

J.-P. BENZÉCRI

Typologie de textes grecs d'après les occurrences des formes des mots outils

Les cahiers de l'analyse des données, tome 16, n° 1 (1991),
p. 61-86

http://www.numdam.org/item?id=CAD_1991__16_1_61_0

© Les cahiers de l'analyse des données, Dunod, 1991, tous droits réservés.

L'accès aux archives de la revue « Les cahiers de l'analyse des données » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

TYPOLOGIE DE TEXTES GRECS D'APRÈS LES OCCURRENCES DES FORMES DES MOTS OUTILS

[TEXTES GRECS]

J.-P. BENZÉCRI

1 Thème de la recherche et préparation des données

1.1 Le corpus des textes

1.1.1 Autour du Nouveau Testament

Le cahier n°1 du Volume XV est entièrement consacré à la statistique linguistique et à la stylométrie. En particulier, l'article [NOUV. TEST. GREC.], présente une "Typologie des chapitres du texte grec du Nouveau Testament d'après les occurrences des formes des mots outil". Sous ce titre sont compris, les chapitres de trois des quatre Évangiles: Marc, Luc et Jean; ceux des Actes des Apôtres; de l'Apocalypse; et de trois épîtres: l'Épître aux Romains, L'Épître aux Hébreux et la première Épître de Saint Jean; avec, en référence externe, les 13 chapitres du Livre I de la *Politique* d'Aristote: soit, au total, 158 chapitres.

Le présent travail porte sur 439 chapitres ou fragments. Dans cet ensemble, prédomine, comme précédemment, le texte grec du Nouveau Testament, saisi maintenant dans son entier; mais on a également multiplié les références externes, dont toutefois la plupart approchent du corpus de base à un titre ou à un autre; l'étude stylométrique de la langue grecque classique dans son ensemble n'étant pas encore abordée. Nous énumérerons donc les textes saisis dans l'ordre de leur distance croissante au Nouveau Testament.

De l'Ancien Testament, nous avons pris les 66 chapitres du Livre du Prophète Isaïe et les 4 du livre de Malachie, dans la version grecque des Septante; en suivant l'édition de la *Sainte Bible Polyglotte* publiée à Paris, de 1900 à 1909, par Monseigneur VIGOUROUX.

1.1.2 Les Antiquités Judaïques

Sur l'histoire du peuple hébreux, l'une des rares sources qui ne dépende pas exclusivement de la Bible, mais utilise des textes anciens qui ne nous sont pas parvenus, est l'œuvre de Flavius JOSÉPHE: *Les Antiquités Judaïques*, ΙΟΥΔΑΪΚΗ ΑΡΧΑΙΟΛΟΓΙΑ. Après avoir exercé des fonctions sacerdotales au

temple de Jérusalem, l'auteur remplit une mission à Rome; revint dans la Ville Sainte en 65 de notre ère, alors que le pays était en pleine révolte; participa à la lutte contre les Romains; puis, s'étant rendu à la discrétion de Vespasien, il revint à Rome, prit le nom de Flavius en hommage à la nouvelle dynastie impériale, et écrivit en langue grecque des ouvrages historiques où il se proposait de faire connaître son peuple, en exhaltant la pureté du monothéisme tout en minimisant l'opposition de ses coréligionnaires au pouvoir romain.

Des XX livres des *Antiquités Judaïques*, qui se déroulent de la Genèse au règne de Néron, nous avons saisi, d'une part, les deux premiers tiers du Livre XVIII; et, d'autre part, la première moitié du Livre X.

La partie retenue du Livre XVIII rapporte une période qui correspond exactement aux événements du Nouveau Testament: depuis le recensement contemporain de la Nativité, jusqu'aux vicissitudes de la vie fastueuse d'un Agrippa, père de celui devant qui comparut Saint Paul. Deux passages célèbres citent même les noms de Jésus et de Jean Baptiste. Jésus étant reconnu comme le Christ dans le texte des manuscrits les plus anciens que nous possédions, il apparaît que ce texte n'est pas tel que Josèphe l'avait écrit, mais a été repris par des chrétiens. La critique contemporaine n'estime cependant pas qu'il s'agisse d'une interpolation pure et simple, mais seulement de retouches; et la grande presse (*Figaro* du 16 Février 1972) nous a, voici quelques années, rendu l'écho des recherches de deux savants israéliens, proposant une version possible du texte original, établie d'après un auteur de langue arabe.

Le début du Livre X rapporte les règnes d'Ézéchias, Manassé, Amos et Joas; et l'on y voit le Prophète Isaïe.

Nous reviendrons sur le contenu des textes de Josèphe au §1.3, quand nous en expliquerons le découpage en fragments, dont plusieurs ne coïncident pas avec les chapitres de la recension DINDORF utilisée par nous.

Quant à la langue ou au style de JOSÈPHE, il vaut la peine de citer ce jugement du *Dictionnaire des Œuvres*:

“Son style, assez bon dans *la Guerre Judaïque* (écrite à l'origine en langue araméenne et traduite par la suite en grec), est dur et négligé dans les *Antiquités*.”

1.1.3 Le Pasteur d'HERMAS

L'un des premiers en date des écrivains chrétiens de langue grecque est HERMAS. On s'accorde à placer dans la première moitié du II-ème siècle son séjour à Rome; mais tout ce qu'on a pu dire d'autre sur sa personne est contesté. HERMAS n'est pas seulement chrétien, il écrit dans la manière du Nouveau Testament (encore qu'avec une platitude qui rehausserait, s'il était besoin, la gloire céleste de son modèle!): nous n'osons dire qu'il s'agisse d'une

falsification, ni même d'un pastiche; mais plutôt d'une convention de style; à laquelle, cependant, plus d'un chrétien se laissa prendre; car le texte de son *Pasteur* dut être officiellement banni du *Canon* du Nouveau Testament, par un décret du Pape Gélase, au V-ème siècle.

Quant au contenu, on lit dans le *Dictionnaire* déjà cité:

“L'œuvre est formée de trois parties distinctes: dans la première, l'Église, sous l'aspect d'une matrone, apparaît à l'auteur en cinq 'visions' successives, pour lui montrer, toujours sous forme allégorique, la gravité de ses péchés et de ceux des autres chrétiens, et lui indiquer comment Dieu, dans son infinie bonté, en considération de la ruine qui menace le monde, donne la possibilité de se purifier une seconde fois après le baptême. Dans la dernière vision de la première partie et dans toute la seconde, qui comprend douze 'préceptes', et a un caractère plutôt moral, un ange apparaît à Hermas sous l'aspect d'un Pasteur - d'où le titre de l'œuvre - pour mieux lui expliquer ce qu'il a déjà vu et doit enseigner aux autres”.

Avec un tel contenu, on attendrait des réminiscences de l'Apocalypse; mais aussi des paraboles évangéliques.

Nous avons saisi, en suivant l'édition de R. JOLY, dans les Sources Chrétiennes, les 25 chapitres, ou §, en lesquels est subdivisée la première partie du *Pasteur*.

1.1.4 Les Éthiopiennes d'HÉLIODORE

Est-il vraisemblable que l'auteur du plus long des romans grecs qui aient été conservés soit mort évêque de Trikka, en Thessalie, après avoir magnifié, dans son diocèse, le célibat ecclésiastique? Assurément, car le XVII-ème français nous a légué, de Monseigneur Jean-Pierre CAMUS, évêque de Belley, *Iphigène, ou la Rigueur Sarmatique*; et de Monseigneur Pierre-Daniel HUET, évêque d'Avranches, une *Diane de Castro*; œuvres romanesques, celle-ci péruvienne, celle-là polonaise dont qui cherche compagnie pour voyager en chambre ne peut manquer de faire ses délices! Nul ne prétend d'ailleurs qu'HÉLIODORE fût évêque alors qu'il conduisait *Théagène et Charichlée*, en dépit de mille obstacles, jusqu'au plus heureux dénouement; mais seulement qu'il se convertit ensuite; et, de païen, devint chrétien puis évêque.

Quoi qu'il en soit de l'auteur, l'œuvre, hérissée d'épisodes piquants où l'on peut tout craindre, ne sort jamais des fragiles limites de la bienséance; et, dans son introduction à l'édition Guillaume Budé, J. MAILLON affirme que:

“L'idéal de chasteté que tout romancier grec propose à l'admiration du public ... atteint sa plus haute expression dans les Éthiopiennes”.

En retenant les 33 chapitres dont se compose le Livre I de l'œuvre d'HÉLIODORE, nous avons cité en témoin la prose dont, en ce temps que Gaston

BOISSIER a appelé *la Fin du Paganisme*, des lecteurs délicats s'étaient flatté le palais, avant de se nourrir de l'Évangile.

1.1.4 Politique et Éthique d'Aristote

Outre les 13 chapitres du Livre I de la *Politique* déjà pris en compte dans [NOUV. TEST. GREC], on trouvera dans la présente étude les chapitres, également au nombre de 13, du Livre I de l'Éthique à Nicomaque.

1.2 L'ensemble des formes de mots

Nous l'avons écrit dans [NOUV. TEST. GREC]: il semble qu'au moins en grec, le dénombrement des formes des mots outil offre enfin au philologue la base d'une stylométrie précise... Comme en latin, les pronoms varient avec le cas, le genre et le nombre: or l'analyse du §4 de [LING. TRI.] a montré que les cas d'un même mot peuvent avoir des profils d'emploi bien différents; mais, de plus, il y a, comme en français, un article défini. Les prépositions sont nombreuses et souvent expriment des nuances précises.

Dans la présente étude comme dans la précédente, toutes les analyses et classifications effectuées se fondent sur un tableau croisant l'ensemble des chapitres ou fragments avec un ensemble de 96 formes de mots outil: le même que celui introduit au §3 de [NOUV. TEST. GREC]. Cet ensemble faisant, au §4, l'objet de commentaires, il suffira de rappeler ici qu'en ont été bannies les formes des 1-ère et 2-ème personnes (du verbe être ou des pronoms); formes dont la distribution est, plus encore qu'au style, liée au contenu de l'œuvre.

On citera, à propos des pronoms, le bref compte rendu (in [TEXTE BIBLE], CAD, Vol. XI, n°1, pp. 14-15, 1986) d'un travail (non publié) de Peng Zheng-Ce:

“L'opposition entre le théâtre et les romans apparaît magnifiquement: mais c'est un classique du genre (qui a touché la Chine par cette analyse): le théâtre se distingue du roman parce qu'il contient plus de dialogue et que, de ce fait, les pronoms WO (je) et NI (tu) y prédominent sur la forme TA (il) de la 3-ème personne. On élimine donc les pronoms. Subsiste une opposition sur le 1-er axe entre théâtre et roman; ... maintenant, sur l'axe 2, les pièces de théâtre et le roman de l'auteur LAO SHE se trouvent groupés du même côté de l'origine, tout en étant séparés sur l'axe 1...”

À propos du dénombrement des formes, voici ce qu'a bien voulu nous écrire Étienne ÉVRARD:

“... si, comme nous le faisons, vous notez les formes de descripteurs relatifs non seulement aux lemmes, mais aussi aux catégories morphologiques, vous êtes en mesure d'étudier *et* les lemmes *et* les particularités telles que les personnes grammaticales, les temps verbaux etc. Il me semble qu'ainsi, les éléments qui entrent dans la composition du texte sont mieux distingués.”

Il n'y a rien à objecter aux remarques d'un Maître! Il n'est pas expédient, non plus de justifier autre chose qu'une hiérarchie de choix dans les voies ardues de la stylométrie. Nous écrivions dans [NOUV. TEST. GREC]:

On peut craindre que les résultats surprenants de finesse dont nous rendons compte, ne se reproduisent pas avec une autre langue où les rapports de subordination et de dépendance ne seraient pas aussi clairement marqués qu'ils ne le sont dans le grec, par les formes des mots outil; ce qui dispense de chercher ces rapports au prix d'une difficile analyse sémantico-syntaxique.

En toute rigueur, il est assuré qu'un texte doit être regardé comme composé d'éléments, divers quant à la forme et à la fonction, que VENDRYÈS comprend tous sous le même nom de morphèmes. Il y a, d'une part, les bases de mots porteuses de sens qui sont, en bref, ce qu'É. ÉVRARD appelle lemmes, et qu'on trouve dans un lexique du français comme noms communs, adjectifs et verbes à l'infinitif; et, d'autre part, des outils qui peuvent être des formes isolées (conjonctions, prépositions...), des désinences ('aient'; où il faut d'ailleurs distinguer: imparfait de l'indicatif; 3-ème personne; et: pluriel), ou des procédés de composition (mettre tel mot avant ou après tel autre...).

Mais comme l'on a d'autant moins achevé l'inventaire des morphèmes d'aucune langue que la limite est imprécise entre mot qu'il faut prendre avec son sens (et rapporter à son lemme) et forme ou locution figée (e.g. la locution prépositionnelle: 'en vertu de') qui échappe à sa base pour n'être qu'un outil, nous avons choisi de commencer par mettre à l'épreuve cette hypothèse que le dénombrement facile de formes dont la tradition orthographique a distingué les limites (d'ailleurs non sans arbitraire) suffit à édifier, sur un fondement purement matériel, une typologie générale des textes; et à suggérer de cette typologie une interprétation.

1.3 De la saisie du texte au tableau de correspondance

On a suivi la voie exposée au §2 de [NOUV. TEST. GREC]: le texte est mis en forme en conservant les esprits et le iota souscrit (en fin de mot), tout en respectant l'ordre alphabétique usuel du grec; tout texte est converti en une longue liste où chaque mot (occurrence) est écrit sur une ligne de rang impair, la ligne suivante, de rang pair, donnant les numéros du chapitre et du verset (ou, plus généralement, de l'alinéa, s'il s'agit d'un texte profane où les repères sont introduits par nous).

La seule innovation concerne le programme 'tridic§'; lequel, étant donné la liste des occurrences d'un texte (triées alphabétiquement et numérotées par chapitres et versets) et un lexique, c'est-à-dire un ensemble de formes rangées dans l'ordre alphabétique, permet de construire (sous forme de texte) un tableau de correspondance croisant l'ensemble *I* des formes avec l'ensemble *J* des chapitres du texte (ou avec l'ensemble des versets).

Ainsi que nous l'annoncions au §1 de [NOUV. TEST. GREC], il ne convient pas de s'arrêter à la division des œuvres en chapitres; laquelle n'est sans doute que le produit de l'activité contingente de copistes et d'éditeurs; il faut expérimenter sur le découpage; éventuellement, adopter une division du texte en épisodes.

Celle-ci peut maintenant être introduite sans modifier directement le texte déjà saisi (où les limites de chapitres sont signalées par le numérotage des versets ou alinéas), à l'aide d'une troisième option 'A' de "cumul par épisodes" que nous avons ajoutée au programme 'tridic§'. De façon précise, si l'on a créé un fichier auxiliaire (dont le nom doit obligatoirement être 'txt§Ax', si le texte à subdiviser est 'txt'), définissant la subdivision du texte en épisodes, 'tridic§' ayant constaté la présence de ce fichier, propose à l'utilisateur l'option 'A', après 'V' et 'C', en affichant à l'écran la question suivante:

"faut-il cumuler les occurrences
par verset (V) ou par chapitre(C) ou autrement(A) "

et, si la réponse est 'A', le programme se base sur le fichier auxiliaire pour ventiler les occurrences des formes.

Voici, par exemple, la structure d'un fichier B:FLWz§Ax, pour subdiviser la partie retenue du Livre X des *Antiquités Judaïques* (fichier B:FLWz). Après

Subdivisions du livre X des Antiquités Judaïques		
FW1a	Sénacherib menace Ézéchias	1 8
FW1b	Oracles d'Isaïe; fin de Sénacherib	1 19
FLW2	Maladie d'Ézéchias; oracles d'Isaïe	2 13
FLW3	Règne de Manassé	3 13
FW4a	Amos roi; puis Joas le pieux	4 8
FW4b	Olda prophétise la ruine du peuple	4 12
FW4c	Pénitence; et célébration de la Pâque	4 23

une ligne de titre, terminée par un aller à la ligne, on a, pour chaque épisode, une ligne commençant par le sigle choisi (e.g. 'FW1a' pour le premier fragment du chapitre 1; sigle de 4 caractères au plus), suivi éventuellement d'un commentaire ne comportant aucun chiffre (pour FW1a: "Sénacherib menace Ézéchias"); puis les numéros de chapitre et de verset pour le dernier verset de l'épisode (ces numéros étant séparés par un nombre quelconque d'espaces). Le début, quant à lui, est, tacitement '1 1' pour le premier épisode; et, pour les épisodes suivants, le verset qui suit immédiatement le dernier verset de l'épisode précédent; par exemple, le deuxième épisode ira de '1 9' à '1 19', bornes incluses.

Un épisode peut enjamber sur plusieurs chapitres.

Si le dernier épisode spécifié ne se termine pas au dernier verset du texte saisi, le programme crée un dernier épisode, dont le sigle est 'FIN', qui comprend les versets restants. Dans le cas présent, il n'y a pas d'épisode 'FIN', car la saisie du texte s'arrête à '4 23'.

Subdivisions du livre XVIII des Antiquités Judaïques	
FLV1	Quirinius en Judée; le recensement 1 9
FLV2	Les écoles de pensée chez les Juifs 2 22
FL3a	En Judée de Quirinius à Pilate 3 18
FL3b	Parthes, Mèdes et Arméniens 3 38
FL4a	Pilate à Jérusalem; Jésus Christ 4 15
FL4b	Scandale à Rome au temple d'Isis 4 37
FLV5	Juifs à Rome; Samaritains; renvoi de Pilate 5 10
FL6a	Vitellius à Jérusalem 6 8
FL6b	Vitellius et Artaban; mort de Philippe 6 25
FL7a	Hérode; Hérodiade; Jean-Baptiste 7 14
FL7b	Campagne interrompue de Vitellius 7 23
FL7c	Hérode et sa famille 7 35
FL8a	Agrippa, ruiné, séjourne en Orient 8 17
FL8b	Agrippa à Rome; temporisation de Tibère 8 32
FL8c	Arrestation d'Agrippa 8 45
FL8d	Agrippa en prison 8 52
FL8e	Tibère établit son héritier 8 65
FL8f	Mort de Tibère et accession de Caligula 8 74

On a de même créé un fichier B:FLVz§Ax pour subdiviser le texte entier (B:FLVz) du Livre XVIII des *Antiquités*.

[Signalons que l'édition Didot de la recension DINDORF propose deux divisions en chapitres; selon notre commodité, nous avons saisi pour le livre XVIII la division la plus fine; et la plus grossière pour le livre X.]

2 Les méthodes taxinomiques

Même s'il est loin de représenter l'héritage littéraire grec, l'ensemble des textes déjà saisis est d'une grande diversité. On ne peut donc escompter en faire la synthèse en deux ou trois dimensions: les graphiques plans issus de l'analyse des correspondances n'en offriront qu'une première vue. Il faut poser dans sa généralité le problème taxinomique.

Les quelques centaines de chapitres qu'on a saisis seront ici regardés comme l'analogie d'une collection de quelques centaines d'insectes. Dans cette collection, dix ou vingt espèces sont représentées chacune par plus de dix spécimens; d'autres par un moindre nombre. Des espèces voisines doivent être réunies au sein d'un genre; si la collection révèle un nombre assez élevé de genres, on pourra, s'élevant encore d'un degré dans la hiérarchie, reconnaître plusieurs familles.

Imaginons qu'un entomologiste défie un statisticien de reconnaître un ordre dans cette collection!

Le statisticien, dont l'œil n'est point averti des détails significatifs que présente un insecte, demandera au moins qu'on ne lui soumette pas telle quelle la collection; mais qu'on lui donne, de chaque spécimen une fiche descriptive d'un format normalisé, comportant un ensemble de mesures, toujours les mêmes, assez rigoureusement définies pour qu'on puisse, en quelques leçons, apprendre à les saisir sans hésitation.

De l'ensemble de fiches, les principes de codage propres à l'Analyse des Données, permettent de passer à un ensemble de points situés dans un espace multidimensionnel. Si deux ou trois dimensions suffisent à étaler les points-insecte sans que se trouvent juxtaposés des spécimens différant grandement entre eux, le statisticien embrassera, en quelque sorte, du regard, la diversité de la collection: il percevra des amas locaux de forte densité qui lui suggéreront des espèces; des groupes d'amas qui seront des genres... ; et peut-être relèvera-t-il le défi de l'entomologiste.

Mais si, pour traduire adéquatement les rapports de similitude en rapports de voisinage (deux individus étant d'autant plus semblables qu'une moindre distance les sépare), il faut construire un espace à huit dimensions, les représentations géométriques usuelles ne révéleront pas des espèces; et encore moins leurs définitions en terme de caractères mesurés. On devra recourir à la taxinomie numérique.

Procédant par pure induction, la Classification Ascendante Hiérarchique (CAH) constitue d'abord, à partir de l'ensemble des individus, des paires de proches voisins; puis des agrégats de plus grand effectif, fondés sur la similitude; jusqu'à obtenir une structure ramifiée comprenant tout l'ensemble. De plus, elle caractérise les agrégats en terme de variables descriptives. Le statisticien a, ici encore, une chance de rencontrer les notions familières à l'entomologiste; voire d'éclairer et de perfectionner la taxinomie en usage.

Si, laissé seul, le statisticien s'égaré dans le labyrinthe qu'il a lui-même construit et ne reconnaît pas de genre ou d'espèce acceptable, il pourra demander qu'on lui communique la répartition en espèces de la totalité, ou seulement de la moitié des spécimens; et, d'après ces données complémentaires, construire une définition géométrique du système des classes. La méthode étant, en bref, de définir pour chaque espèce un centre, ou point moyen des individus qu'on sait être compris dans l'espèce; l'affectation d'un individu à une espèce se faisant en déterminant le centre qui en est le plus proche. C'est l'Analyse Discriminante.

Même si le statisticien connaît *a priori* l'espèce de chaque individu, il n'est pas sûr qu'*a posteriori* l'analyse discriminante réussisse; qu'elle fournisse une règle simple, purement mathématique, pour calculer l'espèce en fonction des mesures; il y aura un taux de succès; et, complémentairement, un taux d'erreur. Si l'espèce n'a été prise en compte que pour une moitié des spécimens, le taux d'erreur résultant de la règle de discrimination sera généralement plus élevé pour l'autre moitié.

Or c'est précisément sur cette autre moitié (appelée, pour cette raison: *échantillon d'épreuve*) que l'analyse discriminante pourrait faire la preuve d'une réelle utilité; laisser espérer qu'à l'avenir, il soit possible, sur des individus nouvellement récoltés, de se fier au statisticien pour réaliser, automatiquement, des déterminations concordant avec celles que ferait un naturaliste.

Revenons aux chapitres. Il n'y a pas, en stylistique, de taxinomie établie; ni même de caractères pertinents dont l'observation serait la base des appréciations formulées par les experts... Nous partons ici de l'hypothèse que le dénombrement systématique des formes de mots outil offre une telle base. Il n'y a pas d'espèce ni de genre *stricto sensu*; mais on peut présumer qu'à moins d'hétérogénéité manifeste, les chapitres d'une œuvre rentrent tous dans une même espèce; et que plusieurs œuvres écrites dans un même but par un seul auteur, ou par des auteurs apparentés, rentrent dans un même genre; ce terme étant, d'ailleurs, employé déjà en stylistique, encore qu'avec moins de rigueur qu'en entomologie.

Ceci étant posé, on suivra avec les chapitres, la même voie statistique qu'on suivrait avec des insectes. On placera au sein d'un espace multidimensionnel l'ensemble des individus et l'ensemble des caractères; on effectuera des CAH dont le succès sera apprécié d'après la concordance existant entre la hiérarchie des classes construites inductivement par calcul et la répartition *a priori* des chapitres en œuvres et des œuvres en genres.

On recourra aussi à l'analyse discriminante; en assimilant provisoirement, faute de mieux, les œuvres à des espèces; le centre d'une œuvre étant calculé soit comme moyenne de tous ses chapitres, soit à partir d'une moitié seulement de ceux-ci; avec, ici encore, l'occasion de mettre à l'épreuve, sur les taux de succès et d'erreurs, la justesse de notre hypothèse qu'une stylométrie peut se fonder sur le dénombrement des formes de mots outil.

C'est de ce point de vue qu'on présentera les résultats obtenus. Nous avons fait de multiples CAH et analyses discriminantes, parce que ces méthodes se prêtent à de nombreuses variantes; d'une part, selon le nombre de dimensions retenues (c'est-à-dire le nombre d'axes factoriels utilisés); et d'autre part selon que, dans les données elles-mêmes, on restreint selon divers critères, l'ensemble des textes ou l'ensemble des formes. Seuls méritent d'être retenus les résultats que l'on retrouve dans toutes les analyses; ou du moins dans toutes celles qui se conforment à des principes clairement formulés issus d'une pratique.

Nous considérerons d'abord (au §3) les groupements de chapitres, à l'exclusion des groupements de formes (de mots outil) qui cimentent en quelque sorte ces classes. En effet, d'une part, les classes de chapitres s'apprécient suivant des critères de cohésion accessibles même au lecteur qui ignore la grammaire grecque; et, d'autre part, les grammairiens eux-mêmes, se bornent à nous signaler l'absence de certaines formes à certaines époques ou dans la langue commune; mais laissent tout à découvrir sur les affinités qu'ont entre elles les formes de mots dont la fréquence s'élève ou s'abaisse simultanément selon le style. Mais l'intérêt de ces recherches est tel que nous soumettrons (au §4) quelques exemples de classifications des formes de mots outil dans leurs rapports avec les classes de chapitres.

Discrimination sans échantillon d'épreuve

.1	Luca	(20+ 9)/24	{Mth3;J2;Act7,9;Jac1;Is36,37;Mlc2;TsB1}
2	Jean	(24+ 0)/26	
3	Acta	(25+ 5)/28	{Rm16;Mth2,3;Hrm10;TmB3}
4	Apoc	(22+ 3)/22	{Act6;Is12,15}
5	PltA	(13+ 2)/13	{Ncm3,13}
6	Hebr	(12+ 7)/13	{Rm15;PtA1,2;Hrm22;Eph1;TsA3;TmB4}
7	Roma	(14+15)/16	{Héb3;CrB9;Jac3;PtA3,5;Glt1,3,4,6;Eph3,4;Php3;Cls1,3;TsB2}
8	Mtth	(23+ 5)/28	{Luc12,17;Hrm24;Mrc13;Mlc1}
9	CorA	(15+10)/16	{J15;Mth6;Jac2,4;IohC;Glt5;Eph5;TsA2;TmA3,5}
10	CorB	(12+25)/13	{CrA1;PtA4;PtB1,2;IohB;Glt2;Eph2,6;Phpl,2,4;Cls2,4;TsA1,4,5;TsB3;TmA1,2,4;TmB1,2;Tit1,2;Phlm}
11	Herm	(22+ 3)/25	{Iuda;TmA6;Tit3}
12	Marc	(15+ 1)/16	{Mth4}
13	EthA	(32+ 1)/33	{FW4a}
14	NicA	(11+ 0)/13	
15	Isaï	(62+ 5)/66	{Luc21;Jac5;PtB3;Mlc3,4}
16	FlvJ	(25+ 1)/25	{Luc3}

3 Classification des chapitres et discrimination entre œuvres

Nous présentons d'abord les résultats d'analyses discriminantes. Puis des classifications purement inductives.

3.1 Analyse discriminante

On part du tableau (349×96) croisant chapitres et formes. On en écarte 5 des formes, lesquelles sont apparues (cf. §4) trop nettement caractéristiques de certaines œuvres pour qu'on puisse, en leur présence, percevoir les nuances qui nous intéressent.

Quant aux chapitres, tout lecteur du Nouveau Testament connaît les généalogies placées dans le chapitre I de Matthieu et le chapitre III de Luc: ces textes stéréotypés sont rendus en français celui-ci avec la formule 'fils de', celui-là avec 'engendra'; en grec, il y a, dans les deux cas, répétition de mots outil. Il est apparu que ces chapitres étaient non seulement écartés du type de l'œuvre à laquelle ils appartiennent, mais susceptibles de troubler nos constructions géométriques en créant des corrélations inopportunes. Dans les analyses discriminantes dont on rend compte ici, les cumuls définissant les centres de Matthieu et de Luc ne comportent pas ces chapitres.

Nous considérons deux analyses discriminantes: dans chacune de celles-ci, on crée des centres uniquement pour les 16 œuvres dont nous avons colligés au moins 10 chapitres ou fragments, soit: les quatre Évangiles {Mtth, Marc, Luca, Jean}; les actes des Apôtres (Acta); 4 épîtres (aux Romains: Roma; aux Corinthiens I et II: CorA & CorB; aux Hébreux: Hébr); l'Apocalypse (Apoc); le livre d'Isaïe (Isaï); le Pasteur (Herm); les Antiquités de Flavius Josèphe (FlvJ, les parties des livres X et XVIII étant ensemble); les Éthiopiennes (EthA: livre I); et deux livres d'œuvres d'Aristote (Politique I: PltA; éthique à Nicomaque I: NicA).

Discrimination avec les lignes paires en épreuve

principaux	supplément	Syn	Act	Ioh	Epi	Apo	Isa	Hrm	Flv	Eth	Ari	Hbr
LUCA (9/11)	Luca (10/12)	2	3	1	3	0	4	1	0	0	0	0
JEAN (10/11)	Jean (9/10)	0	0	5	0	0	0	1	0	0	0	0
ACTA (13/14)	Acta (12/14)	2	0	0	5	0	1	0	0	0	0	1
APOC (11/11)	Apoc (11/11)	0	0	0	0	0	3	0	0	0	0	0
PLTA (6/ 6)	PltA (4/ 7)	0	0	0	1	0	0	0	0	0	1	0
HEBR (7/ 7)	Hebr (5/ 6)	0	0	0	8	0	0	1	0	0	0	0
ROMA (7/ 8)	Roma (6/ 8)	0	0	0	13	0	0	0	0	0	0	0
MTH (14/14)	MctH (10/13)	4	0	0	1	0	0	2	0	0	0	0
CORA (8/ 8)	CorA (6/ 8)	0	0	1	5	0	0	0	0	0	0	0
CORB (6/ 6)	CorB (6/ 7)	0	0	0	23	0	0	0	0	0	0	0
HERM (9/12)	Herm (8/13)	0	0	0	3	0	0	0	0	0	0	0
MARC (7/ 8)	Marc (8/ 8)	0	0	0	0	0	0	2	0	0	0	0
ETHA (16/17)	EthA (15/16)	0	0	0	0	0	0	0	0	0	0	0
NICA (6/ 6)	NicA (6/ 7)	0	0	0	0	0	0	0	0	0	3	0
ISAÏ (31/33)	Isaï (29/33)	1	0	0	1	0	2	1	0	0	0	0
FLVJ (13/13)	FlvJ (12/12)	0	0	0	0	0	0	0	0	2	0	0

Dans une première analyse, il n'y a pas, pour les œuvres, d'échantillon d'épreuve: les centres sont calculés par cumul de tous les chapitres; (mis à part MthI et LucIII). Encouragé par la qualité de la discrimination obtenue, on a fait une deuxième analyse, en réservant comme échantillon d'épreuve les chapitres qui correspondent dans notre tableau (349×91) à des lignes de rang pair (ce qui équivaut à ne prendre en compte dans les cumuls que les lignes de rang impair).

Dans l'une et l'autre analyse, on affecte au centre le plus proche tous les chapitres; ceux des 16 œuvres les plus longues comme ceux des autres œuvres jugées trop brèves pour qu'on puisse en définir valablement un centre typique. La première analyse est dépouillée en donnant d'abord, pour chaque œuvre, une fraction, telle que (20+9)/24 pour Luc: dont le dénominateur est, naturellement, le nombre des chapitres de l'œuvre (l'Évangile selon Saint Luc compte 24 chapitres); et où le numérateur donne, après le nombre de chapitres correctement affectés (20 pour Luc), celui des chapitres divers (de l'une des 15 autres œuvres principales; ou d'une autre œuvre) affectés aussi à ce centre, comme plus proches de lui que d'aucun autre; puis, après la fraction, vient la liste de ces chapitres divers: pour Luc, le Ch. III de Matthieu, le second de Jean, ..., jusqu'au 1-er de la 2-ème Épître aux Thessaloniens. Il faut signaler que le centre de l'Évangile selon Saint Jean est calculé d'après ses 21 chapitres; mais comme les 5 cap. de la première Épître de Jean vont constamment avec son Évangile, on prend, en bref, pour dénominateur 26=21+5.

Pour la deuxième analyse discriminante, on donne d'abord deux fractions concernant respectivement les éléments principaux (lignes de rang impair: avec, pour plus de clarté, un sigle en capitales: LUCA) et le supplément (lignes de rang pair): de Luc, sont correctement affectés 9/11 des premiers (non compris Luc III!) et (10/12) des seconds. Ensuite, au lieu d'une liste des affectations diverses qui pourrait être trop longue, un bilan de celles-ci, ventilées suivant 11 postes, choisis en fonction de connaissances *a priori* et des résultats de multiples

analyses. Les sigles de ces postes, qui nous serviront maintes fois dans la suite, se déchiffrent sans peine:

‘Syn’ est l'ensemble des 3 Évangiles *synoptiques* {Mtth, Marc, Luc} (ainsi nommés parce que leurs récits parallèles ont été, de tout temps, mis en vis-à-vis, ou *synopse*); soit, au total 68 chapitres (66, sans Mth I et Luc III);

‘Act’, comprend les 28 chapitres des Actes des Apôtres;

‘Ioh’, réunit 21 chapitres de l'Évangile de Jean et 5 de sa 1-ère Épître;

‘Épi’, comprend toutes les épîtres du Nouveau Testament; à l'exception de Jean I (comprise dans ‘Ioh’) et de l'Épître aux Hébreux, dont la langue est particulière et qui est comptée à part sous ‘Hbr’; soit 103 chapitres;

‘Apo’ est l'Apocalypse: 22 chapitres;

‘Isa’, compte avec les 66 chapitres d'Isaïe les 4 de Malachie;

‘Hrm’ = 25 cap. du Pasteur; ‘Flv’ = 25 fragments de Flavius; ‘Éth’ = 33 cap. des Éthiopiennes; ‘Ari’ groupe 13 chapitres de la Politique et 13 de l'Éthique à Nicomaque; ‘Hbr’ = 13 cap. de l'Épître aux Hébreux.

Il faut cependant prendre garde que, dans le dépouillement de l'analyse discriminante, ces sigles sont utilisés d'une façon particulière: ils s'entendent à l'exclusion des chapitres de l'œuvre prise pour centre. Par exemple, quand on parle de 2 chapitres des Synoptiques attachés à Luc, il faut comprendre: chapitres de Mth ou de Marc; de même, on dénombre, rattachés à la 2-ème Épître aux Corinthiens, CorB, 23 chapitres des *autres* épîtres; au sein desquelles CorB apparaît donc comme particulièrement représentative d'un type moyen.

On est surpris des taux élevés d'affectations exactes obtenus dans la deuxième analyse: (173/185) pour les éléments de base; (157/185) pour l'échantillon d'épreuve. Quant aux éléments principaux, on peut invoquer pour explication le grand nombre de variables (les 91 formes) qui engendrent un espace de dimension 90, au sein duquel il y a place pour des cloisons séparant nos 16 œuvres: le succès étant toutefois que les hyperplans médiateurs (des segments joignant les centres) constituent précisément ces cloisons. Mais un succès pour l'échantillon d'épreuve ne peut s'expliquer que par le fait que les chapitres pris en compte pour calculer les centres suffisent à caractériser les œuvres dans leur ensemble; qu'il y a, dans les mots outil, une signature propre à chaque œuvre. L'analyse du tableau (16×91) sur lequel est fondée la première discrimination (sans échantillon d'épreuve) permet de caractériser cette signature en mettant en rapport œuvres et mots; et tenant compte de classifications arborescentes: conformément au plan adopté, nous réservons ces considérations au §4.

Ncap	Nfac	Acta	Hermas	Hébreux
349	8	cls:15pur; syn:8	eclaté	clas:7 (&Act...)
349	25	clas:23pur	prph:16pur	eclaté
349	90	syn:21pur+...	syn:16=pur (&Héb)	syn:9 (&16Hrm)
347	8	cls:16pur; syn:11	clas:12pur+...	classique:10...
347	25	syn:19pur + 7syn	syn:19pur	épîtres:9
347	90	syn:21pur+...	syn:15=pur (&Héb)	syn:9 (&15Hrm)
16oe	15	((Act, Héb), Hrm) ->syn	((Act, Héb), Hrm) ->syn	((Act, Héb), Hrm) ->syn

3.2 Analyses factorielles et classifications des chapitres

On ne peut tout tenter! Ni même rendre compte de tout ce qu'on a tenté! De notre expérimentation, nous retenons 6 classifications: fondées sur deux analyses factorielles distinctes, de chacune desquelles on retient soit 8, soit 25, soit 90 facteurs.

Pour les analyses factorielles, on a d'abord gardé en principal Mth I et Luc III: nous estimons avoir gagné quelque peu à les éliminer ensuite; et croyons utile de prendre le lecteur à témoin de ce progrès. Quant au nombre de facteurs, 90 est le maximum possible: classer d'après 90 facteurs équivaut à accepter la distance distributionnelle sans la filtrer par l'analyse factorielle; 25 facteurs correspondent dans nos analyses à environ 60% de l'inertie du nuage; avec 8 facteurs, on a seulement un tiers de l'inertie, mais il y a, entre les facteurs de rang 8 et 9, une certaine dénivellation.

Les CAH ont en commun certaines caractéristiques générales qui sont à la base de leur dépouillement. Les Synoptiques sont intimement soudés; l'Évangile selon Saint Jean est totalement isolé; les Épîtres, (la première Épître de Saint Jean et l'Épître aux Hébreux étant mises à part,) font un bloc, généralement bien reconnu par la méthode inductive de la CAH; le langage prophétique, typiquement représenté dans notre corpus par la version des Septante d'Isaïe, est aussi celui de Malachie et de l'Apocalypse. Quelque distinction qu'on doive faire entre un philosophe d'époque classique comme Aristote, un Flavius Josèphe, historien d'époque impériale, dont le style a pu passer pour négligé (cf. §1.1.2) et les galaneries des brigands d'un roman Hellénistique, les trois se rejoignent, face au reste du corpus pour représenter un certain classicisme.

En revanche, Actes, Hermas et Hébreux font problème. Témoins, peut-être, de manières d'écrire non exactement représentées dans le corpus mais dont certaines tendances sont attestées, ils voguent au gré des analyses, tout en se présentant souvent par gros fragments dont la place correspond toujours à des hypothèses acceptables: telles que: écrits par Saint Luc, les Actes sont apparentés aux synoptiques; mais maints discours prononcés en milieu hellénistique, maints récits d'aventures au sein de ce monde se ressentent d'un air qui n'est pas celui de la Judée ni de la Galilée. Certains groupes de chapitres, où l'on traite en détail de la religion de l'Ancienne Loi sont précisément, presque toujours, classés avec les synoptiques.

classification des 349 chapitres en 17 classes d'après 8 facteurs

	Syn	Act	Ioh	Epi	Apo	Isa	Hrm	Flv	Eth	Ari	Hbr
prph	855	5	1	0	1	8	20	0	0	0	0
	861	2	0	0	1	1	44	4	0	0	1
	824	0	0	0	0	2	4	0	0	0	0
	848	0	0	0	0	11	1	1	0	0	0
	782	0	0	0	0	0	0	4	0	0	0
	786	10	0	0	1	0	1	0	0	0	0
	857	38	3	1	0	0	0	0	0	0	1
	839	9	8	0	0	0	0	0	0	0	0
	836	0	0	8	0	0	0	0	0	0	0
	854	0	0	2	0	0	0	4	0	0	0
	846	1	0	15	0	0	0	0	0	0	0
	856	1	15	0	12	0	0	4	3	0	7
	853	1	1	0	0	0	0	1	17	2	0
	816	0	0	0	0	0	0	0	5	20	1
	858	0	0	0	1	0	0	5	0	11	25
	859	0	0	0	35	0	0	0	0	0	0
	860	1	0	0	52	0	0	2	0	0	1
tot	68	28	26	103	22	70	25	25	33	26	13

On ne s'étonne pas qu'Herma, de par le contenu du *Pasteur*, touche aux synoptiques et aux textes prophétiques; c'est aussi un texte littéraire; dont la langue peut avoir été châtiée sur les mêmes modèles qui ont servi à mettre en grec une Épître aux Hébreux à laquelle on a pu légitimement supposer un original sémitique.

classification des 349 chapitres en 17 classes d'après 25 facteurs

	Syn	Act	Ioh	Epi	Apo	Isa	Hrm	Flv	Eth	Ari	Hbr
prph	859	4	1	0	3	6	5	16	0	0	1
	854	3	0	0	0	2	33	0	0	0	1
	836	0	0	0	0	11	2	0	0	0	0
	827	0	0	0	0	3	2	0	0	0	0
	842	0	0	0	0	0	27	0	0	0	0
	790	0	0	0	1	0	0	4	0	0	0
	822	11	0	0	0	0	0	0	0	0	1
	855	46	4	1	3	0	1	2	0	0	2
	856	0	0	16	0	0	0	0	0	0	0
	846	0	0	9	1	0	0	0	0	0	0
	858	2	23	0	1	0	0	2	5	1	3
	860	0	0	0	0	0	0	1	20	25	2
	857	0	0	0	1	0	0	0	0	7	24
	802	0	0	0	3	0	0	0	0	0	0
	849	1	0	0	33	0	0	0	0	0	0
	861	1	0	0	14	0	0	0	0	0	4
	841	0	0	0	43	0	0	0	0	0	1

Nous laissons le lecteur interpréter par lui-même les résultats des classifications et critiquer le bilan qui en est proposé sur un tableau dont les 6 premières lignes se rapportent aux 6 CAH des quelque 430 chapitres; et la dernière à la CAH des centres des œuvres (dont l'exposé est renvoyé au §4). Il ne nous a pas paru bon d'étaler tous les détails de 6 CAH: nous avons fait des tableaux bilan, en face de chacun desquels est un arbre en réduction qui suffit à donner la structure hiérarchique qui surmonte la partition retenue. À cette

classification de 347 chapitres en 16 classes d'après 8 facteurs

	Syn	Act	Ioh	Epi	Apo	Isa	Hrm	Flv	Eth	Ari	Hbr
proph	844	1	1	0	2	10	7	0	0	0	0
	850	0	0	0	0	10	4	0	0	0	0
	825	0	0	0	0	2	4	0	0	0	0
	856	2	0	0	0	0	50	0	0	0	0
	777	0	0	0	0	0	0	4	0	0	0
	847	0	0	1	2	0	2	12	0	0	3
	853	3	16	0	7	0	2	4	4	0	5
	838	0	0	0	0	0	0	0	17	2	0
	810	0	0	0	0	0	0	0	4	20	1
	857	0	0	0	0	0	3	0	11	25	0
	855	0	0	1	40	0	0	0	0	0	1
	852	1	0	0	51	0	0	2	0	0	1
	821	13	0	0	1	0	1	0	0	0	0
	858	45	11	1	0	0	0	0	0	0	1
	835	0	0	9	0	0	0	0	0	0	0
	846	1	0	14	0	0	0	0	0	0	0

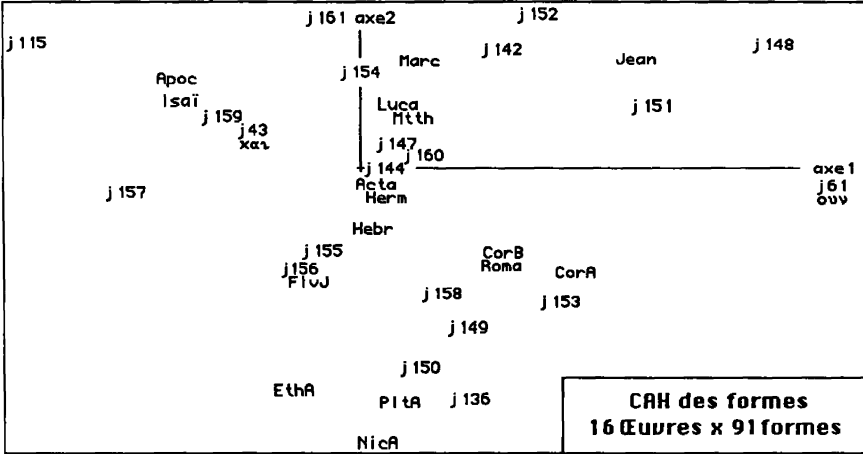
Notre préférence irait aux CAH de 347 chapitres d'après 25 ou 8 facteurs; celle-ci réussit un filtrage efficace de l'information, aboutissant à un tableau parfaitement cohérent, malgré sa complexité à laquelle on s'étonne qu'un espace de dimension 8 puisse suffire; celle-là, avec 25 facteurs donne des résultats d'une agréable compacité; il faut toutefois noter qu'à la différence de ce qui est le cas pour les 5 autres CAH, la partition retenue n'est pas celle définie par les nœuds les plus hauts; mais qu'on a fait son profit de l'occasion qui s'offrait de choisir les nœuds en sorte que disparaissent toutes les subdivisions mineures ou non immédiatement interprétables (même si une étude stylistique approfondie devait ultérieurement en expliquer la pertinence).

classification de 347 chapitres en 13 classes d'après 25 facteurs

	Syn	Act	Ioh	Epi	Apo	Isa	Hrm	Flv	Eth	Ari	Hbr
prph	856	3	0	0	3	4	36	0	0	0	2
	844	0	1	0	2	14	0	0	0	0	0
	864	0	0	0	0	3	28	0	0	0	0
	785	0	0	0	1	0	4	0	0	0	0
	837	0	0	0	0	1	19	0	0	0	1
	851	15	0	0	0	3	0	0	0	0	1
	839	2	19	0	0	0	0	0	0	0	0
	855	46	7	1	0	0	1	0	0	0	0
	862	0	0	24	0	0	0	0	0	0	0
	865	0	0	1	96	0	2	0	0	0	9
	840	0	1	0	0	0	1	22	2	0	0
	829	0	0	0	0	0	0	3	24	1	0
	852	0	0	0	1	0	0	0	7	25	0

La livre d'Isaïe est cependant partagé entre les classes 856 et 864: division qui n'est pas fortuite; car des 28 chapitres que renferme 864, 23 ont un rang ≤ 35 , les autres étant {49,55,58,60,64}; sans aucun des chapitres historiques 36 à 40; lesquels sont dans 856, plus proches de l'Apocalypse (844).

Par souci de brièveté, on n'a rien dit des subdivisions qui, dans les CAH, permettent de retrouver, au moins en partie, une distinction entre les deux livres d'Aristote: Éthique à Nicomaque et Politique; ce dernier volontiers associé à quelques exploits des chevaleresques brigands du delta du Nil...



4 Classification des formes de mots outil et correspondance entre œuvres et formes

On considère désormais les mots et classes de mots, en se bornant à deux analyses. Le §4.1 est fondé sur le tableau (16x91) construit pour la première analyse discriminante (sans échantillon d'épreuve) et croisant 16 œuvres avec 91 formes. Le §4.2 prend pour principal le tableau (437x91), avec en supplément les cumuls des 16 mêmes œuvres; et il retient la partition des chapitres en 13 classes effectuée dans l'espace engendré par les 25 premiers axes.

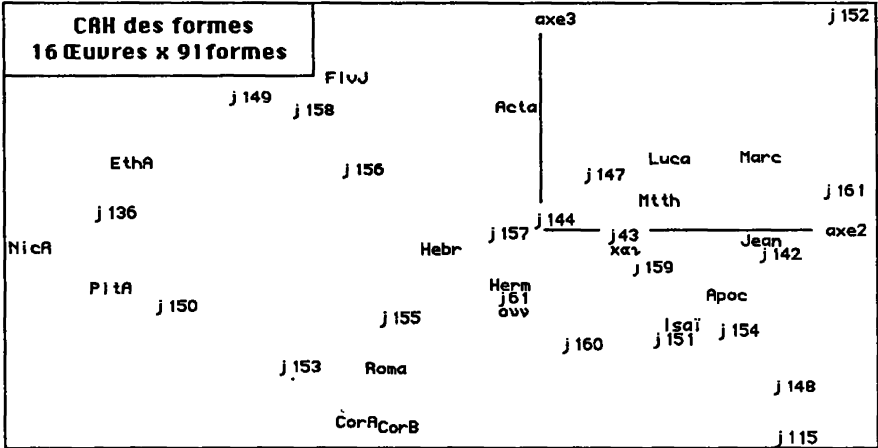
4.1 Croisement de 16 œuvres avec 91 formes

cumuls de 16 oeuvres x 91 formes

trace :	2.358e-1									
rang :	1	2	3	4	5	6	7	8	9	10
lambda :	538	497	361	213	190	125	99	72	66	61 e-4
taux :	2282	2108	1532	903	804	530	419	303	281	257 e-4
cumul :	2282	4389	5922	6824	7629	8159	8578	8881	9162	9419 e-4

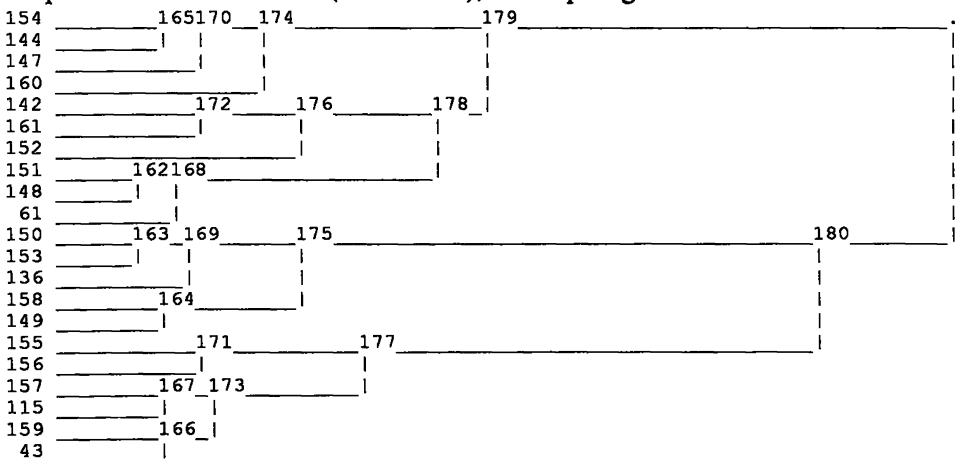
Figurent sur les graphiques d'analyse factorielle, d'une part les 16 centres afférents aux cumuls des œuvres, d'autre part les centres des 21 classes de formes retenues dans la partition définie par les 20 nœuds les plus hauts de la CAH; les classes réduites à un mot sont marquées avec ce mot.

Autour d'Isaïe et de l'Apocalypse (situés dans le quadrant $F1 < 0, F2 > 0$), on a 4 classes caractérisant le style prophétique; en particulier $\kappa\alpha\iota$, 'et', fréquent dans les locutions calquées à partir de l'hébreux; mais la classe la plus excentrée est 115, dont les deux mots traduisent la formule "et il n'y aura pas non plus...". Dans le quadrant ($F1 > 0, F2 > 0$), on trouve l'Évangile selon Saint Jean, associé à donc, $\text{o}\upsilon\upsilon$, et à 148, {où, de nouveau, de même que, pour que}. Le long de l'axe 2, se pressent les autres œuvres, qui s'étaient dans le plan (2,3).



Dans le plan (2,3) on trouve les caractéristiques de la langue classique ($F2 < 0$); des Épîtres ($F3 < 0; F2 < 0$); des Synoptiques ($F2 > 0$). Bien que fortement corrélé à l'axe 3 ($COR3(Acta) \approx 500$), Acta (les Actes des Apôtres) apparaît atypique en ce qu'aucune classe de mots n'en est proche. Quant à Herm et Hébr, ils sont très peu corrélés aux axes 1 à 5: ce explique que, comme on l'a noté au §3, les rapports entre ces deux œuvres et les autres soient incertains.

Le détail du contenu des classes de mots est proposé sur un tableau: les lignes sont groupées en blocs, selon la hiérarchie supérieure de la CAH; et chaque classe est étiquetée par les sigles (en grec) des œuvres où la fréquence en est particulièrement élevée (voire basse); cet étiquetage concorde avec celui de la



ci-dessus: Arbre de la CAH des 91 formes d'après 16 œuvres

Luca		161++			
Mtth	142+		152+		
Marc	152+++		43+		
Hébr	147++	19	147+	24	Hébr:150+,156+
Acta	149+				
Herm	61++	159+++			
Jean	152++		142++	151+	148+
CorA	151++		160++		
Roma	144+	150+			CorA:149+
CorB	148+				CorA,B:153++
FlvJ	144+		151=0	156++	158+
EthA	161-		136+	157+	149+
PltA	155++		136+	155+	156+
NicA	136++				
Apoc	154++	156+			43+
Isaï	157+	115++			

Œuvres étiqueté par les classes de formes

CAH des œuvres, sans que toutefois les deux soient identiques, l'importance d'une œuvre dans la caractérisation d'un groupe de mots n'impliquant pas que, réciproquement, le groupe de mots prédomine dans le caractère de l'œuvre.

ci-dessous: Partition des formes en 21 classes, étiquetées d'après 16 œuvres

154	α< ΕΧ ΕΝΩΡΙΟΝ	ΑΠΟΚ+++++	ΜΑΘ-
144	του οσ< τω, ου<	ΑΠΟΚ+	ΡΩΜ+
147	εξ αυτου αυτουσ μετα εις νυν ου< του		ΕΒΡ+
160	τη, εν ου ουκ μη ουτωσ<		ΡΩΜ+ ΚΟΡαβ+

142	οτι< ο<		ΙΩ++ ΜΑΘ+
161	οτε< μετ αυτου αυτουσ απο απ εωσ< ο<		ΜΑΘ+
152	ην ουτωσ< αυτω, αυτοισ ησαν αυτου	ΙΩ++	ΜΑΘ+ ΜΑΡΚ++ ΛΟΥΚ+

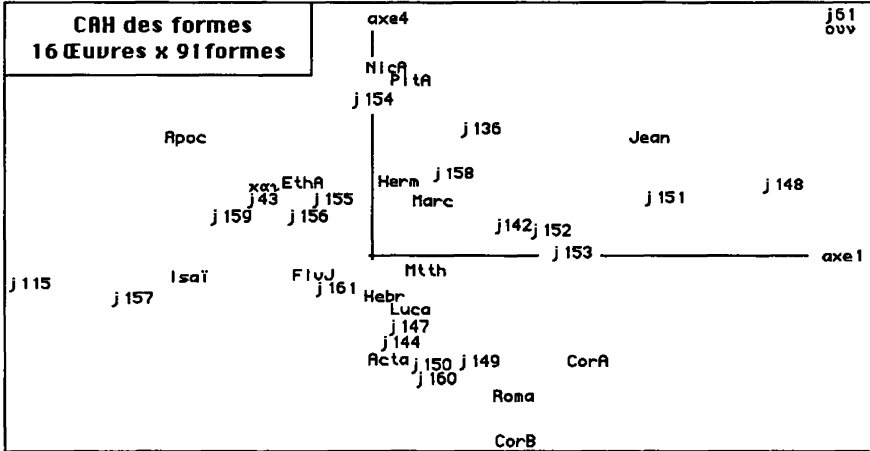
151	ουχι οταν εαν εστι	ΙΩ++	ΚΟΡα++ ΦΛΑ≠0 ΑΙΘ≠0
148	οπου< παλιν καθωσ ινα<	ΙΩ++++	ΚΟΡαβ+ ΜΑΡΚ+ ΑΙΘ≠0
61	ουν	ΑΠΟΚ- ΙΩ++++	ΡΩΜ+ ΕΡΜ++ ΗΣΑ≠0

150	δρα κατα	ΕΒΡ++ ΡΩΜ+++	ΚΟΡβ+ ΝΙΚ+
153	αυτο τουτο τισ πωσ ει αλλ αλλα		ΚΟΡαβ+ ΝΙΚ+
136	αν τι η	ΑΠΟΚ- ΠΟΛ++	ΝΙΚ++++

158	περι ταυτα παρα προσ	ΑΠΟΚ- ΠΟΛ++	ΜΑΘ- ΑΙΘ+ ΦΛΑ+
149	συν υπο< δε	ΑΠΟΚ≠0	ΠΡΑΣ+ ΑΙΘ+ ΦΛΑ+

155	η< επι τα ουτε το	ΠΟΛ++	
156	τησ τουσ τοισ την των	ΑΙΘ+ ΦΛΑ++	

157	επ εψ επι ωσ<	ΑΙΘ+	ΗΣΑ+++ ΦΛΑ++
115	ουδε εσται		ΗΣΑ+++++
159	ουτοι< εισι αυτη, αυτην τασ αυτησ αυτων ταισ αι<		ΕΡΜ++ ΗΣΑ+
43	κα<		ΗΣΑ+



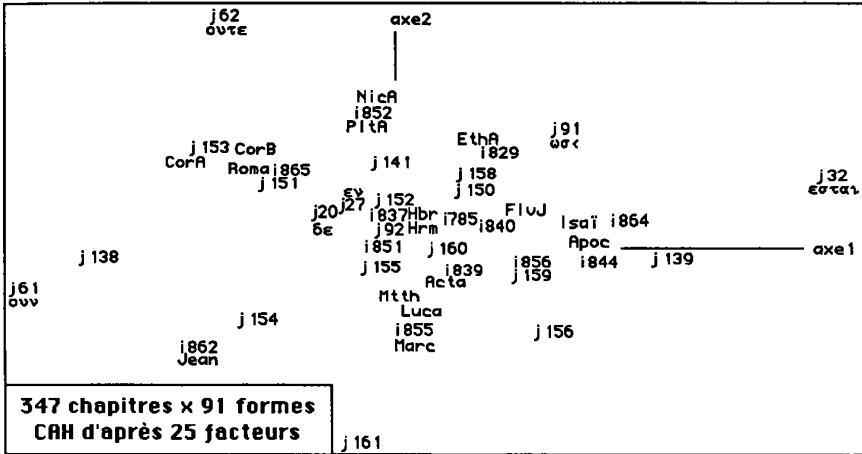
Sur l'axe 4, s'étalent les textes relevant de ce qu'au §3.2 on a appelé, non sans réserve, le genre classique: Antiquités de Flavius Josèphe (ΦΛΑ), Éthiopiens (ΑΙΘ) et textes retenus d'Aristote (ΠΟΛ, ΝΙΚ); ceux-ci associés aux classes de formes 154 et 136 (η: ou; αv: particule conditionnelle; τv: pronom neutre, nominatif ou accusatif, souvent équivalent à notre article indéfini 'un').

La mise en page offre, d'autre part, l'occasion de constater une étroite similitude (au signe près) entre les axes 1 issus des analyses des §§4.1 & 4.2.

437 chapitres x 91 mots de d4; (cl de mots)x(cl de cap); CAH 25facteurs

	Isaï	Apoc	isaï	herm	Herm	syno	Acta	SY+a	Jean	Épit	Flav	Éthi	Aris
13	1856	1844	1864	1785	1837	1851	1839	1855	1862	1865	1840	1829	1852
j150	552	165	286	37	207	304	282	662	264	761	304	183	361
j159	2442	833	1648	130	475	922	1150	3911	1114	2340	794	643	1025
j156	427	182	190	49	67	178	417	936	212	326	195	87	113
j92	17	24	14	50	10	18	29	59	23	54	3	0	11
j155	434	77	181	18	111	159	419	864	525	661	304	267	302
j20	151	20	55	41	82	232	384	1027	205	756	367	243	368
j141	367	119	291	39	109	194	237	646	271	1012	225	197	411
j160	1145	459	667	85	240	629	846	1934	731	2211	872	320	283
j158	670	347	394	65	238	352	525	1162	392	1451	801	528	662
j138	75	26	18	3	43	49	35	186	205	432	12	2	7
j27	425	115	232	25	65	231	211	552	274	1148	118	43	121
j139	109	13	135	5	49	43	21	130	44	46	49	19	32
j91	142	51	133	2	22	29	43	97	31	219	49	90	47
j32	40	3	132	0	1	46	2	30	3	16	2	3	15
j152	301	194	216	12	77	125	85	358	310	525	100	60	204
j151	398	48	245	25	99	363	113	496	351	946	129	78	173
j62	8	7	8	0	0	10	14	6	9	34	3	14	16
j153	253	45	114	16	110	289	193	766	389	1144	167	206	622
j154	922	217	414	43	206	607	435	1992	1571	1659	257	137	350
j61	9	2	1	8	58	36	48	64	199	128	16	12	64
j161	476	146	123	7	75	284	270	1620	726	435	305	71	53

ci-dessus: tableau de contingence; ci-contre (à droite) pourcent par lignes



4.2 Croisement entre 347 chapitres et 91 formes

Nous considérons simultanément les centres de 13 classes de chapitres et de 21 classes de formes construites par CAH dans l'espace engendré par les axes 1 à 25, et les centres, placés en éléments supplémentaires, des 16 œuvres principales retenues au §§3.1 & 4.1. On a attribué aux classes de chapitres des sigles auxiliaires évoquant les noms des œuvres auxquelles elles s'assimilent plus ou moins étroitement: 'Isaï' et 'isaï' pour les classes i856 et i854 composées

437 chapitres x 91 mots de d4; (cl de mots)x(cl de cap); CAH 25facteurs

	Isaï	Apoc	isaï	herm	Herm	syno	Acta	SY+a	Jean	Épit	Flav	Éthi	Aris
13	i856	i844	i864	i785	i837	i851	i839	i855	i862	i865	i840	i829	i852
150%	13	4	7	1	5	7	6	15	6	17	7	4	8
159%	14	5	9	1	3	5	7	22	6	13	5	4	6
156%	13	5	6	1	2	5	12	28	6	10	6	3	3
92%	5	8	4	16	3	6	9	19	7	17	1	0	4
155%	10	2	4	0	3	4	10	20	12	15	7	6	7
20%	4	1	1	1	2	6	10	26	5	19	9	6	9
141%	9	3	7	1	3	5	6	16	7	25	5	5	10
160%	11	4	6	1	2	6	8	19	7	21	8	3	3
158%	9	5	5	1	3	5	7	15	5	19	11	7	9
138%	7	2	2	0	4	4	3	17	19	40	1	0	1
27%	12	3	7	1	2	6	6	16	8	32	3	1	3
139%	16	2	19	1	7	6	3	19	6	7	7	3	5
91%	15	5	14	0	2	3	5	10	3	23	5	9	5
32%	14	1	45	0	0	16	1	10	1	5	1	1	5
152%	12	8	8	0	3	5	3	14	12	20	4	2	8
151%	11	1	7	1	3	10	3	14	10	27	4	2	5
62%	6	5	6	0	0	8	11	5	7	26	2	11	12
153%	6	1	3	0	3	7	4	18	9	27	4	5	14
154%	10	2	5	0	2	7	5	23	18	19	3	2	4
61%	1	0	0	1	9	6	7	10	31	20	2	2	10
161%	10	3	3	0	2	6	6	35	16	9	7	2	1
tot%	11	4	6	1	3	6	7	20	9	19	6	4	6

Isai		866	159++	20--
Apoc	152++			
isaï	139+	91+	32+++	
herm	156++	92++++	20+	161-
Herm	150++	139+	61++	
syno	156++	32+	71+++	154+
Acta	156+	155+	20+	
SY+a	154+	161++		
Jean	138+	158+++	61++	161++
Épit	138+	27++	71+	153+
Flav	160+	158+++		869
Éthp	155+	91+		20+
Aris	141+	160-	153+++	161-

on a, dans le plan (1,2) du §4.1, une classe de formes, qui se trouve (fortuitement) avoir le même numéro; et, avec un contenu voisin, caractérise {CorA, CorB} (lesquelles portent l'étiquette: 153++); et est aussi caractérisée

ci-dessous: Partition des formes en 21 classes, étiquetées d'après 13 classes

150	ακ τα ταυτά τασ τοισ ακ ταισ οτανκ	ERM+
159	καρ ετι εωσκ ουκ απο μετα οικ	HΣΑ+
156	αυτων τουσ ησαν συν αυτοισ αυτουσ	ΠΡΑΣ++ ΣΥΝΠ++
92	ουτοικ εισι	ΑΠΟ++ Ερμ++++
155	προσ νυν ουκ τον περι	ΑΙΟ+ ΠΡΑΣ+ Συν- ΙΩ+
20	δε ΗΣΑ-- Ησα-- ΑΡΙΣ+ ΑΙΟ+ ΠΡΑΣ+ ΣΥΝΠ+ ΦΛΑ+	
141	τουτο δια αυτο το	ΠΡΟΣ++ ΑΡΙΣ++
160	εισ τη, επ εψ επι οσκ τω, του	ΦΛΑ+
158	κατα τησ των υποκ παρα την ΑΙΟ+	ΦΛΑ+++
138	αυτοι ουχι καθωσ ινακ	ΠΡΟΣ++++ Ησα- ΙΩ++
27	εν	ΠΡΟΣ++++
139	αυτην αυτη, αυτησ	ΗΣΑ+ Ησα+++ ΕΡΜ+
91	ωσκ	ΗΣΑ+ Ησα++ ΑΙΟ++
32	εσται	ΗΣΑ+ Ησα++++ Σύν++
152	ενωπιον εκ ηκ	ΙΩ+ ΑΠΟ++
151	εαν ουδε ου ουτωσκ μη	Σύν++ ΠΡΟΣ++
62	ουτε ΕΡΜ≠0 ΣΥΝΠ-- ΠΡΑΣ+ ΠΡΟΣ++ ΑΡΙΣ+ ΑΙΟ+++	
153	αλλ αλλα τισ ει αν εξ πωσ τι η	ΠΡΟΣ++ ΑΡΙΣ++++
154	μετ οπου οτεκ παλιν οκ εστι οτικ ουκ απ αυτουσ	ΙΩ++
61	ουν ΗΣΑ-- Ησα-- ΑΡΙΣ+ ΕΡΜ+	ΙΩ+++++
161	αυτου ην αυτω, ουτοσκ αυτου	ΣΥΝΠ+++ ΙΩ++

par ces deux épîtres (ΚΟΡαβ+): les deux classes de mots ont, en effet, en commun les formes {αλλ, αλλα: mais; πως: comment; ει: si; τισ: équivalent masculin ou féminin du neutre τι, donné ci-dessus comme une sorte d'article indéfini}; ces formes étant liées non seulement aux épîtres mais aussi à Aristote (ΑΡΙΣ), particulièrement à l'Éthique à Nicomaque (ΝΙΚ).

On remarquera aussi d'intéressantes nuances. Dans le tableau des classes de formes du §4.2, la classe j32 {εσται: il sera, il y aura...} est notée comme fortement associée à la classe i864 (Ησα++++), dont on a noté au §3.2 (*in fine*) qu'elle comprend 23 des 35 premiers chapitres du livre d'Isaïe sans aucun des chapitres historiques 36 à 40; le lien entre j32 et le reste d'Isaïe, i856 (ΗΣΑ+), est moins net. Le tableau de contingence et le tableau de pourcentages, présentés face à face au début du présent §, permettent de chiffrer ces différences. On lit sur la ligne 'tot%', que i856 et i864 contiennent, respectivement, 11% et 6% du total des occurrences de mots outil retenues; mais la ligne '32%' (qui se réfère à la classe j32) porte que, pour les seules occurrences de εσται, ces pourcentages sont de 14% et 45%. Les listages d'aides à l'interprétation, produits avec la CAH et non publiés ici, permettent de multiplier ce type de recherches précises.

Prenons un autre exemple simple: la particule δε (j20). Dans la langue classique, cette particule s'oppose à μεν comme 'd'une part' s'oppose à 'd'autre part'. Mais dans le nouveau testament μεν est si peu fréquent qu'on l'a éliminé des analyses rapportées ici (de même que γαρ, 'car'; ειναι, infinitif du verbe être, tous deux trop étroitement dépendants du mode d'exposition d'Aristote...); en l'absence de μεν, δε perd sa valeur d'outil dialectique; comme 'donc' perd sa valeur d'outil logique dans le parler familier français.

En comparant les lignes 20% et tot%, on voit que δε est relativement fréquent dans les textes classiques {Flav, Éthi, Aris} (pour lesquels la part de δε qu'ils contiennent dépasse de moitié la part qu'il représentent de l'ensemble des occurrences) ainsi que pour Acta (nous avons d'ailleurs dénombré, dans le livre des Actes, 48 occurrences de μεν) et SY+a (classe i855, qui contient, avec 7 chapitres des Actes, 46 chapitres des Synoptiques; et, notamment, 15 des 16 chapitres que compte l'Évangile selon Saint Marc); tandis que δε est très rare chez Isaïe, dans l'Apocalypse et aussi dans l'Évangile selon Saint Jean. L'étiquetage choisi pour j20 (δε) signale: ΑΡΙΣ+, ΦΛΑ+, ΑΙΘ+ (les Éthiopiens), ΠΡΑΞ+ (πραξις, les Actes), ΣΥΝΠ+ (classe i855), ΗΣΑ--, Ησα-- (déficit dans les deux classes de chapitres d'Isaïe).

Au §4.2, la plupart des prépositions sont dans la branche j169, particulièrement dans ses subdivisions j160 et j158: {δια εις επι επι εν κατα υπο: παρα}; mais font exception {απο μετα}, qui sont dans j159 (ΗΣΑ+) tandis que leurs formes élidées {ατ μετ} sont dans j154. On a également des groupements non fortuits de formes fléchies d'un même mot outil.

5 Conclusions pour l'étude des textes

Dans [NOUV. TEST. GREC], nous avons déjà noté l'étroite similitude qui unit la 1-ère Épître de Saint Jean au quatrième Évangile; et signalé la distance qui sépare ces deux textes de l'Apocalypse. Nous avons maintenant trouvé dans la version des Septante du livre d'Isaïe un parallèle à la langue de l'Apocalypse; mais l'Évangile de Jean reste seul...

L'éventuelle existence d'un original sémitique pour l'Épître aux Hébreux avait retenu notre attention. On a vu, au §4.2 que cette épître se montre en rapport à la fois avec les Actes, l'ensemble des autres épîtres et certains textes classiques: ceci confirme que l'Épître aux Hébreux a un caractère linguistique particulier, sans infirmer aucune des hypothèses proposées depuis des siècles pour expliquer ce caractère.

L'unité de style des trois Évangiles ne fait aucun doute, même si Marc a certains traits qui lui sont propres (cf. §3.2: classe i855); indépendamment de l'emploi très fréquent de la locution $\kappa\alpha\lambda\ \epsilon\upsilon\theta\upsilon\varsigma$, "et aussitôt", que nous n'avons pas prise en compte; et dont l'insertion ne modifie en rien le cours du texte.

On pourra, en revanche être surpris de la facilité avec laquelle la classification ascendante hiérarchique retrouve une parenté de langue entre des épîtres dont la diversité est réelle.

L'unité de la langue classique vis-à-vis de celle du Nouveau Testament ou des Septante apparaît, bien que nos témoins classiques, d'ARISTOTE à HÉLIODORE en passant par FLAVIUS JOSÈPHE, couvrent plus de cinq siècles et soient d'une grande diversité. Nous en sommes d'autant plus vivement incité à poursuivre des recherches sur l'ensemble des textes classiques; de ceux du moins qui sont compris dans une même tradition dialectale.

Le succès de la CAH et de l'analyse discriminante nous assure que, même si le dénombrement des formes de mots outil n'épuise aucunement la description formelle d'une manière d'écrire, il offre d'ores et déjà, à l'analyse multidimensionnelle, un moyen praticable de déceler la signature stylistique de textes grecs.

Références et sources

DICTIONNAIRE des ŒUVRES de tous les pays et de tous les temps : Ouvrage collectif; LAFFONT - BOMPIANI, éditeur; réédition "BOUQUINS", 1981.

ARISTOTELIS : *Ethica Nicomachea*; recognovit brevique adnotatione critica instruxit L. BYWATER; Oxford University Press; édité en 1898, réimpression de 1979.

ARISTOTELIS : *Politica*; edidit W. D. ROSS; Oxford University Press; 1967.

FLAVII JOSEPHI : *Opera*, græce et latine; recognovit Guilelmus DINDORFIUS; Firmin-Didot, Paris; préface datée de 1845, et impression de 1929.

HÉLIODORE : *Les Éthiopiennes (Théagène et Chariclée)*; texte établi par R.M. RATTENBURY, Rev. T.W. LUMB et traduit par J. MAILLON; Les Belles Lettres, Paris, 1935.

HERMAS : *Le Pasteur*, introduction, texte critique, traduction et notes de R. JOLY; Sources Chrétiennes, N° 53; Les Éditions du Cerf, Paris, 1968.

NOVUM TESTAMENTUM : apparatus critico instructum edidit Augustinus MERK; 8-ème édition, Institut Biblique Pontifical, Rome, 1957.

LA SAINTE BIBLE polyglotte : contenant le texte hébreux original, le texte grec des Septante, le texte latin de la Vulgate, et la traduction de M. l'abbé GLAIRE; avec les différences de l'hébreux des Septante et de la Vulgate; des introductions, des notes, des cartes et des illustrations; par F. VIGOUROUX, prêtre de Saint-Sulpice; in Vol. V, pp. 242-501: Isaïe; in Vol. VI, pp.617-631: Malachie; A. & R. ROGER et F. CHERNOVIZ, Paris, 1900-1909.