi des désaccords avec les autres variantes. Philologiquement, cette méthode ne tient pas, mais nous l'avons consciemment appliquée dans le but d'exagérer les écarts entre les familles et d'obtenir des regroupements de manuscrits plus nets.

³⁴ Ce dernier choix s'explique par la trop grande disparité qui règne parmi les lemmes et par le fait que les lemmes et les ajouts distinguent en général des familles entières, n'apportant aucune information supplémentaire pour le classement des manuscrits à l'intérieur des familles.

b. Lecture et interprétation des graphes

A. Le Disc. 38

1. Toutes les variantes (figure 1)

Il est possible de délimiter deux zones de points, l'une ovale à droite, l'autre plus ou moins circulaire au centre. La zone de droite compte 18 manuscrits, tous d'acolouthie M³⁵ (représentés par un triangle), plus X 24 et X 31. La zone centrale regroupe la majorité des manuscrits (42 manuscrits), tous N (représentés par un tiret) et X (représentés par un cercle) sauf M 7³⁶, M 9 et M 25. Il reste cinq témoins qui n'entrent dans aucune des deux zones : M 16, X 9, X 22, X 25, X 39. Les manuscrits sont plus resserrés dans la zone centrale, certains sont pratiquement confondus. La zone de droite pourrait peut-être être divisée en deux sous-zones, l'une au nord et l'autre au sud de M 10, mais cette subdivision est plutôt subjective à ce stade.

2. Variantes non isolées

La configuration restant pratiquement inchangée par rapport au premier MDS, il a semblé superflu de publier ce graphique. L'élimination des variantes non isolées n'apporte aucune information supplémentaire.

3. Variantes sélectionnées philologiquement (figure 2)

Normalement, la meilleure configuration doit être obtenue sur la base des variantes significatives, puisque ce sont elles qui contiennent le plus d'information sur les rapports entre les manuscrits. Effectivement, ce schéma peut être lu d'une manière qui nous fournit plus de renseignements que dans les autres MDS. Autour de la zone circulaire N-X en bas à gauche, se dessine une zone périphérique, contenant M 9, M 25 et N 30 à l'est ; X 05, X 22, N 11 et N 7 au nord. M 9 et M 25 sont rapprochés ainsi pour la première fois ; Ces deux manuscrits sont apparentés : ils présentent des variantes propres (ou qu'ils partagent avec quelques autres manuscrits) et même une faute commune. Dans le noyau de la zone N-X elle-même, N 6 et N 10, dont nous avons dit plus haut qu'ils étaient unis par un rapport de filiation, sont confondus, alors qu'ils ne l'étaient pas sur les deux autres schémas. Par contre, X 7 et X 27 sont moins proches ici qu'ils ne l'étaient sur les précédents MDS,

³⁵ Par commodité, nous désignerons les groupes de manuscrits par le sigle qui correspond à l'acolouthie majoritaire des manuscrits de ce groupe.

³⁶ Il faut noter que cette collection ne suit un ordre M que pour sa première partie, l'ordre de la deuxième partie est indéterminée : SOMERS, *Collections complètes*, p. 434-438.

mais sont effectivement apparentés. En dehors de la zone N-X et de sa périphérie proche, on retrouve à peu près les mêmes manuscrits que ceux qui étaient restés hors classement dans le premier MDS : X 9, X 25, X 39. Le manuscrit M 16 se trouve cette fois dans la zone d'influence des M, tandis que X 22 rejoint celle des N-X. Trois manuscrits viennent s'ajouter aux inclassables, mais ils restent proches de la périphérie de N-X : X 12, X 26 et X 35. Certains de ces huit manuscrits (M 16, X 09, X 12, X 22, X 25, X 26, X 35 et X 39) sont probablement des témoins contaminés. C'est presque certain en ce qui concerne X 25, car ce manuscrit présente un texte hybride également pour d'autres discours³⁷, et les titres des discours qu'il contient sont mixtes³⁸. D'autres peuvent provenir d'ancêtres qui se placent à un autre niveau par rapport aux deux groupes principaux. C'est peut-être le cas de X 22, qui atteste dans plusieurs discours des variantes qui semblent plus anciennes que les variantes M et N-X³⁹. Ce sont en tout cas des manuscrits fortement individualisés, qui contiennent généralement de nombreuses variantes isolées⁴⁰. Il est intéressant de noter que ces manuscrits inclassables sont tous anciens (Xe-XIe s.), sauf X 12 (XIIe s.) et X 35 (XIVe s.). Le cas de X 22 est curieux, c'est un manuscrit du Xe s., copié dans une minuscule apparentée à la bouletée, peut-être d'origine provinciale⁴¹, qui contient deux pièces en double⁴². Une hypothèse plausible pour expliquer cette anomalie serait que ce manuscrit (et sans doute d'autres) aurait été copié sur base de plusieurs modèles dépecés en cahiers, les cahiers étant répartis entre différents copistes, afin de pouvoir réaliser plusieurs copies simultanément⁴³. Le découpage en cahiers aurait été mal réalisé pour le modèle de X 22, ce qui expliquerait la double copie de quelques pièces. L'hypothèse de la copie simultanée pourrait être également un élément d'explication de la variété des acolouthies et de l'explosion démographique⁴⁴ que connaissent les collections complètes d'acolouthie variable au XIe s., alors que l'apogée des collections de type M et N se situe au Xe s. Cette production abondante devait répondre à une demande plus importante à cette époque. Dans la zone M, les manuscrits sont nettement moins resserrés que dans la zone N-X. M 1,

³⁷ C'est le cas pour les *Disc.* 6 et 27, par exemple. Ce dernier discours est étudié par C. Macé, dans le cadre de sa dissertation doctorale. Le cas du *Disc.* 6 a été traité sommairement dans un article par MACE-SANSPEUR, *Discours 6*.

³⁸ C'est particulièrement spectaculaire pour le *Disc.* 14. Voir SOMERS, *Collections complètes*, p. 146.

³⁹ Voir MACE-SANSPEUR, *Discours 6*, p. 407.

⁴⁰ X 25 a 19 variantes isolées pour le *Disc.* 38 et 10 pour le *Disc.* 27 ; X 22 en compte respectivement 14 et 42.

⁴¹ M.L. AGATI, *La minuscola « bouletée »*, Vatican, 1992 (*Littera Antiqua*, 9), p. 282-283.

⁴² Il s'agit des *Epist.* 101 et 102: SOMERS, *Collections complètes*, p. 466-469.

⁴³ Cette pratique est connue pour quelques grands *scriptoria* occidentaux latins. Cf. N. NETZER, *Cultural Interplay in the Eighth Century. The Trier Gospels and the Making of a Scriptorium at Echternach*, Cambridge, 1994 (*Cambridge Studies in Palaeography and Codicology*, 3), p. 28-33. Voir aussi P. CANART, *Quelques exemples de division du travail chez les copistes byzantins*, in *Recherches de codicologie comparée. La composition du codex au Moyen Âge en Orient et en Occident*, P. Hoffmann (éd.), Paris, 1998 (Bibliologie), p. 49-67.

⁴⁴ Les manuscrits X représentent 48% des collections complètes. On compte 20 collections M (sur 27 au total) et 20 collections N (sur 34 au total) aux IXe et Xe s., puis ce chiffre tombe respectivement à 5 et à 10 pour le XIe s., alors qu'il est de 26 pour la même période en ce qui concerne les collections X.

M 11, M 21 forment un petit groupe⁴⁵, auquel se joint M 12. De même, M 20 et X 31 d'une part et M 2 et M 5 d'autre part constituent des petites entités plus soudées, qui, tout comme M 10, M 22 et M 23, appartiennent certainement au même groupe que M 1, M 11, M 12 et M 21. Viennent ensuite M 3, M 6, M 14, M 15, M 16, M 17, M 24 et X 24. Ces manuscrits semblent isolés les uns des autres, sauf M 15 et M 3.

4. Variantes sélectionnées par un critère externe (figure 3)

Les variantes sélectionnées de cette manière ne sont pas suffisamment discriminantes : de nombreux manuscrits échappent aux groupements, un certain nombre sont pratiquement confondus. Quelques manuscrits ont été éliminés du graphe, parce qu'ils avaient un comportement identique sur toutes ces variantes. Ainsi, N 12 représente aussi N 22 et N 28, N 25 remplace X 11 et X 18, N 6 équivaut à N 10 et N 13. Néanmoins, là encore, les groupes principaux demeurent : d'une part les manuscrits N-X (plus M 7, M 9 et M 25) et d'autre part une zone de manuscrits M (plus X 24, X 25, X 31) assez large et éclatée. Le petit nombre de variantes conservées fait qu'une différence portant sur une variante a des répercussions importantes sur le positionnement des points. Les versions, symbolisées par un manuscrit fictif X 00 (représenté par une croix), viennent se placer plutôt du côté N-X dans ce schéma, tout près de X 22. La tendance des versions anciennes à soutenir en bloc ce groupe de manuscrits est confirmée par les recherches récentes sur d'autres discours⁴⁶. Il s'agit bien d'une tendance et non d'un accord complet, puisque les versions attestent parfois la leçon de l'autre groupe. Cette dernière expérience montre bien la stabilité des groupes principaux déjà repérés dans le premier MDS.

B. Le commentaire au *Disc.* 38

1. Méthode 1: sans tenir compte des \$ (figure 4)

Les trois familles (texte court, texte long, « syllogè ») sont bien apparentes : les manuscrits de type court forment pour ainsi dire une ligne oblique au centre du graphe, qui délimite d'un côté la famille des « syllogè » et de l'autre les manuscrits de type long. La seule exception est le manuscrit S 14, qui figure parmi les manuscrits de la version longue alors qu'il est de type « syllogè ». On ne s'en étonnera pas dans la mesure où l'examen des

⁴⁵ Sur ce groupe de manuscrits, voir S. LUCA, *Scritture e libri della scuola niliana*, in *Scritture, libri e testi nelle aree provinciali di Bisanzio. Atti del Seminario di Erice, 18-25 settembre 1988*, a cura di G. CAVALLO, G. DE GREGORIO, M. MANIACI, 2 vol., Spolète, 1991 (*Biblioteca del « Centro per il collegamento degli studi medievali e umanistici nell'Università di Perugia »*, 5), p. 373.

⁴⁶ Voir MACE-SANSPEUR, *Discours* 6, p. 409.

critères externes, dont les résultats ont été publiés ailleurs⁴⁷, avait déjà révélé une grande affinité de ce manuscrit avec ceux de la version longue.

Le graphe corrobore également les constatations faites à propos de chacune des familles sur la base des critères externes. Ainsi, les écarts relativement importants entre les manuscrits de type long confirment l'idée qu'il s'agit d'un groupe peu homogène, bien que nettement distinct des deux autres. Parmi les manuscrits de type « syllogè », on retrouve également les sous-familles déterminées à l'aide des critères externes, à savoir un premier groupe constitué de S 6, S 7 et S 8, nettement à l'écart, et un autre regroupant les manuscrits S 2, S 3, S 4 et S 5, auquel s'apparente visiblement le manuscrit S 1, comme on le sait du reste par les critères externes. Le graphe confirme aussi la proximité de S 9, S 10 et S 11, et la position isolée des manuscrits S 12 et S 13 qui semblent montrer une affinité avec les manuscrits de type court. Parmi ces derniers, le graphe permet d'isoler les manuscrits B 1 et B 2, mais n'y rattache pas B 3 comme on s'y attendrait en fonction des critères externes. Pour le reste, on retrouve l'affinité de B 9 et B 10 ainsi que le groupe plus large constitué de B 4, B 5, B 6, B 7, B 8 et B 11. D'une manière générale, on peut dire que l'analyse statistique des données textuelles permet de distinguer plusieurs sous-groupes, qui confirment le premier classement obtenu sur la base des critères externes.

2. Méthode 2: accords des \$ (figure 5)

Comme on s'y attend, la deuxième méthode a pour effet de resserrer les manuscrits à l'intérieur des familles, puisque le nombre de leurs accords a augmenté en fonction des \$. Les trois grandes familles se distinguent plus nettement que dans la figure 4. À nouveau, le manuscrit S 14 se singularise par sa présence au milieu de la famille des longs, qui demeure un groupe relativement éclaté. La position isolée de S 12 et S 13 est confirmée et même accentuée : ils se situent nettement à l'écart des manuscrits de type « syllogè ». Parmi ces derniers, les deux groupes constitués pour l'un de S 6, S 7 et S 8, et pour l'autre de S 1, S 2, S 3, S 4 et S 5 sont eux aussi confirmés (avec S 1 plus franchement rattaché à son groupe que dans la figure 4). S 9, S 10 et S 11 restent visiblement proches les uns des autres, mais cette parenté n'est pas accentuée par rapport à la méthode 1. Parmi les manuscrits de type court, B 3 est maintenant clairement rattaché à B 1 et B 2, corrigeant l'anomalie constatée dans la figure 4 ; ces trois manuscrits forment un groupe clairement identifiable. Pour le reste, on constate que le manuscrit B 11 s'est nettement distancié du groupe constitué de B 4, B 5, B 7 et B 8 auquel on pouvait le rattacher dans le premier graphe ; or, textuellement, il s'agit bien d'un manuscrit isolé du reste des manuscrits de type court. En revanche, l'éloignement relatif de B 6 du groupe précité est une petite surprise, étant donné que ce manuscrit partage avec lui des variantes importantes. Enfin, l'affinité

⁴⁷ Voir SCHMIDT, *Basilii Minimi*, p. XXIX - XL.

entre B 9 et B 10 reste visible sur le graphe, sans toutefois être accentuée par rapport à la méthode 1.

3. *Méthode 3: accords et désaccords des § (figure 6)*

La méthode 3 accentue comme prévu les différences entre les trois grandes familles. Pour l'établissement des sous-groupes, elle confirme dans les grandes lignes les résultats obtenus lors des calculs précédents, sur lesquels nous ne revenons pas. Il convient toutefois de noter que le manuscrit B 7 s'est légèrement détaché du groupe B 4, B 5, B 8, à l'instar du manuscrit B 6 dont il semble plus proche. Il se peut que nous ayons à faire à un nouveau sous-groupe ; ces cinq manuscrits restent néanmoins relativement proches. L'affinité de B 9 et de 10 est à nouveau confirmée. Parmi les manuscrits de type « syllogè », signalons que la parenté que les critères externes avaient établie entre les manuscrits S 9, S 10 et S 11 semble voler en éclat dans le graphe 3 ; cet éclatement n'est cependant qu'illusoire, nous semble-t-il, car la méthode utilisée pour le calcul 3 a précisément pour effet d'accentuer les différences au détriment des similitudes. Cette dernière remarque s'applique aussi à ce que nous avons dit plus haut de B 6 et B 7.

c. *Conclusions*

A. En ce qui concerne le *Disc.* 38, les différents MDS ont fourni des résultats cohérents. Les meilleurs résultats ont été obtenus sur base de variantes sélectionnées selon des critères philologiques⁴⁸. Le traitement strictement statistique des données textuelles tirées des collations du *Disc.* 38 a permis de dégager deux groupes principaux dans la tradition manuscrite, l'un comprenant la majorité des manuscrits, tous N ou X, à l'exception de M 7, M 9 et M 25 ; l'autre, plus réduit, comptant 16 manuscrits M, plus X 24 et X 31. Ce deuxième groupe est moins homogène que le premier, on peut y distinguer plusieurs sous-groupes et des individus isolés, entretenant entre eux des rapports de proximité plus ou moins lâches. Un certain nombre de manuscrits X n'entrent dans aucun groupement et se présentent comme des entités isolées, probablement contaminées.

B. Pour les manuscrits de Basile, les trois MDS ont produit des résultats concordants. Si les méthodes 2 et 3 ont offert des graphes distinguant de façon plus nette les trois grandes familles, en revanche, pour l'établissement de groupes à l'intérieur des familles, il s'est avéré que les résultats sont plus ou moins identiques quelle que soit la méthode utilisée. La meilleure appréciation des sous-familles nous semble toutefois

⁴⁸ Cette sélection des variantes n'est pas, comme A. Dees et d'autres philologues le pensent, un moyen arbitraire d'orienter les résultats a priori, mais une étape nécessaire de l'analyse visant à éliminer le plus possible les informations non pertinentes. Cf. MACE-SANSPEUR, *Discours* 6, p. 379-380.

provenir du graphe de la méthode 2, dans lequel on assiste à un resserrement des groupes constatés dans le graphe 1. Quoi qu'il en soit, la conclusion la plus importante est que les résultats obtenus à l'aide des MDS confirment sur presque tous les points les constatations faites à partir des critères externes. Cette concordance est le meilleur garant de la fiabilité de cette méthode.

IV. UNE METHODE PHYLOGENETIQUE: L'APPROCHE PHENETIQUE⁴⁹

Une autre méthode, au départ mise au point pour la phylogénétique⁵⁰, permet également de visualiser les proximités des témoins, mais sous forme d'une représentation arborée⁵¹. Différentes approches phylogénétiques sont possibles, principalement la cladistique et la phénétique. La cladistique est simplement une méthode de classification qui groupe les objets (en systématique biologique, ces objets sont eux-mêmes des groupes, ou « taxons », c'est-à-dire des groupes d'organismes, reconnus en tant qu'unités formelles à quelque niveau de la classification que ce soit) hiérarchiquement en ensembles et en sous-ensembles⁵². La phénétique diffère de la cladistique par sa manière de traiter la similarité⁵³. En effet, la phénétique, ou taxinomie numérique, considère la « similitude globale » entre les individus, c'est-à-dire qu'elle traite tous les caractères de la même façon, sans distinguer la cause de la similitude des caractères de deux individus. La similitude recouvre en effet deux concepts très différents : l'homologie, c'est-à-dire le « partage d'un même caractère en raison d'une ascendance commune », et la convergence, qui est une « ressemblance apparue indépendamment »⁵⁴. Dans notre cas, les variantes sont toutes considérées sur le même pied, quelle que soit la cause de la similitude des variantes entre les manuscrits. En règle général, la sélection des variantes significatives a conservé uniquement des variantes qui ne peuvent se transmettre que par génération et non par hasard. Toutefois, le sens de la variation (ou de la dérivation, puisqu'on passe d'un caractère primitif à un caractère dérivé) n'est pas indiqué, et il est rarement possible de l'indiquer. En cladistique, seul le caractère dérivé est porteur d'information généalogique : le partage d'un caractère primitif ne nous

⁴⁹ Nous tenons à remercier P. Baret, Professeur à la Faculté de Sciences agronomiques de l'Université catholique de Louvain à Louvain-la-Neuve, pour le temps qu'il a passé à nous expliquer les arcanes de la phylogénétique.

⁵⁰ La phylogénétique est « l'histoire de la descendance des êtres organisés ». Cf. P. TASSY, *L'arbre à remonter le temps. Les rencontres de la systématique et de l'évolution (Épistémè Essais)*, Paris, 1991, p. 317 (= TASSY, *Arbre*).

⁵¹ Un arbre est « un graphe non orienté, connexe et sans cycle ». Cf. N.X. LUONG et M. NOVI, *Représentations arborées de données textuelles*, in *Méthodes quantitatives et informatiques dans l'étude des textes. Hommage à C. Muller, Université de Nice, 5-8 juin 1985*, Genève / Paris, 1986, p. 577.

⁵² I.J. KITCHING, P.L. FOREY, C.J. HUMPHRIES et D.M. WILLIAMS, *Cladistics. The Theory and Practice of Parsimony Analysis*, 2^e ed., Oxford / New York / Tokyo, 1998 (*Oxford Science Publications. The Systematics Association*, 11), 1.

⁵³ M. RUVOLO, *Reconstructing Genetic and Linguistic Trees : Phenetic and Cladistic Approaches*, in *Biological Metaphor and Cladistic Classification. An Interdisciplinary Perspective*, H.M. HOENIGSWALD, L.F. WIENER(eds), Philadelphia, 1987, p. 193-216.

⁵⁴ TASSY, *Arbre*, p. 132-134.

apprend rien sur la parenté de deux individus, puisque ce caractère appartenait déjà à l'ancêtre. Dans notre tableau d'analyse du comportement des manuscrits, contrairement aux tableaux utilisés en cladistique, le « 1 » n'est pas plus porteur d'information que le « 0 », et ce d'autant plus que, dans le cas du *Disc.* 38, le texte de base utilisé pour la comparaison est un texte hybride, constitué par les Mauristes en prenant des leçons à différentes sources, manuscrits et éditions⁵⁵.

L'application d'une telle méthode à la philologie nous semble légitime. L'analogie entre la descendance des êtres vivants et celle des manuscrits a été remarquée depuis longtemps : c'est elle qui est à la base de la méthode lachmannienne et du *stemma codicum*⁵⁶. L'utilisation de ces méthodes n'est donc, comme l'a dit B. Salemans⁵⁷, qu'une résurrection de la méthode de Lachmann, revue et nuancée sur de nombreux points cependant. On peut résumer comme ceci les analogies entre la cladistique et l'ecdétique :

	Cladistique	Ecdétique
but	hiérarchie des « groupes frères »	généalogie des manuscrits
objet	taxons	manuscrits
critère de parenté	synapomorphie (partage de caractères dérivés)	fautes communes
méthode	« three-taxon statement » + parcimonie	méthode quentinienne (2 contre 1)
représentation	Cladogramme (dichotomique)	<i>Stemma</i> (le plus souvent bifide)
racine	ancêtre: « morphotype ancestral hypothétique », « la plus petite unité évolutive identifiable »	archétype: le plus ancien commun dénominateur reconstituable de toute la tradition
interprétation	scénario évolutif	histoire du texte

⁵⁵ Cf. J. MOSSAY, *Quelques manuscrits de Paris relatifs à l'édition de S. Grégoire de Nazianze par les Mauristes*, in *L'Antiquité Classique*, t. 50, 1981, p. 545-561. Les Mauristes ont par ailleurs certainement utilisé l'*Editio Princeps* de Marcus Musurus, ou une édition qui en dérive, et cette *Editio Princeps* puise elle-même à plusieurs manuscrits de groupes différents. Cf. M. BERTOLINI, *L'edizione aldina del 1516 e il Testo delle Orazioni di Gregorio Nazianzeno*, in *Studi classici e orientali*, t. 38, 1988, p. 383-390.

⁵⁶ La linguistique historique, la philologie et la systématique biologique naissent dans une atmosphère culturelle et scientifique commune. Cf. S. TIMPANARO, *La Genesi del metodo del Lachmann*, Florence, 1963 (*Bibliotechina del Saggiatore*, 18), p. 73-76. L. Cavalli-Sforza considère la faute de copie comme une « mutation culturelle (...) accidentelle et faible comme l'est une mutation biologique ». Cf. L. CAVALLI-SFORZA, *Gènes, peuples et langues*, Paris, 1996 (*Travaux du Collège de France*), p. 260.

⁵⁷ B. SALEMANS, *Cladistic and the Resurrection of the Method of Lachmann. On Building the Stemma of « Yvain »*, in *Studies in Stemmatology*, P. VAN REENEN, M. VAN MULKEN (éds), Amsterdam / Philadelphia, 1996, p. 3-70.

Ainsi, on a pu dire que le *stemma* était en quelque sorte un « cladogramme inversé ». En effet, le zoologiste est intéressé par les terminaisons de l'arbre, tandis que le philologue, au contraire, veut les éliminer le plus possible pour ne garder que les ancêtres communs des différentes branches, avec comme but théorique de reconstruire l'archétype, cet ancêtre reconstruit étant une fiction pour le zoologiste. En outre, le philologue voit dans le processus d'évolution non un progrès, mais une corruption⁵⁸.

Les programmes que nous avons utilisés sont les programmes Fitch et Kitsch, mis au point par J. Felsenstein et fournis dans la suite de programmes PHYLIP (PHYLogeny Inference Package), disponibles gratuitement sur le Web⁵⁹. Ces programmes peuvent être utilisés en vue de traitements cladistique et phénétique. En raison de la nature de nos données, nous n'avons exploité que ce dernier aspect. Les données de départ sont pratiquement les mêmes que pour les MDS, à savoir les matrices des distances des manuscrits. Les arbres obtenus, ou phénogrammes, sont des arbres non enracinés et non orientés, qui expriment des « distances taxinomiques » entre les individus, c'est-à-dire des relations de proximité et non de filiation. Il ne s'agit donc absolument pas d'un *stemma codicum*. Ces résultats sont indicatifs : c'est leur interprétation par le philologue qui leur donnera tout leur sens. La procédure revient à peu près à ceci : un manuscrit est pris arbitrairement comme point de départ et tous les autres manuscrits sont positionnés par rapport à ce point de départ, qui n'est pas une racine. L'algorithme considère d'abord les objets trois par trois, et les dispose par dichotomies successives. Il s'arrête lorsque tous les objets ont été placés. La méthode d'optimisation est le principe de parcimonie, qui minimise les distances. Pour améliorer le schéma selon ce principe, l'algorithme peut effectuer dans un deuxième temps des réorganisations globales des branches. Dichotomie et parcimonie sont des principes logiques, non des expressions de la réalité. Les phénogrammes sont toujours dichotomiques, sans que cela implique une éventuelle bifidité de la tradition, parce que les branches du phénogramme ne représentent pas des descendance, et parce que la dichotomie tient à la nature de la technique, non à celle des données. Selon le manuscrit de départ, la morphologie de l'arbre peut varier, c'est la stabilité des groupes formés qui importe. Les deux programmes, Fitch et Kitsch, sont ceux qui conviennent le mieux pour traiter des matrices de distances, car ils obéissent tous deux à un « additive tree model », selon lequel les distances doivent évaluer les sommes des longueurs des branches entre les objets. Ce sont ces programmes qui sont utilisés pour

⁵⁸ H. D. CAMERON, *The Upside-down Cladogram. Problems in Manuscript Affiliation*, in *Biological Metaphor and Cladistic Classification. An Interdisciplinary Perspective*, H. M. HOENIGSWALD, L. F. WIENER (eds), Philadelphia, 1987, p. 227-242.

⁵⁹ <http://evolution.genetics.washington.edu/phylip.html>. Sur ce site, J. Felsenstein fournit une longue liste de programmes utiles en phylogénétique et disponibles en ligne, ainsi qu'une liste de sites Web complémentaires. Le Professeur J. Irigoien nous a en outre signalé l'existence d'un autre programme utilisant un algorithme différent : GRAPPA (Genome Rearrangements Analysis under Parsimony and other Phylogenetic Algorithms) : <http://www.cs.unm.edu/~moret/GRAPPA/>.

comparer des séquences d'ADN, par exemple. La seule différence entre les deux algorithmes tient au fait que Kitsch comporte une contrainte supplémentaire, à savoir qu'il tient compte d'une « horloge évolutive ». Cela signifie concrètement que tous les objets doivent se situer à égale distance (en termes de longueurs de branches) du point de départ arbitraire du schéma. En systématique zoologique, cela signifie que toutes les espèces étudiées sont contemporaines. Dans notre cas, cette contrainte a peu d'incidence réelle, puisque nous n'utilisons ce programme qu'en vue d'un traitement phénétique. Il faut savoir que, malgré cette contrainte, la longueur des branches n'est arbitraire pour aucun des deux programmes. Les longueurs peuvent toutefois être exagérées, dans certains cas, pour des raisons de lisibilité. Elles restent relatives et strictement descriptives, mais néanmoins interprétables. Les numéros des nœuds (c'est-à-dire des croisement entre deux branches) n'ont aucune signification, ce sont de simples identificateurs.

V. APPLICATION DE LA METHODE PHENETIQUE

A. Des deux phénogrammes présentés ici, le premier a été obtenu sur base de toutes les variantes grâce au programme Fitch, le second sur les seules variantes significatives grâce au programme Kitsch. Ces deux programmes peuvent, dans notre cas, être utilisés pratiquement indifféremment. Les légères variations observées entre l'un et l'autre tiennent à la petite différence de calcul, que nous venons d'expliquer. Ces variations n'ont pas de signification au niveau philologique.

Le phénogramme basé sur toutes les variantes (figure 7) fait nettement ressortir le groupe que nous avons qualifié de groupe M lors de la lecture des MDS. Si l'on prend par exemple M 1, 18 nœuds⁶⁰ le séparent du manuscrit N le plus proche, qui est N 11. Les manuscrits de ce groupe sont les mêmes que dans le MDS, et on retrouve aussi les mêmes manuscrits intermédiaires. Un sous-groupe comprenant M 23, M 22, X 31, M 20, M 12, M 11, M 21 et M 1 est assez nettement dégagé, les parentés entre M 1 et M 21, proches aussi de M 11, d'une part, et X 31 et M 20 d'autre part, sont bien soulignées. Le programme a divisé l'autre groupe en plusieurs sous-groupes, mais qui ne sont jamais séparés les uns des autres par un grand nombre de nœuds. Par exemple, M 9 et M 25 semblent assez isolés des manuscrits N-X, mais il ne faut que sept nœuds pour atteindre le couple N30-X36 ou cinq nœuds jusqu'à N 25, alors qu'il y a 13 nœuds entre ces manuscrits et le plus proche M, qui est M 24. Dans le groupe des N-X, le phénogramme distingue un petit sous-groupe de X (X 7, X 18, X 19, X 27, X 34, X 48), un sous-groupe à majorité N (N 2, N 3, N 4, N 6, N 9, N 10, N 12, N 13, N 19, N 22 et N 28, ainsi que X 23 et X 10), un troisième sous-groupe des N-X restants, et un petit sous-groupe contenant M 9, M 25 et N 26. En outre,

⁶⁰ Les nœuds sont indiqués par des chiffres au croisement de deux branches.

N 7 et N11 apparaissent séparés des autres manuscrits N-X par un plus grand nombre de nœuds. Quelques couples se distinguent par la brièveté des branches qui séparent les deux éléments du couple : X 7 et X 27, X 10 et N 9, N 2 et N 12, N 6 et N 10, X 31 et M 20, M 1 et M 21, M 9 et M 25. Tous ces couples sont confirmés par l'examen des variantes.

Le second phénogramme (figure 8) montre mieux encore les deux groupes principaux : les M d'une part et les N-X d'autre part. Le groupe M est divisé en deux sous-groupes au niveau du nœud numéro 37. Les manuscrits X qui n'étaient pas classés dans les MDS apparaissent sans lien les uns avec les autres et ne sont pas plus proches des M que des N-X. L'agencement des sous-groupes à l'intérieur du groupe N-X est un peu changé par rapport au premier phénogramme et apparaît moins bien. Il semble qu'il ne restait pas assez de variantes discriminantes pour distinguer correctement les manuscrits N-X entre eux.

B. Dans le cas de Basile, les trois méthodes de calculs décrites à propos des MDS ont été utilisées également pour l'établissement des phénogrammes, aussi bien avec Fitch qu'avec Kitsch. Globalement, les deux programmes donnent des résultats similaires, mais dans le détail, ceux de Fitch nous ont paru meilleurs parce qu'ils sont parfaitement cohérents avec les constatations faites sur la base des critères externes. Pour chacune des trois méthodes, les arbres produisent exactement les mêmes groupements que ceux que nous avons constatés dans les graphes des MDS. Pour ne pas répéter ce qui a été dit plus haut, nous avons choisi de ne présenter ici que les résultats obtenus avec Fitch selon la méthode 2, dont nous pensons qu'elle a le plus de chance de correspondre à la réalité.

Le phénogramme (figure 9) fait clairement apparaître les trois grandes familles (texte court, texte long, « syllogè »). Il convient de rappeler que l'arbre n'a pas de racine et n'est pas orienté, si bien qu'il peut être lu depuis n'importe quel endroit ; si on le lit à partir du milieu, au niveau du nœud 7, on constate que l'arbre comporte trois embranchements principaux, qui correspondent à chacune de trois grandes familles. Le phénogramme confirme aussi que les manuscrits de type long constituent un groupe éclaté : bien que clairement apparentés, ils occupent chacun une branche individuelle ; le seul lien plus étroit qu'on puisse reconnaître est celui qui unit L 6 et L 7, confirmant l'impression qui s'était dégagée au moment des collations. Notons aussi que S 14 se retrouve à nouveau parmi les manuscrits de type long. Quant à L 4, il est parmi les manuscrits longs celui qui se rapproche le plus de l'embranchement de la famille des « syllogè », modeste confirmation de la possible affinité que les critères externes avaient fait apparaître. En revanche, l'affinité de L 1 avec la sous-famille constituée de B 1, B 2 et B 3, suggérée par les critères externes, n'est pas confirmée. Dans la famille des « syllogè », les trois groupes attendus se dégagent nettement : le premier, formé de S 6, S 7 et S 8, se démarque des deux autres qui constituent les deux branches d'un sous-groupe commun. Le phénogramme

confirme également la position isolée de S 12 et S 13 et leur relative affinité avec les manuscrits de type court. Ces derniers se répartissent également en trois sous-familles, avec un premier groupe, composé de B 1, B 2 et B 3, se distinguant des deux autres, qui sont les subdivisions d'un même embranchement. A noter aussi la position isolée de B 11, à l'écart de tous les autres manuscrits de type court.

VI. CONCLUSION

A partir des mêmes données, purement textuelles, les deux méthodes mathématiques appliquées permettent de visualiser les groupements des manuscrits d'une manière complémentaire. La lecture de l'information est meilleure dans un phénogramme, parce que les groupements sont plus fins et plus nets. En revanche, le MDS fait mieux apparaître les manuscrits qui s'isolent des groupes, qu'il s'agisse de témoins contaminés ou fortement individualisés. Il est hautement significatif, en tout cas, que les deux méthodes, basées sur des algorithmes totalement différents, génèrent des classifications similaires. Il convient toutefois de rappeler que ces méthodes fournissent une information purement descriptive, non pas analytique. L'interprétation du philologue reste donc indispensable et, en dernier recours, déterminante. Néanmoins, la concordance des résultats obtenus par les deux méthodes avec ceux basés sur les critères externes nous semble offrir une garantie de leur fiabilité. Le philologue est donc en droit d'utiliser ces résultats, soit au moment de choisir les manuscrits pour son édition, soit pour établir l'histoire du texte. Encouragé par les réalisations obtenues, le Centre d'Études sur Grégoire de Nazianze travaille actuellement à améliorer les techniques utilisées. D'une part, une procédure d'encodage permettant de générer des tableaux d'analyse sur base de différentes sélections des données a été mise au point. La sélection pourra porter sur l'analyse philologique des variantes, les critères codicologiques des témoins et la portion du texte à analyser, avec une combinaison possible de ces différents critères⁶¹. Une telle procédure serait utilisable pour n'importe quel type de tradition manuscrite, dans n'importe quelle langue. D'autre part, les recherches ont été poursuivies dans le domaine de la phylogénétique. Nous espérons vivement que ce travail pourra ainsi servir à d'autres philologues.

Caroline Macé - Thomas Schmidt - Jean-François Weiler
Institut Orientaliste
Université Catholique de Louvain
Place Blaise Pascal, 1
B-1348 Louvain-la-Neuve
Belgique

⁶¹ Pour de plus amples renseignements sur les programmes utilisés ou les recherches en cours au Centre d'études sur Grégoire de Nazianze, on peut s'adresser au Prof. B. Coulie (coulie@ori.ucl.ac.be).

ANNEXE
BIBLIOGRAPHIE DES METHODES INFORMATISEES
APPLIQUEES A L'ECDOTIQUE

Depuis les travaux fondateurs de Dom Froger⁶², différentes tentatives d'utilisation des statistiques et de l'informatique dans le domaine de l'ecdotique ont vu le jour. J. Duplacy⁶³ avait dressé une bibliographie raisonnée de ces travaux jusqu'en 1975. Il serait bon que ce travail soit repris et continué. Tel n'est pas l'objet de ces pages, où nous ne donnons que quelques repères utiles.

Un tour d'horizon de ce domaine de la recherche avait été fait en 1978, à l'occasion d'un colloque à Paris:

La pratique des ordinateurs dans la critique des textes. Actes du Colloque international 579, Paris, 29 - 31 mars 1978, J. GLENISSON (éd.), Paris, 1979.

Un recueil de travaux publié récemment sous le titre « Studies in Stemmatology » fait en quelque sorte le point sur la situation aujourd'hui. On ne saurait trop en recommander la lecture:

Studies in Stemmatology, P. VAN REENEN, M. VAN MULKEN (eds), Amsterdam / Philadelphia, 1996.

Les méthodes statistiques, sont bien exposées dans les travaux déjà anciens :

J.G. GRIFFITH, *Numerical taxonomy and some primary manuscripts of the Gospel*, in *Journal of Theological Studies*, N.S., t. 20, 1969, p. 389-406.

—, *The Interrelation of some Primary Manuscripts of the Gospels in the Light of Numerical Analysis*, in *Studia Evangelica*, t. VI (T. U., 112), Berlin, 1973, p. 221-238.

—, *Non-stemmatic classification of manuscripts by computer methods*, in *La pratique des ordinateurs dans la critique des textes. Actes du Colloque international 579, Paris, 29 - 31 mars 1978*, édités par J. GLENISSON (éd.), Paris, 1979, p. 74-86.

—, *A Three-Dimensional Model for Classifying Arrays of Manuscripts by Cluster Analysis*, in *Studia Patristica*, t. 15 (T. U., 128), Berlin, 1984, p. 79-83.

Plus récemment, T.J. Finney a appliqué la même méthode à l'épître aux Hébreux:

T.J. FINNEY, *Mapping the Textual History of Hebrews*, in *Revue Informatique et Statistiques dans les Sciences Humaines*, t. 33, 1997, p. 125-147.

⁶² J. FROGER, *La critique des textes et son automatisation*, Paris, 1968 (*Initiation aux nouveautés de la science*, t. 7). Dom Froger a lui-même une dette à l'égard de H. QUENTIN, *Essais de critique textuelle (ecdotique)*, Paris, 1926.

⁶³ J. DUPLACY, *Classification des états d'un texte, mathématiques et informatique : repères historiques et recherches méthodologiques*, in *RH.*, t. 5, 1975, p. 249-309 (réimprimé dans J. DUPLACY, *Etudes de critique textuelle du Nouveau Testament*, présentées par J. DELOBEL, Leuven, 1987 (*Bibliotheca Ephemeridum Theologicarum Lovaniensium*, 78), p. 193-257).

Une méthode, purement mathématique, a été mise au point dans les années septante:

A. DEES, *Considérations théoriques sur la tradition manuscrite de Lai de l'ombre*, in *Neophilologus*, t. 60, 1976, p. 481-504.

—, *Over Stambomen van handschriften*, in *Forum der Letteren*, t. 18, 1977, p. 63-78.

—, *Ecdotique et informatique*, in *Actes du XVIII^e Congrès International de Linguistique et de Philologie Romanes. Université de Trèves (Trier) 1986*, D. KREMER (ED.), t. VI, Tübingen, 1988, p. 18-27.

—, *Analyse par l'ordinateur de la tradition manuscrite du Cligès de Chrétien de Troyes*, in *op. cit.*, , p. 62-75.

Elle a été utilisée et critiquée, entre autres :

K.-H. UTHEMANN, *Ordinateur et stemmatologie. Une constellation contaminée dans une tradition grecque* (traduction A. SPOORENBERG), in *Distributions spatiales et temporelles, constellations des manuscrits. Etudes de variation linguistique offertes à A. Dees à l'occasion de son 60^{ème} anniversaire*. Textes présentés par P. VAN REENEN et K. VAN REENEN-STEIN, Amsterdam / Philadelphia, 1988, p. 265-277.

—, *Editionstechnik ohne « Trenn- und Bindfehler »? Zur Rekonstruktion von Stammbäumen mittels des Computerprogramms von A. Dees*, in *JÖB.*, t. 39, 1989, p. 49-54.

—, *Which Variants are Useful in Discovering the Deep Structure of the Manuscript Tradition of a Text? Contra a so-called Essentially Quantitative Approach*, in *Studies in Stemmatology*, P. VAN REENEN, M. VAN MULKEN (eds), , Amsterdam / Philadelphia, 1996, p. 249-261.

La cladistique a été plusieurs fois expérimentée comme méthode de classement des manuscrits:

A.R. LEE, *Numerical Taxonomy Revisited : John Griffith, Cladistic Analysis and St. Augustine's Quaestiones in Heptateuchum*, in *Studia Patristica*, t. 20. *Papers presented to the Tenth International Conference on patristic studies held in Oxford 1987. Critica, Classica, Orientalia, Ascetica, Liturgica*, E.A. LIVINGSTONE (ed.), Leuven, 1989, p. 24-32.

H.D. CAMERON, *The Upside-down Cladogram. Problems in Manuscript Affiliation*, in *Biological Metaphor and Cladistic Classification. An Interdisciplinary Perspective*, H.M. HOENIGSWALD, L.F. WIENER (eds), Philadelphia, 1987, p. 227-242.

P. ROBINSON, *Computer-Assisted Analysis and « Best-Tex » Historical Editing*, in *Studies in Stemmatology*, P. VAN REENEN, M. VAN MULKEN (eds), Amsterdam / Philadelphia, 1996, p. 123-134.

R.J. O'HARA et P.M.W. ROBINSON, *Computer-Assisted Methods of Stemmatic Analysis*, in *The Canterbury Tales Projects Occasional Papers*, t. I, N. BLAKE and P. ROBINSON (eds), Oxford, 1993, p. 53-74.

B. SALEMANS, *Cladistic and the Resurrection of the Method of Lachmann. On Building the Stemma of « Yvain »*, in *Studies in Stemmatics*, P. VAN REENEN et M. VAN MULKEN (eds), , Amsterdam / Philadelphia, 1996, p. 3-70.

P. ROBINSON, *New Methods of editing, exploring, and reading The Canterbury Tales*, Talk at the Conference *I Nuovi orizzonti della filologia*, Accademia Nazionale dei Lincei, Rome, May 28 1998, in *The Canterbury Tales Projects. Websit* by M. Pidd (<http://www.shef.ac.uk/uni/projects/ctp>).

Enfin, plusieurs méthodes ont été expérimentées de manière isolée :

P. DERRON, *Classement sur Mc Intosh des manuscrits des « Vers d'or » pythagoriciens*, in *RHT.*, t. 26, 1996, p. 253-268.

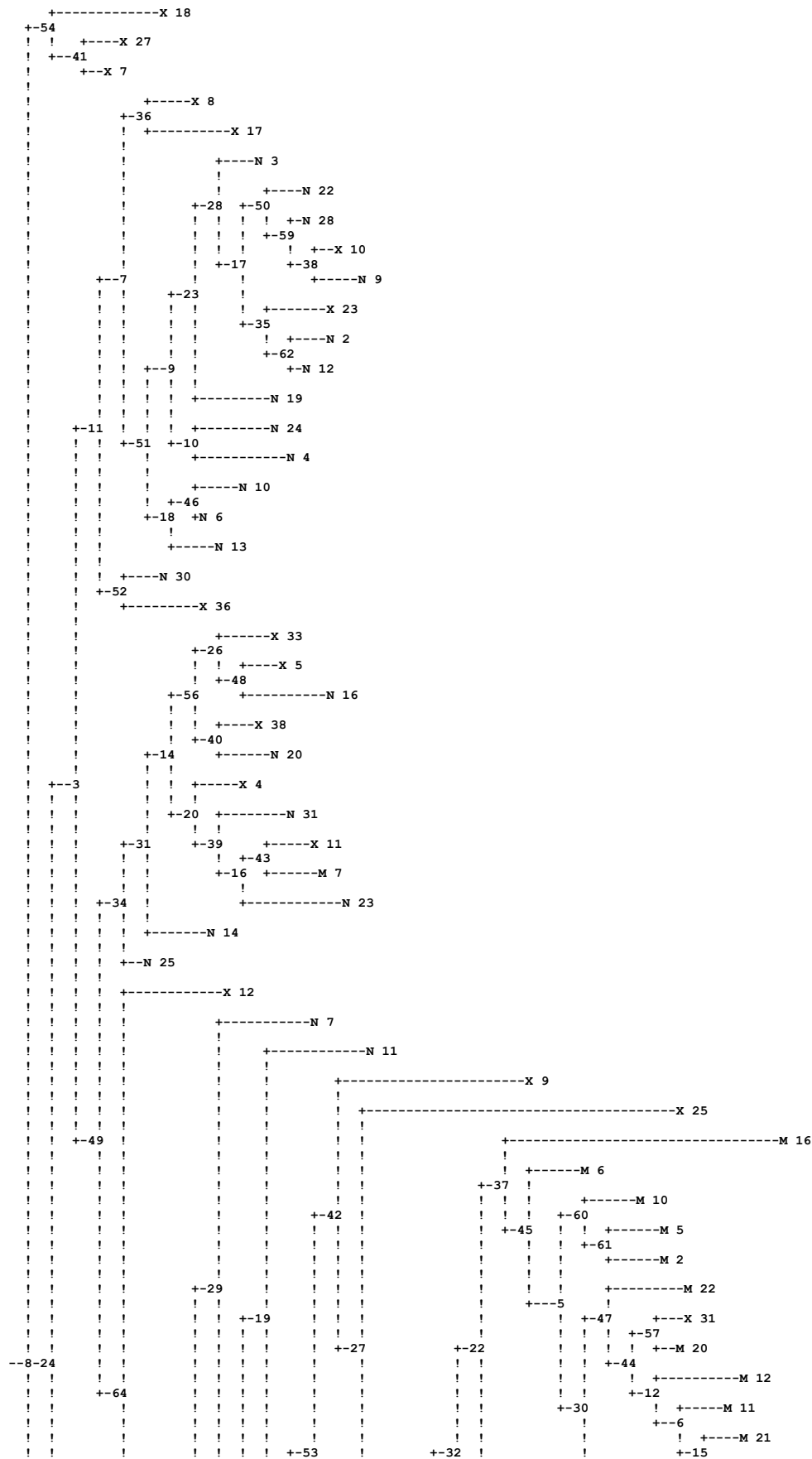
Intéressant surtout pour la démarche philologique :

C. FLIGHT, *Stemmatic Theory and the Analysis of Complicated Traditions*, in *Manuscripta*, t. 36, 1992, p. 37-52.

—, *A Complete Theoretical Framework for Stemmatic Analysis*, in *Manuscripta*, t. 38, 1994, p. 95-115.

L'auteur prétend fournir une méthode formalisée valable pour toutes les traditions, mais elle n'est guère utilisable.

Phénogramme (Fitch) – Discours 38 - Toutes les variantes



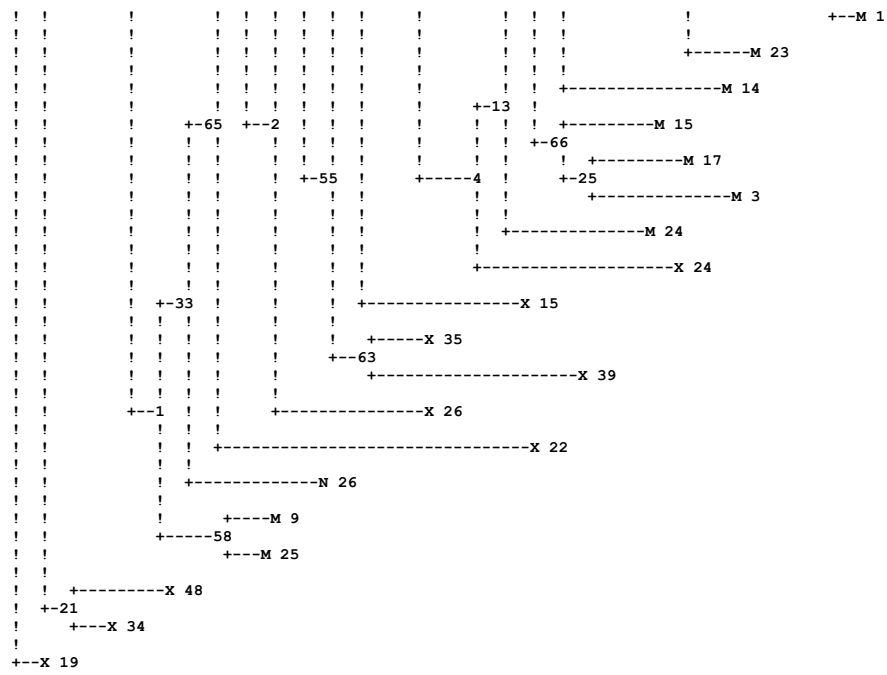
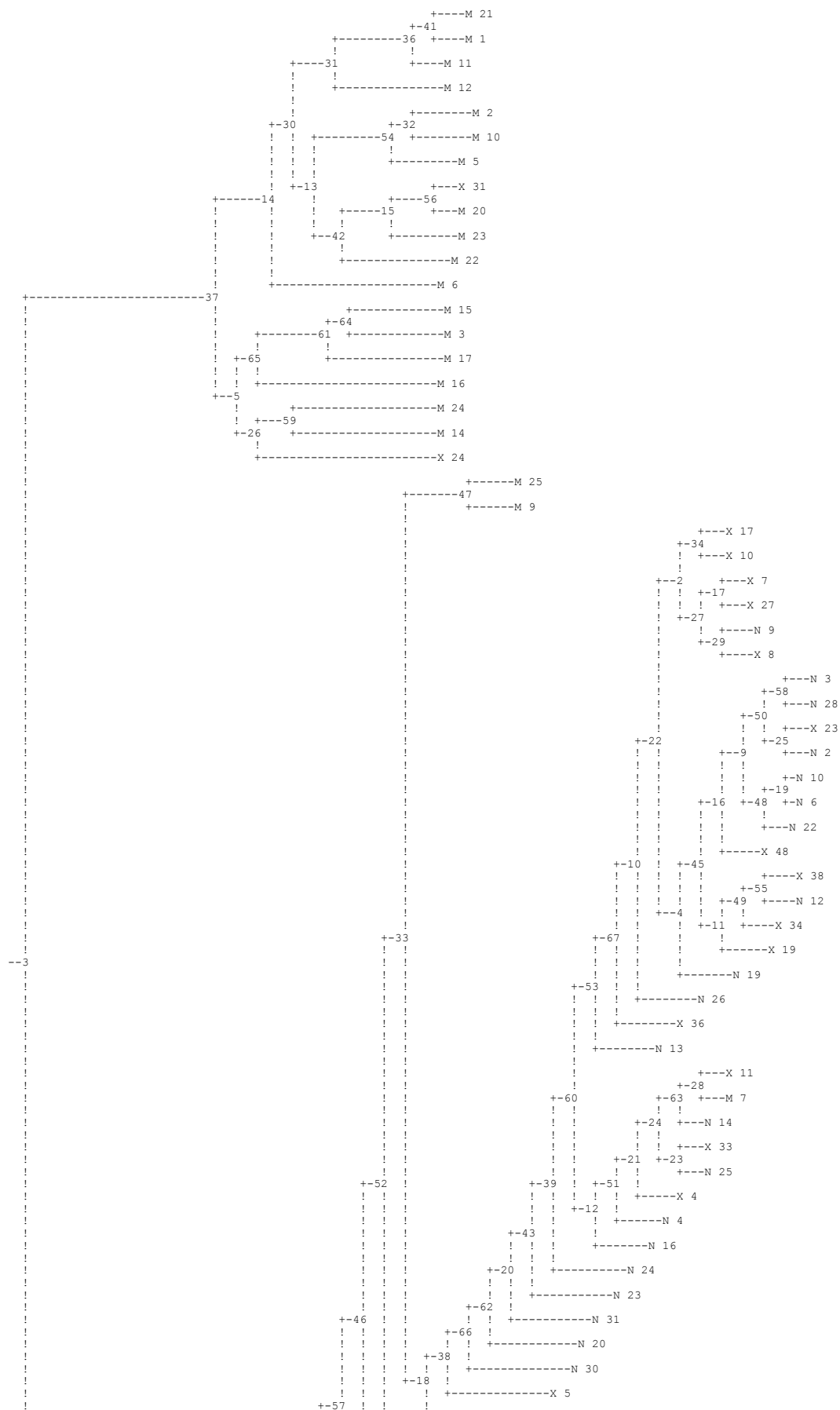


Figure 7

Phénogramme (Kitsch) - Discours 38 - Variantes significatives



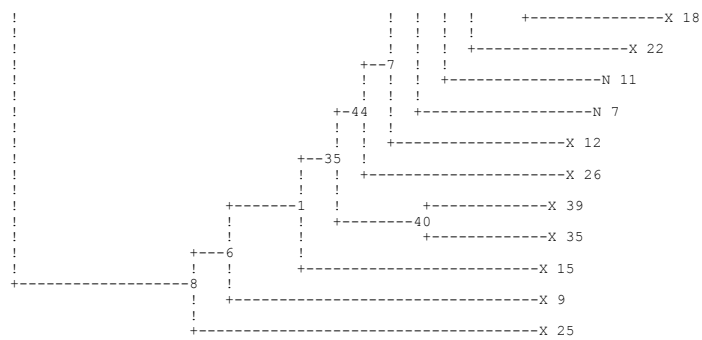


Figure 8

Phénogramme (Fitch) – Commentaire au Disc. 38 – Méthode 2

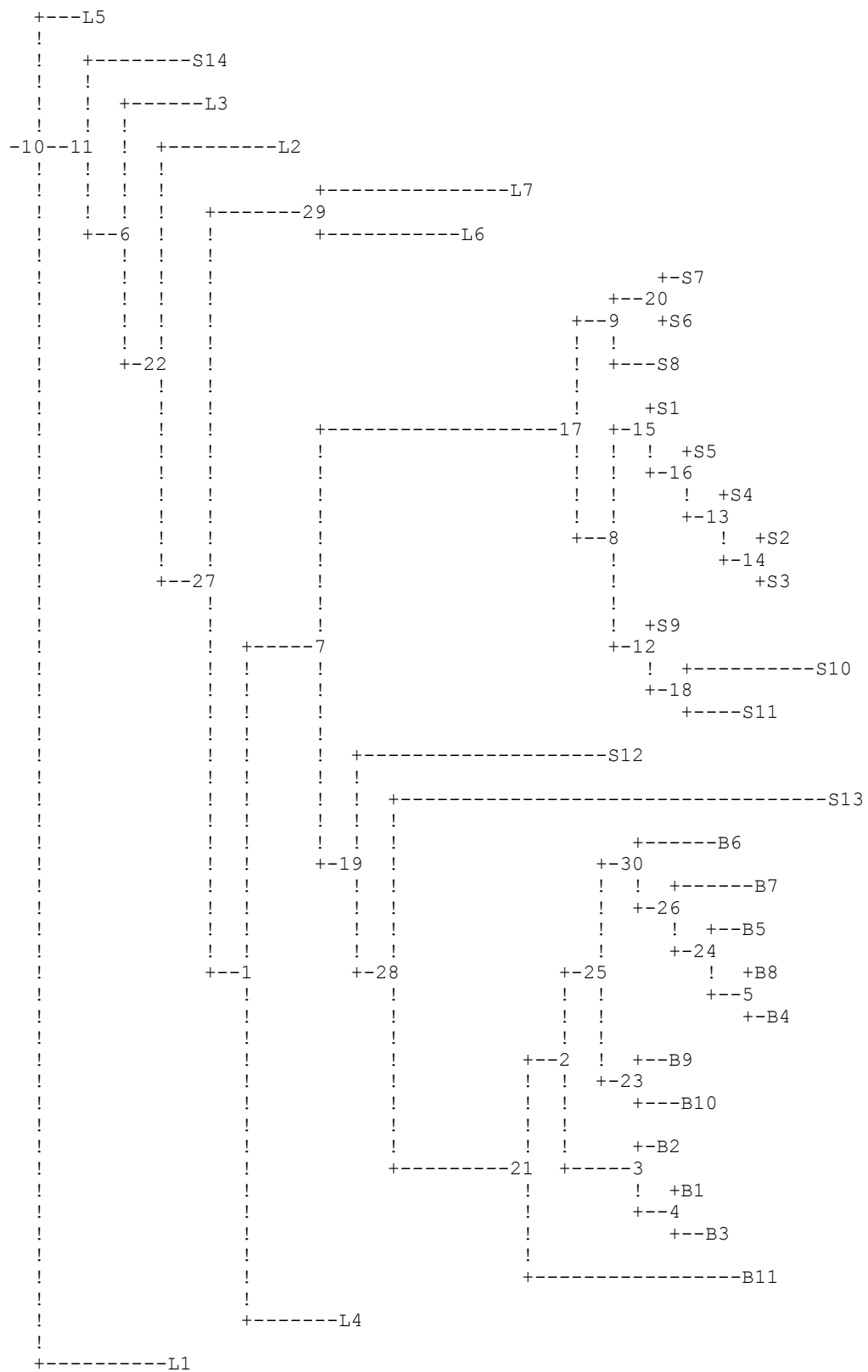


Figure 9